EFFECTS OF THAI HEALTHCARE POLICY ON HOUSEHOLD DEMAND,

HOSPITAL EFFICIENCY AND HOUSEHOLD EARNINGS

By

RAJITKANOK PUENPATOM

A dissertation submitted in partial fulfillment of
the requirements for the degree of

DOCTOR OF PHILOSOPHY

WASHINGTON STATE UNIVERSITY
School of Economic Sciences

DECEMBER 2006

To the Faculty of Washington State University:

The members of the Committee appointed to examine the dissertation of RAJITKANOK PUENPATOM find it satisfactory and recommend that it be accepted.

____

Chair

____

____

ACKNOWLEDGMENTS

I remain deeply indebted to my major academic advisor, Dr. Robert Rosenman, for his unconditional and infinite support and dedication throughout my Ph.D. study. His comments, criticism, and numerous discussions have always been an intellectual learning experience. This dissertation would not have been successful without his support. I am also very grateful to Dr. Ron Mittelhammer for his constructive comments and invaluable advice that substantially improved my analyses. My appreciation is also extended to Dr. Fred Inaba, who has constantly provided recommendations and insights.

I wish to thank all graduate students in the School of Economic Sciences and all friends in Pullman who have made my time at WSU enjoyable. A special thank you to my best friend, Kelley Cullen, and her family for their energetic help and tireless support. My warm thanks go to Cris Gonzalez and Christina Quinn for their endless friendship. I also would like to express my gratitude to every SES staff for their assistance and encouragement.

My deepest appreciation goes to my husband and colleague, Tosmai Puenpatom, and my son, Pat, who was born when I was in the PhD program, for giving me courage and love throughout my academic career. I also owe a debt of gratitude to my parents, Kasem and Chintana Chitmunchaitham, and my in-laws, Sophaphan and Mesayon Puenpatom, for their constant love and encouragement. This dissertation is dedicated to them.

EFFECTS OF THAI HEALTHCARE POLICY ON HOUSEHOLD DEMAND,

HOSPITAL EFFICIENCY AND HOUSEHOLD EARNINGS

Abstract

by Rajitkanok Puenpatom, PhD
Washington State University
December 2006

Chair: Robert E. Rosenman

This dissertation evaluates how healthcare and the health insurance policy in Thailand affect both supply and demand sides. This research applies different econometrics methods such as a nonparametric boostrapping Data Envelopment Analysis method, a Quadratic Almost Ideal Demand (QUAIDS) model, a random-effect Tobit regression, etc. The first chapter investigates the short-term impact of the new national health insurance program or Universal Coverage (UC) in Thailand on technical efficiency in provincial public hospitals. By measuring efficiency before and after the reform, the study applies a two-stage analysis with bootstrapping Data Envelopment Analysis (DEA) and a censored Tobit model. The results indicated that UC improved efficiency in larger public hospitals across the country. The effect differed by region around the country, and hospitals in provinces with more wealth not only started with greater efficiency, but improved their relative position after UC was implemented.

The second chapter investigates the effects of health status and healthcare utilization on agricultural household earnings in Thailand. A utility-maximization production model in which health status and education affect household resource allocation is formulated. Using the Box-Cox transformation, the 2SLS and OLS

estimations are applied. A key finding indicated that education appeared to increase farm household income, while the effect from health was unclear. However, the disaggregated analysis showed that health may be a determinant to income for rice farming household in which farm income appeared to increase by 0.3% with a 10% increase in health investment.

The third chapter explores how national health insurance affects the allocation of household expenditures on consumption goods (i.e., housing, food, etc.) changed by comparing expenditure patterns before and after the health insurance reform. A Quadratic Almost Ideal Demand System (QUAIDS) model developed by Banks, Blundell, and Lewbel (1997) is used incorporating with a two-step approach introduced by Shonkwiler and Yen (1999). The programming was done on GAUSS 7.0 in order to solve the nonlinear least squares problems applying the Gauss-Newton optimization algorithm.

TABLE OF CONTENTS

APPENDIX

LIST OF TABLES

FIGURES

CHAPTER 1

EFFICIENCY OF THAI PROVINCIAL PUBLIC HOSPITALS AFTER THE INTRODUCTION OF NATIONAL HEALTH INSURANCE

Abstract

This paper investigates the short-term impact of the new Universal Health Coverage (UC) program in Thailand on technical efficiency in larger public hospitals. This paper measures efficiency before and after universal coverage using a two-stage analysis with bootstrapping Data Envelopment Analysis (DEA) and a censored Tobit model. The results indicated that universal coverage improved efficiency in larger public hospitals across the country. The effect differed by region around the country, and hospitals in provinces with more wealth not only started with greater efficiency, but improved their relative position after UC was implemented.

1. Introduction

Many international institutions, including the World Bank and the World Health Organization (WHO), have recommended that countries adopt universal health care coverage, believing that adequate health care is a basic human right. Thailand became the first developing country to introduce Universal Health Coverage (UC) in 2001. Six of 76 provinces adopted UC in April 2001, while the remaining provinces implemented UC in October of that year. During the early phase, Thailand has struggled with implementing UC. One of the primary problems is the financial stress of public hospitals due to the mostly unfunded government mandate requiring these hospitals to meet the service needs of the enrolled population.

UC has brought at least two significant changes in Thai health care system. First, public hospitals face increased demand from the 75% of the population previously not covered by any formal insurance system. The government believes this immense demand for health care can be met by increased efficiency rather than increased capacity. Second, the hospital funding system has moved from almost no capitated payments to nearly full capitation. Before 2001, the only public health insurance program using capitation was the Social Security Scheme (SSS), which covered only 9% of the population in 2000. With UC fully implemented almost 90% of the population is now covered by capitation. Since UC capitation is geographically mandated, hospitals have fixed revenues; thus, any hospital's financial viability depends on its ability to control costs.

One goal of using capitation is to provide a financial incentive for increased efficiency among public hospitals. The purpose of this study is to investigate the short-term effect of the capitated system on hospital efficiency by comparing the technical efficiencies of public hospitals before and after the transition period during which universal coverage was implemented. In addition, the paper evaluates other hospital and service area characteristics, which might help explain technical efficiency. Among the factors investigated are geographic regions, religion, and competitions from private sector hospitals.

Our analysis focuses on regional and general public hospitals outside of Bangkok. Regional and general hospitals are the main referral hubs (for more complicated discharges) from community hospitals, most of which are located in rural areas. Large hospitals require expensive high-technology-related medical services, which the capitation payments to hospitals under UC may not sufficiently cover. Early analysis indicates that large regional and general hospitals have been more significantly affected by financial pressure from the budget allocation of the national health insurance program than community hospitals (Na Ranong et al., 2002). These public hospitals, unlike private hospital, are obligated to enroll in the UC program. While private hospitals may be voluntary enroll in UC, few have chosen to do so.[1] This study excludes Bangkok from the sample because its health care market is too different from the rest of the country to treat it similarly, mostly because competition from private hospitals in Bangkok is significantly higher than those in other provinces. In fact, about 40% of all private hospitals in Thailand are in Bangkok.

A technical approach of this study is to measure efficiency using bootstrap Data Envelopment Analysis (DEA), a nonparametric approach based on linear programming, and then use statistical methods to find those hospital and community characteristics that affect hospital efficiency. This is a methodological contribution to efficiency analysis of health care institutions. Banker (1993) provides a statistical foundation for the estimates of

1 Private hospitals that have voluntary enrolled in the program accounts for less than one percent of all hospitals currently enrolled in the plan. Most that have enrolled are in Bangkok.

efficiency based on the Data Envelopment Analysis (DEA) that shows that they are biased for finite samples; thus inferences based on such estimates are unreliable. However, DEA estimates do exhibit the asymptotic property of consistency, so bootstrap methods provide one way to overcome the bias. Although a large number of studies have focused on the efficiency for various health care institutions, to our knowledge, none has as yet applied the bootstrap method. This paper provides an early study of hospital efficiency incorporating the DEA bootstrap model (Badin and Simar, 2003) into a two-stage analysis to identify sources of inefficiency.

The remainder of this paper is organized as follows. Section 2 introduces and briefly reviews the Thai health care system and its national insurance reform. The general literature on hospital efficiency measurement is reviewed in section 3. Section 4 discusses the empirical methodology and efficiency estimation. Section 5 describes the sample selection and variable measurements while section 6 discusses the analytical approach for identifying sources of inefficiency. Section 7 provides the empirical results, and a final section presents conclusions and implications.

2. Background: Healthcare and Health Insurance in Thailand

In 2000 there were 1,293 hospitals in Thailand comprised of 939 public hospitals, 9 state-enterprise hospitals, 14 municipal hospitals and 331 private hospitals. Of the 939 public hospitals and community hospitals, 92 are regional/general hospitals consisting of 25 regional hospitals, 48 large general hospitals and 19 small general hospitals.[2] Public hospitals are under the Ministry of Public Health (MOPH) and are operated as not-for-profit

[2] State enterprise hospitals are hospitals under state jurisdiction. Four of the nine such hospitals are located in Bangkok. Most municipal hospitals (11 out of 14), which are under provincial control, are located in Bangkok. Designation as a regional, large general, or small general hospital is based primarily on size.

organizations, accounting for almost 75% of the nation's hospital beds (see Table 1.1). Community hospitals services are limited to only primary care, range from less than 10 to 150 beds. They are mostly located in districts or minor-districts in rural areas. General hospitals consist of 200 to 500 beds, while regional hospitals are equipped with over 500 beds. Both regional and general hospitals provide tertiary care and primary care services. Public hospitals have been mandated by MOPH to provide medical services for the poor and those who enroll in welfare programs. Physicians in Thai public hospitals are employees of the hospital and as such are paid by the MOPH, according to budgetary structures, through the hospitals.

UC was gradually introduced starting in April 2001. It was implemented nationwide (except some areas of Bangkok) in October, and by April of 2002; all 76 Thai provinces were included. Before the introduction of the UC health insurance programs were classified into four main categories according to their target group (Table 1.2). The Civil Servant Medical Benefit Scheme (CSMBS) is a health insurance program offered as a fringe benefit to government employees, state enterprises employees, and their dependents. It covers less than 10% of the Thai population. CSMBS, which continues under UC, provides more extensive coverage than other insurance programs. It is fee-for-service plan, which reimburses public hospitals based on actual patient care, and pays a considerable share of the costs if the insured chooses to use private rather than public health care services. A second form of insurance that existed prior to UC and also still continues after its implementation is the Social Security Insurance Scheme (SSS). SSS provides insurance for private sector employees and the self-employed. It is a capitated system, covering about 13% of the population, half through private sector providers.[3]

Prior to UC there were two public insurance programs which offered limited health care coverage to those not covered under SSS or CMSB. The Public Welfare Scheme (PWS) provided free medical care for the poor, the elderly, children, and war veterans. The voluntary health card (VHC) covered people who were not eligible for PWS. In 2000, this amounted to approximately 17.5% of the population. VHC offered only limited coverage, and was seen as a temporary measure in the pre-UC period (Tangcharoensathien et al., 2002). Approximately one percent of the population purchased their own private insurance. During the last decade, the number of Thais with no insurance dropped significantly from almost 70% in 1991 to 20% in 2000. However, that still left about 12 million Thai people without any health insurance coverage until the advent of UC. To sum up, access to care depended upon the ability to pay, and most citizens were not afforded equal access, despite some inadequate welfare programs.

The reform combined the PWS and VHC with the uninsured into the UC program and improved. It now accounts for more than 75% of the population. Thus, after 2001, the three public health insurance programs; CSMBS, SSS and UC, provide health coverage to almost the entire population (Table 1.2). Donaldson et al. (1999) and Suraratdecha et al (2005) provide a good review of the Thai health insurance programs and benefits.

Under the new reform, UC employs a fixed capitation payment that is financed by general taxes and a co-pay of 30 baht ($0.75) per hospital visit regardless of actual expenses.[4]

---

[3] There were 137 public providers and 132 private providers in SSS in 2003.

The capitation payment covers a wide-range of benefits packages, including most ambulatory and hospital care, and preventive care and promotion except cosmetic care, obstetric delivery beyond two pregnancies, organ transplantation, infertility treatment, and other high cost interventions. Two differences between the UC and the PWS are that PWS benefit coverage is limited comparing to UC and PWS reimbursed designated public health providers based on a fee-for-service basis, not by capitation.

3. Literature Review

3.1. Capitation and Efficiency

Efficiency in general is defined as the absence of waste. An efficient unit utilizes all of its available inputs and produces the maximum amount of output, given present technological knowledge. Equivalently, the Pareto-Koopmans notion of efficiency states that a decrease in any input must require an increase in at least one other input or a reduction in at least one output (Koopmans, 1951).

Although policy makers have often used capitation in an attempt to improve efficiency in medical care delivery, the literature on the effect of capitation is inconclusive. Chu et al. (2004), using the data from California hospitals, found that less efficient hospitals are more likely to be in capitated contracts. Conrad et al. (1996) studied that impact of individual dimension of hospitals' managed care strategies on hospital efficiency using the cost per discharges as a dependent variable. They found that the proportion of hospital revenues that came from capitation payments was negatively correlated with costs per

---

[4] A capitation payment is made to every hospital depending on the UC population. In 2002, the capitation payment was 1,204.30 baht (approximately $30 per person). The capitation rate has been increased slightly thereafter.

hospital discharge. Heflinger and Northrup (2000), exploring a children's mental health services project, found decreases in access to services and the length of stay because of the capitated contract, but they concluded that the overall effect of capitation funding was unclear.

Worthington (1999) argued that public hospitals may be relatively inefficient because of governmental budgetary constraints; thus the ability of public hospitals to provide an acceptable service depends mainly on the level of funding and the extent of pressures on health care spending, which would argue for increased efficiency if capitation is low. In that regard, Barnum and Kutzin (1993) suggested that the capitation payment could ensure quality of care and cost containment for compulsory insurance program. Although not a study of the effect of capitation on hospital efficiency, Mills et al. (2000), looked at how other Thai providers responded to capitation payment, and found that some evidence of lower treatment quality. Leger (2000), applying a game theory model, indicated that capitation encourages the under-provision of medical care.

3.2. Efficiency Measurement in Healthcare

Two different techniques have primarily been used to measure efficiency of healthcare institutions; stochastic frontier analysis (SFA) and data envelopment analysis (DEA). SFA is a parametric regression based approach, while DEA is nonparametric – thus avoiding the need to specify a functional form and make distributional assumptions regarding residuals in the regression analysis. DEA readily incorporates multiple outputs, so it is particularly useful for measuring efficiency for hospitals which usually have multiple outputs and multiple inputs, and can calculate both technical and scale efficiency using only information on output and input quantities (Abbott and Doucouliagos, 2003). Moreover, DEA is likely to be more appropriate than stochastic frontiers in the non-profit service sectors where prices are difficult to define (Coelli, Prasada Rao, and Battese, 1998).

Thus, DEA has frequently been used to measure efficiency in studies of health care organizations. Valdmanis et al. (2004), Abbott and Doucouliagos (2003), Chang (1998), Rosenman et al. (1997) among others, have used DEA in recent studies of hospital efficiency in industrial countries and developing countries. In a particularly relevant application, Valdmanis et al. (2004) used DEA to investigate the performance of 68 Thai public hospitals in 1999 on the care of poor and non-poor patients. She found that all types of patients are treated equally. Chilingerian and Sherman (2004) provide a comprehensive review of health care applications using DEA.

Most studies of efficiency in health care organizations using DEA have applied a two-stage approach. Efficiency is estimated in the first stage using DEA. Then, in the second stage, the efficiency estimates obtained from the first stage are used as a dependent variable in a regression equation (usually a censored Tobit) to identify environmental variables which affect efficiency (Chilingerian (1995), Grootendorst (1997), Kirjavainen et al. (1998), Hamilton (1999), Worthington (2001), Wang et al. (2003), Scheraga (2004) among others). However, as noted earlier, Banker (1993) showed that statistically analyzing DEA estimates is appropriate only asymptotically (see also Desli and Ray, 2004). In addition, Simar and Wilson (2003) indicated that DEA efficiency estimates are serially correlated, thus using conventional DEA in the two-stage approach is invalid. Therefore, statistical inference and hypothesis tests cannot be conducted directly with the estimated efficiency scores. However, bootstrap methods may be used to resolve these problems. Although several studies including Xue and Harker (1999) and Ferrier and Hirschberg (1997), apply a naive bootstrap method based on resampling from an empirical distribution, in attempt to correct the statistical problems with DEA, the naive bootstrap method is inconsistent in the context of nonparametric efficiency estimation (Simar and Wilson ,1999a, 1999b, and 2000).

Simar and Wilson (2003), building upon earlier DEA estimation by bootstrapping by Ferrier and Hirschberg (1997), and Simar and Wilson (1998, 2000), among others, suggest a bootstrap DEA method for inference and hypothesis testing in the case of DEA estimators with multiple inputs or outputs. Recently, Badin and Simar (2003) propose a simple way of bootstrap DEA to construct confidence intervals for the efficiency scores. The Monte Carlo experiments, estimating the coverage probabilities of the estimated confidence intervals, confirm that the coverage probabilities are as good as those reported in the Monte-Carlo experiments for the full bootstrap approach (see Simar and Wilson, 2004).

4. Efficiency measurement

4.1 The basic concept of efficiency[5]

The concept of Technical efficiency can be shown conceptually using a simple example of a two-input production process in Figure 1.1. The isoquant shows the technically efficient hospital service levels associated with each combination of inputs. Technical efficiency compares how actual output compares to the ideal or the best production of this isoquant. Thus, technically efficient production assumes that reducing the use of one type of input without adding more of another type would result in reduced output. If a hospital uses the combination of inputs (*K* and *L*) indicated by point *Y* to produce the level of services associated with the shown isoquant, it is using more inputs than is technically needed.

5 This description and the following technical section is based on Rosenman and Friesner (2004), which borrowed from Coelli, Prasada, and Battese (1998).

Technical efficiency (TE) is measured by the ratio *0X/0Y*. When TE is equal to one, a firm's actual production point lies on the frontier, which is efficient. If it lies below the frontier then it is technically inefficient.

A related issue is scale efficiency. The output frontier from a single input production function provides the easiest insight into the calculation of scale efficiencies. Figure 1.2 shows a production function where some single input produces an output generically called hospital services. Two production frontiers are shown, one assuming constant returns to scale (labeled "CRS Frontier") and one assuming variable returns to scale (labeled "VRS Frontier"). Scale efficiencies are found by comparing efficiency on the variable returns to scale frontier to efficiency on a constant returns to scale frontier. For example, if a hospital is producing at point B (output B_O with P_b physician FTEs) it is technically inefficient assuming either constant returns to scale or VRS. If there are constant returns to scale, technical efficiency is given by the ratio $TE_{CRS} = B_O B_C / B_O B$. Technical efficiency assuming variable returns to scale is measured as $TE_{VRS} = B_O B_V / B_O B$. Scale efficiency calculated as the ratio of these two measures: SE = $B_O B_C / B_O B_V = TE_{CRS} / TE_{VRS}$. Essentially, scale efficiency gives a rough comparison of the average product of the firm at *B* compared to the average product at the technically optimal point (*D*). Comparison to point *D* tells us if the firm has scale inefficiency due to being too small (in the increasing returns to scale portion of the production function, like point *B*), or too large (in the decreasing returns to scale portion of the production function, like point *C*).

4.2 Data Envelopment Analysis (DEA)

DEA is a non-parametric technique based on linear programming. It establishes an efficiency frontier by solving a series of mathematical programming problems to find the most efficient production units and measure the relative efficiency of each decision making unit (DMU). DEA originated with Farrell (1957) and was further developed by Fare and Lovell (1978), Charnes, Cooper and Rhodes (1978), Banker, Charnes and Cooper (1984) and Coelli (1996), among others. The production frontier of DMUs that are producing a given number of outputs with the fewest number of inputs is identified. Measured against this frontier, efficiency is measured from 0 to 1 (the most efficient). Input oriented technical efficiency measures how much a firm produces relative to the isoquant frontier that is possible with the inputs it has chosen to use. Output oriented efficiency measures how well the firm does in minimizing the amount of inputs it uses, again relative to the isoquant frontier, given the output is has chosen.

Data Envelopment Analysis (DEA) proceeds as follows. Let y_i be a vector of m outputs and x_i a vector of k inputs for the i^{th} firm. If we have data for n firms, then X is a $k \times n$ matrix of input data for all firms and Y is a $m \times n$ matrix of output data. The *envelope*, or efficiency frontier, is derived by solving the following constant returns to scale problem:

$$\min_{\theta,\lambda} \ \theta_i \ \text{subject to} \quad -y_i + Y\lambda \geq 0$$

$$\theta_i x_i - X\lambda \geq 0$$

$$\lambda \geq 0. \qquad (2)$$

where λ is a $n \times 1$ vector of constants and θ_i is a scalar. The value $\theta_i \leq 1$ is the technical efficiency (TE) score for the i^{th} firm with a value of 1 meaning the firm is on the frontier, thereby efficient. The problem is solved once for each firm in the sample, giving technical efficiency scores for each.

The variable returns to scale (VRS) efficiency frontier is derived by solving the following problem:

$$\min_{\theta,\lambda} \; \theta_i \;\; \text{subject to} \;\; -y_i + Y\lambda \geq 0$$

$$\theta_i x_i - X\lambda \geq 0$$

$$I_v \lambda = 1$$

$$\lambda \geq 0 \qquad (1)$$

where I_v is a $n \times 1$ vector of ones. The convexity constraint, $I_v\lambda = 1$, ensures that an inefficient firm is compared against firms of a similar size.

To find scale efficiency, one must first solve the constant returns to scale technical efficiency model (equation (1)). Any difference between the technical efficiency score calculated from the constant returns to scale model, θ_C, and the technical efficiency score from the variable returns to scale model, θ_V, shows scale inefficiency. Scale efficiency is measured by θ_C/θ_V.

Finally, returns to scale is found by running one final technical efficiency model which imposes nonincreasing returns to scale, by changing the third constraint in the variable returns to scale model to $I_v\lambda \leq 1$. If the technical efficiency score found from this problem is equal to the technical efficiency score found in the variable returns to scale model the firm is in its increasing returns to scale area of production. If the two scores are equal, but not equal to the technical efficiency score from the constant returns to scale model, then decreasing returns to scale apply. Obviously, if the technical efficiency score from the variable returns to scale model equals the score from the constant returns to scale model, constant returns to scale are in effect.

In this study we use an input oriented model because Thai public hospitals must meet the market demand given a level of inputs, especially hospital beds and medical staff, which is approved by MOPH. Lovell (1993) argues that such an input-orientated is appropriate in this situation. In addition, because it is more general, we use a VRS model, allowing variable returns to scale. The hospitals in our sample vary quite a bit in the number of authorized beds and the size of medical and other staff, as well as in output quantities. With such a variation in size, it would be inappropriate to assume constant returns to scale over the range of our data. Coelli, Prasada Rao, and Battese (1998) indicate that the VRS specification has been the most used specification in the 1990s.

4.3 DEA bootstrap estimation

Bootstrapping, developed by Efron (1979), uses computer-based simulations to obtain a sample of random variables that mimic the sampling properties of a parent population. Simar (1992) introduced a DEA bootstrap approach which was developed further by Simar and Wilson (1998 and 2000). It applies a smoothed distribution of efficiency values to generate bootstrap samples of efficiencies. Smoothing is performed by an application of a kernel estimate based on the reflection method (Silverman, 1986). Badin and Simar (2003), using the statistical model in Simar and Wilson (2000), proposed a simple bootstrap DEA to construct confidence intervals for the efficiency scores.

This paper applies the technique developed in Badin and Simar (2003).[6] First, using the input-output vectors, we construct efficiency estimates, $\hat{\theta}_i$ for each DMU_i for i =1,…,n. Second, we apply a kernel smoothing of the empirical distribution of the efficiency estimates

[6] A completed description of the algorithm can be found in Appendix 1.6.

to generate smoothed efficiencies. To obtain a consistent estimator, the choice of smoothing parameter (the bandwidth parameter) has to be chosen appropriately. For that reason, the bandwidth function rule for univariate data is recommended by Silverman (1986, eq.3.31).[7] Next pseudo-data is simulated by generating values of $\hat{\theta}$ from a smooth estimate of the continuous density of $\hat{\theta}$, noting that the $\hat{\theta}_{(n)}$=1 after eliminating all the spurious efficiency scores equal to 1 from the pseudo-sample. Third, giving *n* is the number of decision-making units (DMUs), the model estimates $f(\hat{\theta})$ from the remaining $\hat{\theta}$ and generate B samples of the boundary condition $\hat{\theta} < 1$ (of the size n-1), which is $\{\hat{\theta}_1^{*b},...,\hat{\theta}_{n-1}^{*b}\}_{b=1}^{B}$ from $f(\hat{\theta})$. For the process, GAUSS is programmed to construct bootstrapping DEA using these procedures.

5. Data and Variables

The sample, consisting of yearly observations of 92 regional and general hospitals located throughout Thailand, but outside Bangkok, is chosen for two primary reasons.[8] First, these hospitals comprise all major public hospitals in each province that provide tertiary care. They are the main referral hubs from community hospitals, thus admit more complicated discharges and provide more expensive medical services, which the capitation payments to hospitals under UC may not sufficiently cover because of limited budget allocation.[9]

---

7 The Silverman approach finds the proper bandwidth by determining the optimal tradeoff between dispersion and ranges.

8 The sample from Bangkok are excluded because its hospital market is the only one in Thailand facing significant private competition.

9 We exclude community hospitals from the sample because they are in different markets in that they offer less technological services than do the general and regional hospitals. Almost 700 hospitals are community hospitals with 70 beds or less which provide only primary care service. Regional and general hospitals tended to experience financial problem because community hospitals can off-load expensive and difficult cases onto them.

Second, public hospitals are obligated to enroll in the UC program, while private hospitals are not obligated but can voluntarily enroll, although very few choose to do so.

Data are available for the three fiscal years from October 1999 to September 2002 (two year prior to the reform and one year after the reform).[10] The primary sources of data, including financial and activities database, is the Bureau of Health Service System Development, Ministry of Public Health (MOPH). Variables collected include the number of patient visits under different health insurance plans, the number of surgeries, number of patient visits by specialties, number of hospital beds and detailed data about health care personnel. The financial database includes revenue from different sources (including UC funding), expenses, and debts.

Gross provincial product per capita (GPPCR) and number of private hospital beds were obtained from the National Economics and Social Development Board (NESDB) and MOPH, respectively. Since UC officially started on October 1, 2001, the pre-UC period is defined as the fiscal year 2000 (October 1999 to September 2000) and 2001 (October 2000 to September 2001) and the post-UC period is the 2002 fiscal year (October 2001 to September 2002).

5.1 Variables used in the DEA model

The DEA model includes five inputs and five outputs (Table 1.3). Inputs consist of four categories of labor and one category of capital. Labor is measured in Full-time equivalent (FTEs) and differentiated by primary care physicians, ancillary professional care providers (dentists and pharmacists), nurses and other personnel. Since health personnel in

---

[10] MOPH has collected financial and activity database by administrative purpose. Although the 2003 fiscal year data are available, we were not able to use it in this study since the MOPH stopped collecting number of inpatients and outpatients visits categorized by specialties.

public hospitals are paid the same across region and by tenure during the studying period, the wages of labor do not differ across regions. Capital is captured by the number of beds in each hospital.[11] Revenue generation is not part of a hospital's performance criteria, so hospitals maximize output subject to this budget constraint; thus our focus on production efficiency and input oriented DEA.

Output variables include three measuring inpatients and two measuring outpatients, adjusted for hospital-wide severity.[12] Inpatient variables, include INSUR, the number of adjusted inpatient visits in acute surgery (General surgery and Orthopedic surgery); INPRI, the number of adjusted inpatient visits in primary care, (Pediatrics, Medical, and Obstetrics and Gynecology); and INOTHER, the number of adjusted inpatient visits in others (Dental, ENT, Ophthalmology, Rehabilitation medicine, and others). The two outpatient outputs are the number of surgical outpatient visits (OUTSUR) and number of non-surgical outpatient visits (OUTNONSUR).

Table 1.4 provides sample means and standard deviation of all DEA variables by year. Appendices 1.1 and 1.2 disaggregate the sample means by type and by region. Output variables for the most part (INSUR, INPRI, INOTHER, OUTSUR, and OUTNONSUR) are the highest in 2002, and are the lowest in 2001. It is interesting that while the number of

---

[11] Management in public hospitals in Thailand is highly centralized. The government, through line-item budgets, determines the budget allocation to each hospital. Any operating budget remaining at the end of the fiscal year is surrendered back to the government so that hospitals tend to use up money at the end of each fiscal year. Thus, hospital expenses are determined by the hospital's revenue allocation from MOPH.

[12] For outputs, we need to consider severity, which may influence utilization and therefore measured efficiency. One customary approach is to adjust outputs by casemix. However, MOPH does not provide direct measures of patient severity. As an alternative we estimate overall severity within a hospital by the ratio of number of large surgeries to total surgeries, which is used to adjust all outputs. The number of adjusted inpatient visits in each group is defined as $\frac{\textit{number of large surgeries / small surgeries}}{\textit{maximum of the numerator}} \times \textit{number of inpatient visits}$. A justification for adjusting all patients with this ratio is this; hospitals that attract a larger share of more complicated (i.e. large) surgeries likely attract more severe cases of other types of patients as well.

outpatient visits dramatically increased from 225,952 to 370,325 or 64% from 2001 to 2002 immediately after the UC was introduced, the number of in-patient visits slightly decreased by 4% after the UC has been introduced. In 2002, the average size of the sample hospitals is 439 beds, ranging from 85 beds to 1,143 beds. The average number of physicians, nurses, and other personnel are 50, 404, and 78 persons respectively.

MOPH classifies hospitals by number of beds. The number of health personals and beds are positively correlated with size of hospitals. The average number of beds (year 2002) in regional hospitals, large general hospitals, and small general hospitals are 689, 395, and 221 beds respectively (see Appendix 1.1). Regional hospitals, which are the largest hospitals in Thai health care service, generated highest outputs while; small general hospitals produced the lowest amount of services. Larger hospitals are mostly located in the eastern and northeastern regions.

6. Identifying Sources of Inefficiency

6.1. A Censored Tobit regression

The second step of the analysis is to relate the inefficiency scores, acquiring from the bootstrap DEA, to a number of explanatory variables, including observed characteristics of the hospitals and environmental variables. Since efficiency scores computed from the bootstrap DEA model, are censored at zero and one, an OLS regression that assumes a normal and homoscedastic distribution of the disturbance and the dependent variable would produce biased and inconsistent parameter estimates because the expected error will not equal to zero (Maddala, 1983). Therefore, a Tobit model is more appropriate.

Tobin (1958) suggested the Tobit technique by using a left-censored variable as a convenient normalization; therefore, this paper assumes a censoring point at zero. DEA measures of technical efficiency are between 0 and 1. To move to a one-sided truncation the DEA scores were transformed with the formula $INEFF_j$ = (1/DEA)-1. Thus, a negative sign on a coefficient indicates a positive association with efficiency. The exact model specification is defined as follows (Chang ,1998; Chu et. al., 2003; and Kirjavainen and Loikkanen, 1998):

$$y_{jt}^* = \beta' x_{jt} + \varepsilon_{jt}, \qquad \text{where } j = 1,\ldots, N_t \text{ and } t = 1, 2, \text{and } 3.$$

$$y_{jt} = \begin{cases} y_{jt}^* & \text{if } y_{jt}^* > 0, \\ 0 & \text{otherwise.} \end{cases}$$

where $\varepsilon_{jt} \sim N(0,\sigma^2)$, x_{jt} and β are vectors of explanatory variables and unknown parameters, respectively, y_{jt}^* is a latent variable and y_{jt} is the observed inefficiency scores. Chilingerian (1995) points out that once DEA scores (DEA) have been transformed to the inefficiency score, the slope coefficients of Tobit are interpreted in the same way as an ordinary least squares regression.

6.2 Variables used in the regressions

Table 1.5 reports the list of all variables used in the Tobit analysis. The inefficiency score (INEFF) is employed as a dependent variable. INEFF is calculated based on the result from the bootstrap DEA according to the formula above. For explanatory variables, this paper includes a variety of hospital specific characteristics and market factors. The paper measures the mix of labor inputs by the ratio of FTE physicians to other full-time personnel (PHYRATIO), since the proper mix of inputs can affect a hospital’s efficiency. Hospital service variables include two factors.[13] First, the number of referrals (REFER) reflects the level of services and resource consumption of each hospital. Most referrals are tertiary and emergency discharges that exist when a treatment could not be managed at the lower level health center. Having higher-level technological equipment and more physicians, large regional and general hospitals are more likely to admit patients who are referred from other small hospitals. This results in an increase in output, which could improve the hospital's efficiency, especially if the fixed equipment is "lumpy". Second, the length of stay (LOS) is often used to represent the efficiency with which individual patients are treated, although there is a potential tradeoff between length of stay and quality of care (Carey, 2000).

This paper also include three external market factors as explanatory variables; a Herfindahl-Hirschman index of all public and private hospitals (HI), the number of private hospital beds in each province (PRIBED) (both help capture market competition) and the Gross Provincial Product per capita (GPPCR).[14] Public hospitals in different regions generally do not compete with each other because people tend to visit public hospitals based on their geographical areas. However, hospitals in each province encounter different levels of market concentration. HI is defined as the sum of the squares of the market shares of each individual hospital where the market share is calculated by the ratio of number of beds of hospital *i* to the total number of beds in each province.[15] Higher HI values reflect less competitive pressure. PRIBED is a proxy of market competition from private hospitals. Chirikos and Sear (1994) showed that inefficiency scores are higher in markets with more

---

13 Most full-time personnel are nurses.

14 The value of GPPCR is in a real term (1988 constant price).

15 The formula is defined as $HI = \sum_{i}^{n} \Pi^2$ where Π is the market share of a firm *i*, and n is a number of firms in that province.

intense inter-hospital competition. Our model hypothesizes that the greater the number of private hospital beds, the higher the competition from private hospitals. Then, given the more health care choices available, this would result in a reduction in number of visits to public hospitals, and hence decrease efficiency of the public hospitals.

Gross Provincial Product per capita (GPPCR) represents the population's wealth in each 75 province. Provincial wealth gives at least two impacts in hospital efficiency. First, wealth may affect people behavior in seeking health care. Before UC, people who live in less wealthier region may have avoided visiting doctors if they were able to find cheaper alternative treatments, while wealthier people may seek care from either public hospitals or private hospitals; however, since private sector services are principally located in the urban areas alternative sources of health care are not always available. Thus, hospital in wealthier provinces may experience higher efficiency because of higher number of visits comparing to poorer provinces. Second, because the Thais usually provide donations for good deeds to temples and hospitals, provincial wealth may affect hospital revenues. Although the major source of public hospital revenue is from the MOPH budget, part of the revenue is from donations of people residing in the province, which increases the hospitals' reserves. This additional financial reserve could help stabilize the hospital's financial status and loosen performance, but decrease in efficiency. The influences from provincial wealth to hospital efficiency are mixed because it affects both number of hospital visits and hospital financial ability.

The location dummy variables, categorizing the six regions in Thailand, are included to account for geographic heterogeneity. Figure 1.4 shows the map of Thailand, which comprise of northern region, northeastern region, central, east, west, and southern region.

Moreover, because Thailand is predominantly Buddhist (95%), with its 4.4% Muslim population concentrated in the southernmost provinces - Pattani, Yala, Songkla, and Narathiwat, the model includes an Islam dummy variable (ISLAM) to capture religious differences. ISLAM indicates 1 if a hospital is located in the Muslim-dominated provinces, and 0 otherwise. Note that out of 13 Southern provinces, only the four southernmost are Muslim-dominated provinces.

Since UC has changed hospital financial sources, the other two variables are included to assess UC usage. UC usage variables include the ratio of the number of UC inpatients to UC enrollees (INUC) and the ratio of the number of UC outpatients to UC enrollees (OUTUC). Before UC, public hospitals in Thailand received different levels of budget allocations depending on available resources; after UC reform, hospitals obtained capitation payments based on the population in their areas, which has been used to regulate reimbursement. Hospitals in a highly populated area tend to be more financially stable. However, studies such as Pannarunothai et al. (2004) and Na Ranong et al., (2002) argue that the capitation rate in 2002 was not adequate. Ngorsuraches and Sornlertlumvanich (2006) indicates that various managerial variables such as patients to employees ratio, service mix, and market variables were determinants of hospital loss on the first year of UC implementation. They confirm that unprofitable hospitals tended to experience higher number of the UC inpatient days and also were located in the provinces with higher proportion of UC beneficiaries. Thus, hospitals that cared for more UC patients relative to UC enrollees are more likely to have financial shortfalls that should pressure hospital performance and increase efficiency.

Table 1.6 and Appendix 1.4 provide the descriptive statistics of explanatory variables in the 92 large public hospitals used, broken down by hospital type. The average length of stay per admission decreased slightly from 4.91 in 2000 to 4.84 days in 2001, but grew to 5.06 days in 2002. Gross Provincial Product per capita increased from 44,227 baht in 2000 to 46,931 baht in 2002. The ratio of physicians to other medical staff increased slightly from 8% in 2000 to 8.2% and 8.3% in the subsequent years. Furthermore, the average number of referrals from other hospitals (REFER) grew slightly from 9,246 to 9,762 visits over the time period 2000 to 2002. The number of referrals was the highest in the northeastern region (16,300), which was 4 times higher than the central and the west, which are smaller regions. Regarding people's wealth by region, the East had the highest gross provincial product per capita, while the northeastern region was the poorest. In addition, the UC usage ratio was approximately 20% for outpatients and 4.6% for inpatients in its first year. Northeastern hospitals admitted the highest number of UC patients relative to UC enrollments (26% for outpatients and 6% for inpatients), while the central region experienced the lowest UC outpatient utilization (16%).

6.3 Including a UC variable

In this section, this paper proposes a model of the determinants of technical efficiency in Thai public hospitals over the period 2000 to 2002. We include a new variable (OUTUC) that existed after UC has implemented in 2002. Suppose that period 1 is defined as a pre-UC period (2000 and 2001), and period 2 (2002) as a post-UC period. In order to assess the effect of the UC variable whether it changed an intercept, a slope, and/or an error term, the model can be written as follows;

Period 1: $Y_1 = a + bX_{ij}^1 + e_1$ (3)

Period 2: $Y_2 = a_0 + b_0 X_{ij}^2 + cZ_i + e_2$ (4)

where X_{ij}^1 is a $n_1 \times m$ matrix of m explanatory variables of n_1 DMUs in period 1, X_{ij}^2 is a $n_2 \times m$ matrix of explanatory variables in period 2, *Z* is a $n_2 \times 1$ vector of UC variable, and e_1 and e_2 are $n_1 \times 1$ and $n_2 \times 1$ vectors of error terms of each period, respectively.

The full model, which is the unrestricted model, can be written as follows:

$$\text{Y} = (a + a^* D) + (b + b^* D) X_{ij} + cZ_i + e_1(1 - D) + e_2(D) \quad (5)$$

where $a_0 = a + a^* D$, $b_0 = b + b^* D$, D is a time dummy variable indicating 1 if after-UC period.

Five hypotheses tests are 1). $H_0 : a^* = 0$, 2). $H_0 : b_i^* = 0$, 3). H_0 : c = 0, 4). $H_0 : a^* = b_i^* = c = 0$, and 5). $H_0 : \sigma_1^2 = \sigma_2^2$. The first four hypotheses testing examine how the intercept, slope coefficients, and the new UC variable affect the structure of the model, employing a log-likelihood ratio test, which can be calculated by $-2\log\lambda$ where $\log\lambda$ is the difference between the log of likelihood function of a restricted model and an unrestricted model. Note that equation (5) is an unrestricted model. In order to test for the fifth hypothesis, a Modified Levene Test is applied to examine whether the error terms have constant variances. If we cannot reject the null hypothesis of homogeneity of variances, the variances in both equations are statistically equal.

7. Empirical results

7.1 DEA results

The DEA result indicates that UC improved efficiency across the country. Table 1.7 shows the mean of the DEA *efficiency* estimates by type of hospital during the 2000 - 2002 fiscal years. Overall, mean efficiencies in all types of hospitals slightly decreased from 0.83 in 2000 to 0.78 in 2001 immediately after the UC program was introduced, and rebounded to a higher level of efficiency in 2002 (0.86). The average efficiency score was 0.82, implying that hospitals use on average approximately 18% more inputs per unit of output than if they were all efficient. Regional hospitals, in particular, improved their efficiency the most in 2002. On average, small general hospitals were the most efficient hospitals, followed by large general hospitals and regional hospitals (0.90,0.82, and 0.75 respectively). In addition, the UC program affected six regions differently. As shown in Table 1.8, UC affected the southern region's efficiency the most (13.4%) while affected the North's the least by 5%.

Table 1.9 reports the distribution of changes in technical efficiency over time. Comparing 2000 to 2002, 50 out of 92 hospitals increased their efficiency, 14 were unchanged, and 28 decreased. But, during the transition year, from 2000 to 2001, 59 hospitals reported a decrease in efficiency while only 25 hospitals showed an increase in efficiency. After UC was fully implemented, 70% of all hospitals (65 hospitals) experienced an improvement in efficiency from 2001 to 2002, 10% reported no change, and about 20% of hospitals experienced a decrease in efficiency.

It is surprising that so many hospitals experienced a decline in efficiency from 2000 to 2001 because UC was not widely implemented until October 2001 (which is the start of the 2002 fiscal year). The implementation of UC was tied to the election victory of the Thai Rak Thai party in February 2001. The exact chronology is given in the Appendix 1.4. It is possible that people, expecting that the out of pocket cost of care would decline with the implementation of UC, delayed the use of hospitals when possible. For the outpatient visits, Table 1.10 and 1.11 show that the number of both non-surgery and surgery services in both outpatients and inpatients care decreased before the beginning of UC, but increased significantly after the reform has started. It can be seen that changes in efficiency were mostly caused from changes in output since inputs (number of health personnel) increased slightly (Table 1.12). When inputs are rather fixed, this decrease (increase) in output would decrease (increase) efficiency.

This paper employs two nonparametric tests to study for average differences in efficiency by time period. First, a Kruskal-Wallis Analysis of Variance (K-W) test is conducted on the null hypothesis that there is no median difference in technical efficiency across the three years. As shown in Table 1.13, the chi-square is 14.9, which is greater than the 0.05 level of significance, allowing us to conclude that at least one pair of the technical efficiency medians is not equal. The statistical evidence shows that technical efficiencies in Thai provincial hospitals changed after the introduction of UC. Next this paper utilizes the Mann-Whitney U test to conduct pairwise comparisons since the Kruskal-Wallis one way ANOVA test is significant (Sheskin, 1997) of year 2000 to year 2001, year 2001 to year 2002, and year 2000 to year 2002. The results reported on Table 1.14 indicate that the population medians of technical efficiency are different at all pairs; 2000 and 2001, 2001 and 2002, and 2000 and 2002. This is a key finding, which suggests hospital efficiency improved from before the introduction of UC (2000) to after its introduction (2002).

Many of the efficient hospitals, those with a technical efficiency score of one, experienced a decrease in efficiency in 2001, but regain their efficiency in 2002. As can be seen in Table 1.15, the total number of efficient hospitals was highest in 2002 (34 hospitals), compared to 13 hospitals and 22 hospitals in 2001 and 2000 respectively. Large regional hospitals were more likely to be efficient in 2002. However, some efficient large regional hospitals in 2000 experienced a decline in efficiency in 2001, although they all regained full efficiency in 2002. Appendix 1.5 shows the information in detail.

After UC, more northeastern hospitals were efficient than in other areas (over 70%, 14 out of 19). The performance of southern hospitals was the same before 2001, but improved in the post-UC period. Out of 19 hospitals, the number of efficient hospitals from the South increased from five hospitals in 2000 and 2001 to eight hospitals in 2002. All four efficient small hospitals in 2000 were the same hospitals in 2002.

7.2 Comparing DEA and bootstrap DEA results

Table 1.16 and 1.17 report the descriptive statistics for the bootstrap DEA scores for B = 1,000 replications from 2000 to 2002. The average of the bootstrap efficiency estimates was 0.76, which is lower than the average of the (0.82) original efficiency scores. Also, the minimum and the standard deviation of the original DEA estimates for each of the years are higher than the bootstrap values (except the minimum in 2000). Efron (1982) indicates that the bias of the statistic is not a serious problem when the ratio of the estimated bias to the standard error is less than 0.25. Our result shows that approximately 65% of hospital ratios are greater than 0.25, indicating a bias problem of the original scores. Therefore, the bootstrap DEA estimates are likely better indicators of hospital technical efficiency. Figure 1.3 compare the original DEA and bias-corrected bootstrap DEA efficiency scores.

7.3 The regression results

Section 6.3 formulated a model of the determinants of technical efficiency in Thai public hospitals over the pre-UC period (2000 and 2001) and the post-UC period (2002). To test whether the UC variable (OUTUC) changes the structure of the regression model, this paper performed five hypotheses tests.[16] Table 1.18 shows the Likelihood-ratio tests for parameters of the Tobit model. For the first hypothesis, testing a potential change in the intercept term over the two periods, the results show that the log likelihood ratio (3.98) is greater than a chi-square statistic with one degree of freedom ($\chi^2_{(0.95;1)} = 3.84$). Thus, we can reject the null hypothesis at a 0.05 level of significance that a coefficient of a time dummy variable *D*, a^*, is equal to zero, implying the intercepts of two periods are not equal.

For the second hypothesis, which is all slope coefficients are jointly equal to zero, the results indicate that the model is significant with a LR test of the restriction that all the slope coefficients are jointly zero rejected at a 0.05 level [LR=22.56 ~ χ^2_{12}]. This allows us to conclude that at least one pair of the slope coefficients is not equal and the result provides statistical evidence that OUTUC may cause a change in efficiency after the introduction of UC. We also reject the third hypothesis of the coefficient of UC variable, c, being equal to zero, which implies that there is statistical evidence that OUTUC affected technical efficiencies in Thai provincial hospitals. In addition, we reject the fourth null hypothesis that the intercept, slope coefficients, and the UC coefficient are zero. Furthermore, the modified Levene test (Table 1.19), examining the homogeneity of variances hypothesis, does not reject the null hypothesis that the variances of both periods are equal, implying that the inclusion of

---

16 We dropped INUC due to a high collinearity. The reason for leaving OUTUC in the model is that the number of outpatients has significantly changed during the period of study, while the number of inpatients was rather constant (refer to section 5.1).

the UC variable does not alter the error terms in both periods. To sum up, there is evidence that including the UC variable (OUTUC) changed the intercept, and no statistical evidence that it changed the error terms of the model, but the slope coefficients did change. Thus, the model used for the Tobit regression in this paper is written as follows;

$$Y = (a + a^{*}D) + (b + b^{*}D)X_{ij} + cZ_{i} + e \qquad (6)$$

This paper, then, performed a Tobit regression using bootstrap DEA scores as a dependent variable. A set of explanatory variables includes the outpatient UC usage ratio (OUTUC), physicians to other staff ratio (PHYRATIO), the length of stay (LOS), the gross provincial product per capita (GPPCR), number of referrals (REFER), the Herfindahl index (HI), number of private hospital beds of each province (PRIBED), the Islam dummy variable (ISLAM), and six region dummy variables. The estimated coefficients and standard errors of the pre-UC and post UC parameters are shown in Table 1.19. Also included in the same table are statistics for Log-likelihood ratio tests. The level of significance ($\chi^2_{(1)}$ = 18.80) indicates that a random-effect Tobit model is more appropriate than the pooled Tobit regression.

As shown in Table 1.19, the regression confirms that the introduction of UC increase hospital efficiency. The UC variable, outpatient UC usage ratio (OUTUC), has a negative and statistically significant effect on hospital technical inefficiency as expected, suggesting that an increase in number of UC patients per enrollees tends to increase efficiency.[17] After UC was implemented, those hospitals with larger UC utilization were more efficient. Because hospitals with a high percentage of UC usage tend to experience more financial problems, this result is consistent with the hypothesis that hospitals responded to the financial pressures associated with higher UC utilization by increasing efficiency.

---

17 This variable (OUTUC) is zero before UC.

The other key finding about how UC affected hospital behavior shows up in the variable PHYRATIO (the percentage of physicians to other full-time personnel). PHYRATIO shows a significant positive effect in the pre-UC period. The result suggests that the hospitals that have a larger ratio of physicians to other medical staff are less efficient. Thailand has experienced a shortage in medical professionals such as physicians, dentists, and nurses for years especially in the northeast and the northern regions. The regression result implies that given a fixed number of physicians, an increase in the number of other medical professionals, especially nurses, improves efficiency. Because D*PHYRATIO has a significant negative effect in the post-UC period, we conduct the Log-likelihood ratio test (LR) to examine whether the sum of both period coefficients is significantly different from zero (Table 1.18). With the LR test, we cannot reject the hypothesis that the sum of PHYRATIO's and D*PHYRATIO's coefficients is zero at a 0.05 level [LR= 1.1 ~ $\chi^2_{(1)}$], indicating that the effect of PHYRATIO (lowering efficiency) disappeared after the implementation of UC. Such change may cause from a fact that the physician shortage problem in public sector has become more severe after UC since some physicians have switched to the private sector. With more data, a further investigation on the effect of medical personnel ratio on efficiency is desirable.

The base equation provides more general information about what improved hospital efficiency in Thailand, both before and after the introduction of UC. The results show that number of referrals (REFER) is positively related and statistically significant with the level of efficiency in Thai public hospitals while the effect of D*REFER is not significant. The result suggests that large hospitals, which admit more tertiary cases, are likely to be more efficient in both pre-UC and post-UC periods and this efficiency does not change with introduction of UC. The result is consistent with the hypothesis that an increase in the number of referrals enhances output and efficiency. Furthermore, the regression results indicate that the length of stay (LOS) is positive and statistically significant, which conforms to the a priori hypothesized signs. The shorter length of stay appears to improve the level of efficiency. With UC, the effect of LOS on technical efficiency does not change. The regressions also show that the number of beds (BED) is not statistically significant showing that the size of hospitals does not determine hospital efficiency. This may appear to be at odds with our earlier finding that regional hospitals were less efficient than large and small general hospitals. However, the dependent variable in this regression measures the marginal effect with number of beds, while in the earlier analysis we measured the average effect.

Although gross provincial product per capita (GPPCR) and D*GPPCR are not statistically different from zero at a 0.1 level of significant, they are statistically significant at a 0.15 and 0.12 level of significant respectively. Although the coefficient is small, the negative coefficients show that hospitals located in wealthy areas are more likely to be efficient. The result shows that the provincial wealth factor is positively correlated to efficiency in both periods, and the impact was stronger after UC was implemented.[18] Patients from more affluent areas on average have a greater ability to pay for hospital services than patients from poorer areas. In fact, one conjectures that people in lower income provinces tend to prefer self-care; using over-the-counter drugs or obtaining traditional care, because of the lower cost of these alternative treatments. Therefore, in the pre-UC, hospitals in poorer areas tended to admit fewer patients, which lower hospital efficiency when controlling for inputs.

[18] We reject the hypothesis that the sum of GPPCR's and D*GPPCR's coefficients is zero at a 0.05 level

Most regional dummy variables were statistically significant, indicating general patterns of efficiency by geographical location. Compared with the (excluded) southern region, in the pre-UC period hospitals from the northeastern region were the most efficient, followed by the southern region. Hospitals in the west were the least efficient. The Northeast is the poorest region where people in the region have the highest ratio of population per physician, nurse and hospital than the rest of the country. Ngorsuraches and Sornlertlumvanich (2006) report that 30% of unprofitable hospitals after UC was implemented were in northeastern. An increase in outputs as well as the higher pressure from financial difficulty induced efficiency. In contrast, it appears that there are more public hospitals in western region relative to other regions; of the 76 provinces, four out of 92 hospitals are located in Rachaburi, which is a small affluent province in the western region. It is possible that some hospital resources may not have been utilized efficiently due to the small amount of services. After UC was introduced, only the East dummy variable is statistically significant at a 0.05 level of significance, indicating that only in that region, on average, did hospital efficiency change (becoming less efficient) relative to the south. The Islam dummy variable (ISLAM) is not statistically significant in either period showing that hospital efficiency is unaffected by religious composition of population. .

Neither competition variable (PRIBED and HI) had a coefficient that statistically differed from zero. Although the model hypothesizes that an increase in private hospitals may improve public hospital efficiency because patients have more choices to choose where they visit, the empirical results do not support that idea. The insignificance may imply that private and public hospitals in Thailand serve different markets. Competition among private and public hospitals is limited because private hospitals tend to focus on upper-middle income to high-income market because they usually provide more courteous and luxurious services while public hospitals provide care at a lower cost for a lower income group.

8. Summary and Conclusions

This paper investigates the short-term effect of the new national health insurance known as Universal Coverage on hospital efficiency by comparing the technical efficiencies of public hospitals before and after the transition period during which universal coverage was implemented. This paper studied the efficiency differences among 92 Thai provincial public hospitals using a two-stage analysis, including the Data Envelopment Analysis, bootstrapping DEA, and a censored Tobit model.

In all, the DEA results indicate that UC improved efficiency across the country. Regional hospitals, in particular, improved their efficiency the most. On average, small general hospitals were the most efficient hospitals, followed by large general hospitals and regional hospitals. Comparing the original DEA and the bootstrapping DEA, the result confirms that the bootstrap DEA estimate is the better indication of hospital technical efficiency because of its unbiasedness and consistency.

The Tobit regression shows that the reform is a source of efficiency, which is consistent with the DEA results. Because access of care, especially by those with lower incomes and the uninsured improved, an increase in the number of UC patients per enrollees increased hospital efficiency. This also implies that the capitation budget system which has replaced the incremental financing supply-sided cost, improved efficiency. Considering hospital input allocations, the results showed that the physicians to other medical staff ratio hurt efficiency before UC, after UC no such effect was evident. An increase in health professional school’s capacity may help lessening the shortage of medical personnel in the public health sector.

The results show with marginal significance that provinces with more wealth were more efficient relative to those in less wealthy areas. The impact of provincial wealth on efficiency became stronger after UC started in that hospitals in wealthier provinces tended to be more efficient. Considering the effect of number of referrals, the results indicate that more referrals improve efficiency in regional and general hospitals, which may be an area for further research. In terms of market competition, the Herfindahl index and the presence of private hospitals in the local market, representing degree of competition, do not affect technical efficiency. Finally, the efficiency change depends on geographical locations. Hospitals in the East become the least efficient instead of hospitals in the West after the reform started.

This study shows preliminary results, analyzing only at the short-term immediate effects of UC on the efficiency of regional and general hospitals in Thailand. The program implementation is still in the transitional stage. We are able to explain some aspects of hospital efficiency that transcend UC. If the efficiency of regional and general hospitals is considered important, referrals from community hospitals should be encouraged. In addition, this paper showed regional and income differences in efficiency that may be amenable to policy interventions. Further study, after more time is available for implementation and adjustment, is needed in order to reveal the full impact of how efficiency changed with the reform.

REFERENCES

Abbott, M., Doucouliagos, C. (2003), "The efficiency of Australian universities: a data envelopment analysis", *Economic of Education Reviews* 22, 89-97

Banker R.D. (1993), "Maximum Likelihood, Consistency and Data Envelopment Analysis: A Statistical Foundation", *Management Science* 39 (10), 1265-1273.

Banker, R.D., Charnes, A., Cooper, W.W. (1984), "Some models for estimating technical and scale inefficiencies in data envelopment analysis", *Management Science* 30 (9),1078–92.

Barnum, H. and Kutzin, J. (1993), *Public Hospitals in Developing Countries: Resource Use, Cost Financing,* (The Johns Hopkins University Press. Baltimore and London).

Biorn, E. , Hagen, T. , Iversen, T., and Magnussen, J. (2003), "The Effect of activity-based financing on hospital efficiency : A panel data analysis of DEA efficiency scores 1992-2000", *Health Care Management Science* 6, 271-283.

Badin, L. and Simar, L. (2003), "Confidence Intervals for DEA-type Efficiency Scores:How to Avoid the Computational Burden of the Bootstrap", Discussion paper 0316, Universit'e Catholique de Louvain.

Carey, K. (2000), "Hospital Cost Containment and Length of Stay: An Econometric Analysis", *Southern Economic Journal* 67 (2), 363- 380.

Chang, H. (1998), "Determinants of Hospital Efficiency: the Case of Central Government-owned Hospitals in Taiwan", *Omega*, International Journal Management Science 26 (2) 307-317.

Charnes A. Cooper WW, Rhodes E. (1978), "Measuring the efficiency of decision making units", *European Journal of Operational Research* 2, 429-44.

Chilingerian, Jon A, (1995), "Evaluating physician efficiency in hospitals: A multivariate analysis of best practices", *European Journal of Operational Research* 80 , 548-574.

Chinlingerian J.A. and Sherman H.D. (2004), "Health care Applications: from Hospitals to Physicians, From Productive Efficiency to Quality Frontiers", in: *Handbook on Data Envelopment Analysis*, W.W. Cooper and J. Zhu, eds., (Kluwer Academic Publishers).

Conrad, D., Wickizer, T., Maynard, C., Klastorin, T., Lesser, D., Ross, A., et al. (1996), "Managing Care, Incentives, and Information: An Exploratory Look Inside the Black Box of Hospital Efficiency.", Health Services Research 31, 235-59

Chu, H. L, Lui, H. Z., Romeis, J.C., Yaung, C. L. (2003), "The Initial Effects of Physician Compensation Programs in Taiwan's hospitals: Implication for Staff Model HMOs", *Health Care Management Science* 6, 17-26.

Chu, H. L, Lui, H. Z., Romeis, J.C. (2004), "Does Capitated Contracting Improve Efficiency? Evidence from California Hospitals", Health Care Management Review, 29(4), 344-352.

Coelli, T. (1996). "A Guide to DEAP version 2.1 : A Data Envelopment Analysis (Computer) Program", *CEPA Working Paper* 96/08. University of New England. Australia.

Coelli, T. , Prasada Rao D.S., and Battese G.E. (1998), *An introduction to efficiency and productivity analysis*. (Kluwer Academic Publishers, Boston).

Culyer, A. and Newhouse, J. edit (2000), *Handbook of Health Economics*, 1B.

Delsi, E. and Ray, H. (2004), "A Bootstrap-Regression Procedure to Capture Unit Specific Effects in Data Envelopment Analysis", *Working Paper 2004-15,* University of Connecticut.

Dixon, J. and Harrison, A. (1997), Funding the NHS: A little local difficulty, BMJ, 314: 216.

Donaldson, D., Pannarunothai, S., and Tangcharoensathien, V. (1999), Health Financing in Thailand: Technical Report, Health Management and Financing Study Project, ADB no. 2997, Boston. (http://www.msh.org/resources/online_reports/pdf/thaihf0.PDF).

Efron, B. (1982), “The Jacknife, the Bootstrap, and Other Resampling Plans”, Monograph 38, Society for Industrial and Applied Mathematics, Philadelphia.

Efron, B. and Tibshirani, R.J. (1993), “An Introduction to the Bootstrap”, Monographs on Statistics and Applied Probability, No. 57, (Chapman and Hall, New York).

Fahrer, J. and Justin G. (2000), “Public Hospital Costs in Two Australian States”, *The Australian Economic Review* 33 (2) , 182-192.

Fare, R. and Lovell, C.A.K. (1978), “Measuring the Technical Efficiency of Production”, Journal of Economic Theory, vol. 19, no. 1, Oct. 1978, pp. 150-62

Farrell, M.J. (1957). “The measurement of productivity efficiency”, *Journal of the Royal Statistical Society* 120, 253-281.

Ferrier, G. D., and Hirschberg, J.G. (1997). “Bootstrapping Confidence Intervals for Linear Programming Efficiency Scores: With an Illustration Using Italian Bank Data.”, Journal of Productivity Analysis 8,19-33

Giokas , D. (2001), “Greek Hospitals : How well their resources are used”, *Omega, Int. Journal of Management Science* 29, 73-83.

Greene, W. (2003), *Econometric Analysis*. 5th ed. (Macmillan Publishing).

Grootendorst, P., V. (1997), “Health care policy evaluation using longitudinal insurance claims data : an application of the panel Tobit estimator”, *Health economics* 6, 365-382.

Hamilton, B., H. (1999), “HMO selection and medicare costs : Bayesian MCMC estimation of a robust panel data tobit model with survival”, *Health Economics* 8, 403-414.

Heflinger, C.A., and Northrup, D.A. (2000). “What Happens When Capitated Behavioral Health Comes to Town? The Transition from the Fort Bragg Demonstration to a Capitated Managed Behavioral Health Contract.”, *Journal of Behavioral Health Services and Research* 27, 390-405

Hollingsworth, B. (2003), “Non-parametric and parametric applications measuring efficiency in health care”, *Health Care Management Science* 6, 203-218.

Hollingsworth, B., and Wildman, J. (2003), “The efficiency of health production : re-estimation the WHO panel data using parametric and non-parametric approaches to provide additional information”, *Health Economics* 12, 493-504.

Kirjavainen , Tanja and Loikkanen, Heikki A. (1998), “Efficiency differences of Finnish senior secondary schools: An application of DEA and Tobit analysis”, *Economics of Education Review* 17 (4), 377-394.

Koopmans, T.C. (1951), “An Analogus of Production as an Efficient Combinations of Activities,” in Koopmans, T.C. (edi), Activity Analsysis of Prodcution and Allocation, Cowles Commission for Research in Economics, Monograph 13, (Wiley, New York).

Leger, P.T. (2000), “Quality Control Mechanism under Capitation for Medical Services”, *The Canadian Journal of Economics* 33 (2), 564-586

Lovell, C. A. K. (1993), Production frontiers and production efficiency in *The Measurement of Productive Efficiency: Techniques and Applications* (Eds) H. O. Fried, C. A. K. Lovel and S. S. Schmidet, (Oxford University Press, New York), 3-67.

Maddala, G.S. (1983), *Limited dependent and qualitative variables in econometrics,* (Cambridge University Press, New York).

Mill, A., Bennett, S., Siriwanarangsun P., and Tangcharoensathien V. (2000), "The response of Providers to Capitation Payment: A Case-study from Thailand", *Health Policy* 51, 163-180.

Ministry of Public Health, Thailand, (2001). Health Insurance Systems in Thailand, Health Systems Research Institute. Nonthaburi, Thailand

Na Ranong, V., et. al. (2002), *Evaluation of the Universal Coverage Scheme in Thailand FY2001,* Thailand Development Research Institute, Thailand.

Ngorsuraches, S. and Sornlertlumvanich, A. (2006), "Determinants of Hospital Loss in Thailand: Experience from the First Year of a Universal Coverage Health Insurance Program", Health Care Management Science 9, 59–70.

Pannarunothai, S., Patmasiriwat, D., Srithamrongsawat, S. (2004), "Universal health coverage in Thailand: ideas for reform and policy struggling", *Health Policy* 68 (1), 17-30

Parkin D. and Hollingsworth B. (1997), "Measuring production efficiency of acute hospitals in Scotland, 1991-94: validity issues in data envelopment analysis", *Applied Economics* 29, 1425-1433.

Pauly, M. and Redisch M. (1973), "The Not-for-profit Hospital as a Physician's Cooperative", *American Economic Review* 63, 87-100.

Rosko, M. (2001), "Cost Efficiency of US Hospitals : A Stochastic Frontier Approach", *Health Economics* 10, 539-551.

Rosenman, R., Siddharthan, K. and Ahern, M. (1997), "Output Efficiency of Health

Maintenance Organizations in Florida”, *Health Economics* 6(3), 295-302

Rosenman, R. and Friesner, D. (2004), “Scope and scale inefficiencies in physician practices”, *Health Economics* 13(11), 1091-1116

Ruggiero, J. (1996), “On the measurement of technical efficiency in the public sector”, *European Journal of Operational Research* 90, 553-565.

Scheraga, C.A. (2004), “Operational efficiency versus financial mobility in the global airline industry : a data envelopment and Tobit analysis”, *Transportation Research : Part A.* 38, 383-404.

Sheskin, D. J. (1997), *Handbook of Parametric and Nonparametric Statistical Procedures*, (CRC Press. Florida).

Simar, L., Wilson, P.W., (1998), “Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models”, *Management Science* 44(1), 49-61.

Simar, L., Wilson, P.W., (1999a), “Some Problems with the Ferrier/Hirschberg Bootstrap Idea”, *Journal of Productivity Analysis* 11,67-80.

Simar, L., Wilson, P.W., (1999b), “Of Course We Can Bootstrap DEA Scores! But Does it Mean anything? Logic Trumps Wishful Thinking”, *Journal of Productivity Analysis* 11,93-91

Simar, L., Wilson, P.W., (2000), “Statistical Inference in Nonparametric Frontier Models: The State of the Art”, *Journal of Productivity Analysis* 13, 49-78.

Simar, L., Wilson, P.W., (2003), “Estimation and Inference in Two-Stage, Semi-Parametric Models of Production Processes”, *Discussion paper 0307*, Universit’e Catholique de Louvain.

Simar, L., Wilson, P.W. (2004), “Performance of the bootstrap for DEA estimators and iterating the principle", in: *Handbook on Data Envelopment Analysis*, W.W. Cooper and J. Zhu, eds., (Kluwer Academic Publishers).

Silverman, B. W. (1986), Density Estimation for Statistics and Data Analysis, (Chapman and Hall, New York).

Steninmann, L. and Zweifel, P. (2003), "On the (in)efficiency of Swiss Hospitals", *Applied Economics* 35, 361-370.

Supachutikul, A. (1996), *Situation analysis on health insurance and future development. Bangkok*, Thailand Health Research Institute, National Health Foundation

Suraratdecha, C., Saithanu, S., and Tangcharoensathien, V. (2005), "Is the Universal Coverage a Solution for Disparities in Healthcare? Finding from three low-income provinces of Thailand", *Health Policy* 73, 272-284.

Tangcharoensathien, V. et al. (2002), "Universal Coverage and Its Impact on Reproductive Health Services in Thailand", *Reproductive Health Matters* 10 (20), 59-69.

Tobin, J. (1958), "Estimation of Relationships for Limited Dependent Variables", *Econometrica* 26 (1), 24-36.

UNFPA (2005), *Reproductive Health of Women in Thailand: Progress and Challenges towards attainment of International Development Goals*, Country Technical Services Team for East and South-East Asia Bangkok, Thailand.

Valdmanis, V., L. Kumanarayake, and J. Lertiendumrung (2004) "Capacity in Thai Public Hospitals and the Production of Care for Poor and Non-Poor Patients" *Health Services Research*, 39 (6) Part II, 2117-2134

Wang K. L., Tseng, Y.T., and Weng, C.C. (2003) , "A study of production efficiencies of integrated securities firms in Taiwan", *Applied Financial Economics* 13, 159–167

Wagstaff, A. and Barnum, H. (1992). Hospital Cost Functions for Developing Countries. Population and Human Resources. *WPS 1044*, The World Bank.

Worthington, A. (1999), "An empirical Survey of Frontier Efficiency Measurement Techniques in Healthcare Services". Working paper, School of Economics and Finance, Queensland University of Technology, Australia.

Worthington, A. (2001), "Efficiency in Pre-Merger and Post-Merger Non-Bank Financial Institutions", *Managerial and Decision Economics* 22, 439-452

Xue, M. and Harker, P.T. (1999), "Overcoming the Inherent Dependency of DEA Efficiency Scores: A Bootstrap Approach," unpublished working paper, Wharton Financial Institutions Center, University of Pennsylvania.

TABLES

Table 1.1: Hospital and Medical Establishments with beds by Type of Administration in 2000

Type of Administration	Number of hospitals	Number of beds
1. Government	939	102,122
- Ministry of Public Health	868	87,752
- Other Ministries	71	14,370
2. State Enterprise	9	2,439
3. Municipality	14	2,279
4. Private	331	29,361
Total	1,293	136,201

Source: Ministry of Public Health

Table 1.2: Health insurance coverage in Thailand 1991-2003 (%)*

	Before Oct. 2001						After Oct. 2001	
Insurance program	**1991**	**1992**	**1995**	**1997**	**1999**	**2000**	**Insurance program**	**2004**
Civil Servant Medical Benefit Scheme (CSMES)	10.2	11.3	11.0	10.8	10.8	12.0	CSMES	8
Social Security Scheme (SSS)*	3.2	4.4	7.3	7.6	9.2	9.4	SSS	13.16
Public Welfare scheme for the poor (PWS)	16.6	35.9	43.9	44.7	42.1	40.8	UC	75.24
Voluntary Health Scheme	2.9	3.9	9.8	15.3	15.8	17.5		
The uninsured	67.1	44.5	28.0	21.6	22.1	20.3	The uninsured	~4

Note : * There were approximately 62.6 millions people in Thailand in 2004.
** Excluded the Workmen Compensation Fund (WCF) and the Car-accident Compensation Scheme, which is considered supplementary schemes where funds are collected from those who are liable for the workplace or traffic accidents.

Source: Health Insurance Systems in Thailand (2001), Health Systems Research Institute, Ministry of Public Health. and The 2004 Universal Coverage Report (2004), National Health Security Office, Ministry of Public Health.

Table 1.3: Definitions of DEA variables

Variables		Definition
Output	INSUR*	Number of adjusted number of inpatient visits in acute surgical – general surgery and Orthopedic surgery
	INPRI*	Number of adjusted number of inpatient visits in primary care -- Pediatrics, Medical, and Obstetrics and Gynecology
	INOTHER*	Number of adjusted number of inpatient visits in others -- Dental, ENT, Ophthalmology, Rehabilitation medicine, and others
	OUTSUR	Number of surgical outpatient visits
	OUTNONSUR	Number of non-surgical outpatient visits
Input	BED	Number of beds
	PHYSICIAN	Number of physicians FTEs
	NURSE	Number of nurses
	DENPHAR	Number of dentists and pharmacists
	OTHERS	Number of other personnel

Note : * Adjusted inpatient variables of each group are defined as :
ratio of large surgeries to total surgeries x number of inpatients in acute surgical, or primary care or others
maximum amount of the numerator

Table 1.4: Descriptive statistics: mean (standard deviation) per year

	INSUR	INPRI	INOTHER	OUTSUR	OUTNON SUR	BED	PHYSI-CIAN	NURSE	DEN PHAR	OTHERS
2000	6,635.98 (4638.09)	15,316.20 (7233.29)	2,534.59 (3538.75)	3,200.65 (4,299.13)	218,966.4 (101,352.8)	431 (196)	43 (30)	369 (155)	21 (9)	65 (26)
2001	6,338.22 (4367.66)	14,977.59 (7088.26)	2,163.44 (2041.67)	2,986.67 (3,855.13)	222,951.9 (103,369.2)	431 (196)	46 (33)	383 (167)	22 (8)	70 (28)
2002	6,853.25 (4849.56)	15,620.97 (7277.47)	2,166.13 (2794.22)	2,880.45 (3,665.88)	370,325.0 (157,762.8)	439 (201)	50 (38)	404 (176)	24 (10)	78 (34)

Table 1.5: Definition of explanatory variables

Categories	Variables	Definition
Dependent variables	INEFF	Inefficient scores
Input mix	PHYRATIO	Ratio of FTE physicians to other full-time personnel
	BED	Number of beds
Services	REFER	Number of patients referring from the other hospitals ('000)
	LOS	Length of stay
Geographic influences	NORTH	1, if from the northern region
	NORTHEAST	1, if from the northeastern region
	CENTRAL	1, if from the central region (exclude Bangkok)
	EAST	1, if from the eastern region
	WEST	1, if from the western region
	SOUTH	1, if from the southern region
	ISLAM	1, if a hospital located in Muslim-dominated provinces
Effect from UC	INUC	Number of UC inpatients/UC enrollees of that hospital (persons)
	OUTUC	Number of UC outpatients/UC enrollees of that hospital (persons)
Market factors	PRIBED	Number of beds in private hospitals of each province
	HI	Herfindahl index
	GPPCR	Real Gross Provincial Product per capita

Table 1.6: Descriptive statistics of explanatory variables

YEAR	OUTUC	INUC	GPPCR	LOS	BED	PHYRATIO	REFER	HI	PRIBED
2000	0	0	44,226.92	4.91	430.80	0.0802	9,246.10	0.21	244.73
2001	0	0	44,253.32	4.84	430.80	0.0823	9,487.73	0.21	279.07
2002	0.205	0.046	46,931.04	5.06	439.49	0.0831	9,762.28	0.212	313.41

Table 1.7: Mean technical efficiency in 2000 to 2002 by type of hospitals

Type	2000	2001	2002	Average by type
Regional hospitals (N = 25)	0.729	0.683	0.845	0.752
Large general hospitals (N = 48)	0.835	0.790	0.837	0.821
Small general hospitals (N = 19)	0.898	0.875	0.935	0.903
Average by year	0.819	0.779	0.860	0.819

Table 1.8: Mean technical efficiency by region

\Region Period	North	Northeast	Central	East	West	South
pre-UC (2000 and 2001)	0.81	0.91	0.76	0.70	0.71	0.80
post-UC (2002)	0.84	0.96	0.81	0.74	0.78	0.91
%change	5%	5.4%	7.4%	5.5%	9.7%	13.4%

Table 1.9: Number of hospitals experiencing a change in efficiency during 2000-2001, 2001-2002 and 2000-2002

Changes in efficiency	2000 V.S.2002	%	2000 V.S.2001	%	2001 V.S.2002	%
Increase	50	54.3%	25	27.2%	65	70.7%
Unchanged	14	15.2%	8	8.7%	9	9.8%
Decrease	28	30.4%	59	64.1%	18	19.6%

Table 1.10: Changes in number of surgical and non-surgical outpatient visits in 2000 to 2001 and 2001 to 2002

year Type\	2000-2001				2001-2002				
	Non-surgery					Non-surgery			
Surgery	Change in visits (number of hospitals)	Increase	Decrease	Unchanged	Surgery	Change in visits (number of hospitals)	Increase	Decrease	Unchanged
	Increase	15	13	0		Increase	46	4	0
	Decrease	33	30	0		Decrease	42	0	0
	Unchanged	1	0	0		Unchanged	0	0	0

Table 1.11: Changes in number of surgical and non-surgical inpatient visits in 2000 to 2001 and 2001 to 2002

year Type\	2000-2001				2001-2002				
	Non-surgery					Non-surgery			
Surgery	Change in visits (number of hospitals)	Increase	Decrease	Unchanged	Surgery	Change in visits (number of hospitals)	Increase	Decrease	Unchang ed
	Increase	23	10	0		Increase	48	15	0
	Decrease	18	41	0		Decrease	11	18	0
	Unchanged	0	0	0		Unchanged	0	0	0

Note: Non-surgical inpatient visits consist of INPRI and INOTHER. Surgical inpatient visits are INSUR.

Table 1.12: Changes in number of health personnel (physicians and other medical staffs) in 2000 to 2001 and 2001 to 2002

year / Type\	2000-2001				2001-2002				
	Other staff					Other staff			
	Change in number of health personnel (number of hospitals)	Increase	Decrease	Unchanged		Change in number of health personnel (number of hospitals)	Increase	Decrease	Unchanged
Physicians	Increase	36	13	4	Physicians	Increase	46	13	0
	Decrease	14	11	5		Decrease	16	11	1
	Unchanged	5	3	5		Unchanged	5	0	1

Table 1.13: Kruskal-Wallis test of technical efficiency (original DEA) by year

Statistical name	Value
Chi-Square	14.902
Degree of freedom	2
p-value	.001

Table 1.14: Pairwise comparisons: Mann-Whitney test

Technical efficiency (DEA efficiency scores)	t-statistics
2000 versus 2001	-1.777*
2001 versus 2002	-3.884***
2000 versus 2002	-2.027**

Note : * = significant at a 0.10 level of significant
** = significant at a 0.05 level of significant
*** = significant at a 0.01 level of significant

Table 1.15: Number of efficient hospitals (DEA = 1): by type and by region

Type	Region	Year			Total
		2000	2001	2002	
Regional hospitals (N = 25)	North	1	-	3	4
	Northeast	3	1	5	9
	Central	-	-	1	1
	East	-	-	-	-
	West	-	-	-	-
	South	-	-	2	2
	Total	4	1	11	16
Large general hospitals (N = 48)	North	3	1	2	6
	Northeast	5	3	7	15
	Central	1	1	2	4
	East	-	-	-	-
	West	-	-	-	-
	South	1	2	1	4
	Total	10	7	12	29
Small general hospitals (N = 19)	North	-	1	1	2
	Northeast	2	1	2	5
	Central	1	-	1	2
	East	1	-	1	2
	West	-	-	1	1
	South	4	3	5	12
	Total	8	5	11	24
All efficient hospitals		22	13	34	69
Total hospitals		92	92	92	276

Table 1.16 : Efficiency result - Original DEA estimates and bootstrap estimates

DEA	N	Minimum	Maximum	Mean	Std. Deviation
Original DEA scores	276	0.39	1.00	0.8192	0.1450
Bootstrap DEA scores	276	0.40	1.00	0.7633	0.1183

Table 1.17: Descriptive statistics - Original DEA and bootstrap DEA estimates, 2000-2002

YEAR	Statistics	Original DEA	Bootstrap DEA	BIAS
2000 (N = 92)	Mean	0.8193	0.7503	0.069
	Std. Deviation	0.1470	0.1037	0.049
	Minimum	0.47	0.46	0.006
	Maximum	1.00	0.88	0.152
2001 (N = 92)	Mean	0.7787	0.7294	0.049
	Std. Deviation	0.1380	0.1088	0.0337
	Minimum	0.39	0.40	-0.013
	Maximum	1.00	0.90	0.148
2002 (N = 92)	Mean	0.8596	0.8100	0.0496
	Std. Deviation	0.1399	0.1272	0.030
	Minimum	0.48	0.47	0
	Maximum	1.00	1.00	0.151
Total	Mean	0.8192	0.7633	0.056
	Std. Deviation	0.1450	0.1183	0.039
	Minimum	0.39	0.40	-0.013
	Maximum	1.00	1.00	0.152

Table 1.18: Likelihood-ratio tests of null hypotheses for parameters of the Tobit model

Null hypothesis	Test statistics, λ	Result	Implication
$H_0 : a^* = 0$	3.98 $(\chi^2_{(0.05,1)} = 3.84)$	Reject	The intercepts of both periods are not equal.
$H_0 : b_i^* = 0$	22.56 $(\chi^2_{(12)} = 21.03)$	Reject	At least one of slope coefficients is not equal to zero.
$H_0 : c = 0$	5.14 $(\chi^2_{(1)} = 3.84)$	Reject	Include the UC variable in the Tobit model.
$H_0 : a^* = b_i^* = c = 0$	54.12 $(\chi^2_{(14)} = 23.68)$	Reject	At least one of the coefficient is not equal to zero.
$H_0 : b_{PHYRATIO} + b^*_{PHYRATIO} = 0$	1.1 $(\chi^2_{(1)} = 3.84)$	Cannot reject	The sum of the coefficients is equal to zero.
$H_0 : b_{GPPCR} + b^*_{GPPCR} = 0$	4.9 $(\chi^2_{(1)} = 3.84)$	Reject	The sum of the coefficients is not equal to zero.

Table 1.19: Tobit regression results

Variables	Coefficient	Std. Error	t-statistics	p-value
Constant	0.021625	0.1107537	0.2	0.845
OUTUC	-0.48927	0.2042082	-2.4**	0.017
BED	5.99E-05	0.0001273	0.47	0.638
GPPCR	-4.49E-07	3.09E-07	-1.45	0.146
LOS	0.039212	0.0157615	2.49**	0.013
PHYRATIO	2.263139	0.7021813	3.22***	0.001
REFER	-6.32E-06	1.70E-06	-3.71***	0.000
HI	0.232416	0.1984475	1.17	0.242
NORTH	0.025885	0.0436509	0.59	0.553
NORTHEAST	-0.1118	0.047162	-2.37**	0.018
CENTRAL	0.087413	0.0479739	1.82*	0.068
EAST	0.063152	0.0606847	1.04	0.298
WEST	0.191863	0.0573754	3.34***	0.001
ISLAM	0.035815	0.0571966	0.63	0.531
PRIBED	-6.3E-05	0.0000568	-1.11	0.267
D	0.349	0.1742015	2**	0.045
D*BED	-0.0001	0.0001735	-0.59	0.555
D*GPPCR	-8.11E-07	5.10E-07	-1.59	0.112
D*LOS	-0.01271	0.0285234	-0.45	0.656
D*PHYRATIO	-2.83071	1.120029	-2.53**	0.011
D*REFER	3.07E-06	2.72E-06	1.13	0.258
D*HI	-0.25212	0.3347051	-0.75	0.451
D*NORTH	0.02855	0.0770217	0.37	0.711
D*NORTHEAST	0.03633	0.0803209	0.45	0.651
D*CENTRAL	0.076444	0.0847467	0.9	0.367
D*EAST	0.225565	0.105738	2.13**	0.033
D*WEST	0.047963	0.0987488	0.49	0.627
D*ISLAM	-0.12788	0.0993048	-1.29	0.198
D*PRIBED	-1.9E-05	0.0000864	-0.22	0.825
Log likelihood = 102.30				
Likelihood-ratio test of sigma_u=0 (pooled V.S. random effect): Chi-square(1)= 19.50, p-value= 0.000				
Levene Statistic = 2.521 (p-value =0.114)				

Note :
1. The dependent variable = inefficiency score = (1/DEA) -1.
2. D is a period dummy variable indicating 1 if post-UC, and 0 if pre-UC.
3. * = significant at a 0.10 level of significant
   ** = significant at a 0.05 level of significant
   *** = significant at a 0.01 level of significant

FIGURES

Figure 1.1: Technical efficiency

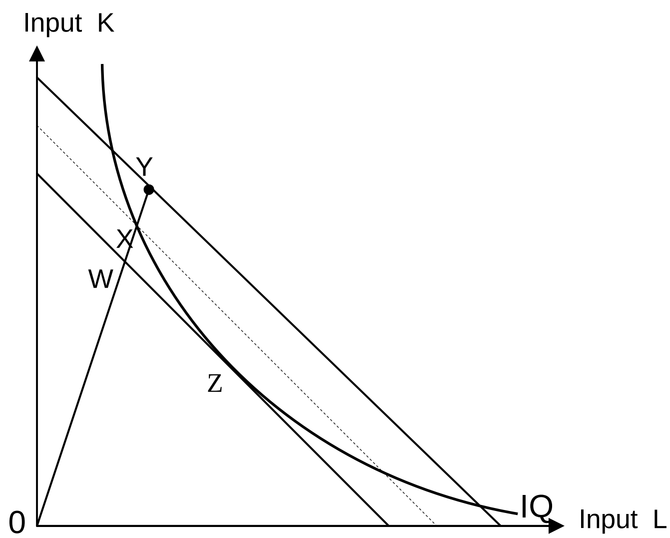


Figure 1.2: Scale efficiencies

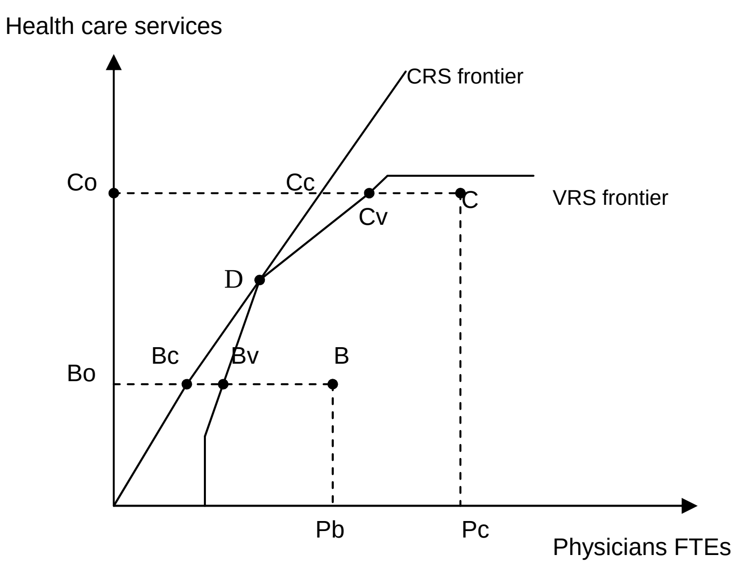


Figure 1.3: Original DEA and bootstrap DEA efficiency scores

(N = 276)

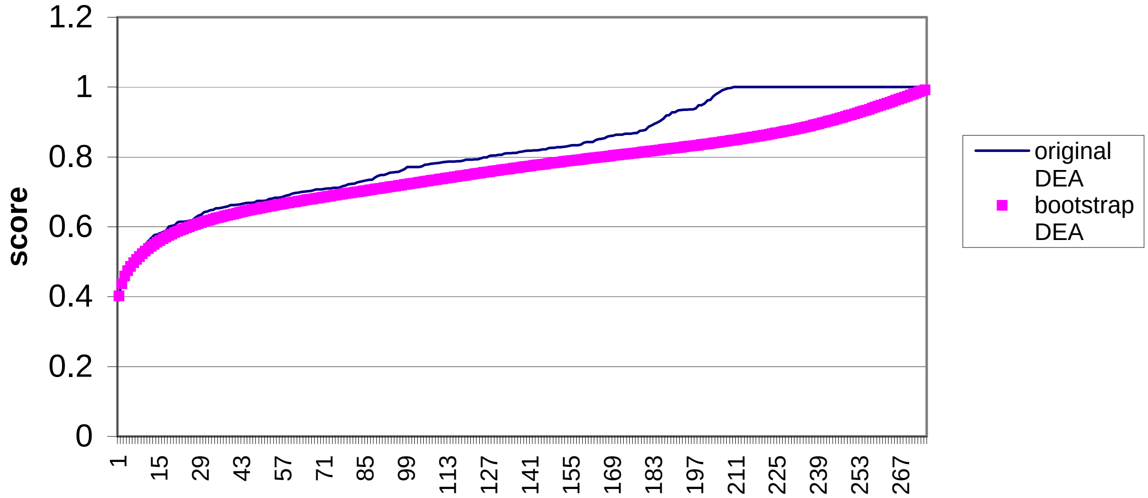


Figure 1.4: Map of Thailand

APPENDIX

Appendix 1.1: Descriptive statistics of input and output variables: mean by year and type of hospitals

N = 92 each year

TYPE	YEAR	INSUR	INPRI	INOTHER	OUTSUR	OUTNON-SUR	BED	PHYSICIAN	NURSE	DENPHAR	OTHERS
Regional hospitals	2000	11,504.4	21,905.2	4,303.1	6,374.0	336,922.5	673.6	80.5	561.6	32.3	95.5
	2001	10,895.1	21,235.9	3,678.4	5,750.8	338,280.9	673.6	86.7	592.2	32.8	102.0
	2002	11,663.8	22,138.1	4,524.5	5,665.4	530,429.6	689.4	95.4	619.1	35.0	120.7
	Total	11,354.5	21,759.7	4,168.7	5,930.1	401,877.7	678.9	87.5	591.0	33.4	106.1
Large general hospitals	2000	5,950.0	14,979.0	2,370.0	2,160.1	197,879.9	391.8	33.3	333.8	18.4	60.5
	2001	5,666.0	14,599.3	1,943.2	2,104.5	203,124.6	391.8	34.9	341.3	19.9	64.5
	2002	6,142.9	14,869.7	1,594.5	2,018.8	344,187.0	395.8	38.0	366.4	22.8	70.1
	Total	5,919.6	14,816.0	1,969.2	2,094.5	248,397.2	393.1	35.4	347.2	20.4	65.0
Small general hospitals	2000	1,963.2	7,498.3	623.5	1,654.0	117,032.2	210.1	17.8	202.3	12.2	38.3
	2001	2,040.7	7,698.6	726.5	1,578.2	121,292.9	210.1	19.8	212.5	13.2	41.3
	2002	2,318.1	8,943.8	507.1	1,392.9	225,693.8	221.0	19.3	214.8	12.9	41.1
	Total	2,107.3	8,046.9	619.0	1,541.7	154,673.0	213.7	18.9	209.8	12.8	40.2
Total	2000	6,636.0	15,316.2	2,534.6	3,200.7	218,966.4	430.8	42.9	368.5	20.9	65.4
	2001	6,338.2	14,977.6	2,163.4	2,986.7	222,951.9	430.8	45.9	382.9	22.0	69.9
	2002	6,853.3	15,621.0	2,166.1	2,880.5	370,325.0	439.5	49.7	403.8	24.0	77.8
	Total	6,609.1	15,304.9	2,288.1	3,022.6	270,747.7	433.7	46.2	385.1	22.3	71.0

Appendix 1.2: Descriptive statistics of input and output variables: mean by year and location of hospitals

N = 92 each year

RELIGION	YEAR	INSUR	INPRI	INOTHER	OUT SUR	OUTNON-SUR	BED	PHYSICIAN	NURSE	DEN PHAR	OTHERS
North	2000	7,381.4	15,703.4	3,132.2	3,600.9	216,926.0	446.9	41.2	374.6	22.1	68.5
	2001	6,646.9	14,452.3	2,721.6	3,515.8	217,247.2	446.9	44.1	400.7	22.5	74.9
	2002	7,527.9	14,168.4	2,652.8	3,094.6	392,024.7	455.7	51.5	416.9	26.5	83.7
	Total	7,185.4	14,774.7	2,835.5	3,403.8	275,399.3	449.8	45.6	397.4	23.7	75.7
North east	2000	9,548.3	19,879.2	5,124.3	5,334.3	244,051.0	519.3	48.6	403.8	22.3	66.1
	2001	9,006.3	19,040.7	3,390.0	4,608.2	247,684.4	519.3	52.4	414.5	24.9	71.8
	2002	10,106.7	20,084.9	3,212.7	4,734.7	409,618.1	522.7	57.8	446.6	27.3	85.5
	Total	9,553.8	19,668.3	3,909.0	4,892.4	300,451.2	520.4	52.9	421.6	24.8	74.5
Central	2000	4,979.7	13,767.9	1,172.1	2,109.0	222,394.9	375.3	38.9	347.1	19.7	63.7
	2001	4,768.5	13,116.5	1,218.5	1,878.2	234,652.6	375.3	42.3	355.7	21.3	66.9
	2002	5,071.4	13,998.5	1,613.8	1,758.6	385,393.4	386.2	44.6	381.8	22.5	71.4
	Total	4,939.9	13,627.6	1,334.8	1,915.3	280,813.6	378.9	41.9	361.5	21.2	67.3
East	2000	7,296.4	16,101.6	1,981.4	2,232.3	234,033.4	499.0	68.0	433.4	24.7	75.6
	2001	7,756.4	16,694.0	2,384.7	1,997.4	244,059.1	499.0	70.1	431.7	23.3	80.4
	2002	7,300.2	17,246.8	2,034.4	1,984.7	382,859.7	526.3	72.4	454.6	25.3	93.4
	Total	7,451.0	16,680.8	2,133.5	2,071.5	286,984.1	508.1	70.2	439.9	24.4	83.1
West	2000	4,866.6	12,124.8	1,190.1	2,677.5	186,946.1	400.3	39.6	330.5	19.1	59.1
	2001	4,477.0	11,824.5	1,302.7	3,174.0	187,704.8	400.3	41.9	335.5	21.3	62.5
	2002	4,324.7	11,711.8	1,294.6	3,014.8	311,606.6	406.5	44.0	356.3	21.4	64.5
	Total	4,556.1	11,887.0	1,262.5	2,955.4	228,752.5	402.3	41.8	340.8	20.6	62.0
South	2000	5,097.0	12,948.2	1,448.2	2,314.4	200,532.3	368.7	35.3	340.5	18.9	62.2
	2001	5,176.1	14,023.8	1,575.2	2,202.3	199,588.1	368.7	37.5	361.8	19.3	64.9
	2002	5,571.4	15,355.5	1,575.1	2,196.0	313,227.2	374.5	39.0	370.2	20.5	70.3
	Total	5,281.5	14,109.2	1,532.8	2,237.6	237,782.5	370.6	37.2	357.5	19.5	65.8
Total	2000	6,636.0	15,316.2	2,534.6	3,200.7	218,966.4	430.8	42.9	368.5	20.9	65.4
	2001	6,338.2	14,977.6	2,163.4	2,986.7	222,951.9	430.8	45.9	382.9	22.0	69.9
	2002	6,853.3	15,621.0	2,166.1	2,880.5	370,325.0	439.5	49.7	403.8	24.0	77.8
	Total	6,609.1	15,304.9	2,288.1	3,022.6	270,747.7	433.7	46.2	385.1	22.3	71.0

Appendix 1.3: Descriptive statistics of explanatory variables: mean by type and location

Type	Region	OUTUC	INUC	GPPCR*	LOS	PHY RATIO	REFER	PRIBED	HI
Regional Hospitals	North	0.21	0.06	23,809.07	5.53	0.10	29,533.20	435.50	0.24
	Northeast	0.30	0.09	17,567.98	5.21	0.11	28,155.00	399.50	0.14
	Central	0.17	0.04	96,184.65	5.49	0.10	8,488.25	371.75	0.19
	East	0.13	0.04	135,105.06	5.67	0.12	9,785.92	391.38	0.23
	West	0.09	0.05	47,163.82	5.84	0.12	10,404.67	405.50	0.15
	South	0.24	0.04	35,486.19	5.30	0.09	15,198.73	319.00	0.19
	Total	0.21	0.06	54,968.28	5.44	0.10	19,043.64	385.10	0.19
Large general Hospitals	North	0.18	0.04	25,094.41	4.63	0.07	9,558.56	257.62	0.21
	Northeast	0.20	0.04	14,047.34	4.51	0.07	9,265.88	85.82	0.20
	Central	0.15	0.04	80,443.19	5.15	0.08	3,267.58	533.73	0.21
	East	0.18	0.05	57,673.67	5.24	0.08	4,217.67	156.50	0.30
	West	0.33	0.04	43,747.51	5.18	0.08	3,696.67	358.25	0.16
	South	0.17	0.04	39,948.94	4.88	0.07	5,094.49	179.57	0.24
	Total	0.19	0.04	40,325.06	4.83	0.08	6,687.77	274.31	0.21
Small general Hospitals	North	0.12	0.02	16,361.53	4.21	0.06	3,356.50	50.00	0.25
	Northeast	0.46	0.04	10,328.20	4.17	0.07	19,426.17	72.50	0.28
	Central	0.17	0.03	98,391.75	5.19	0.07	1,428.50	299.13	0.17
	East	0.42	0.05	20,691.18	4.16	0.10	3,128.33	0.00	0.23
	West	0.16	0.04	42,192.42	4.67	0.08	2,801.11	288.67	0.13
	South	0.23	0.05	35,512.55	4.39	0.06	1,995.14	81.79	0.29
	Total	0.23	0.04	44,358.04	4.55	0.07	4,040.89	151.58	0.23
All hospitals	North	0.18	0.04	23,899.79	4.82	0.08	13,932.02	281.33	0.22
	Northeast	0.26	0.06	14,767.63	4.69	0.08	16,300.37	183.47	0.19
	Central	0.16	0.04	87,535.83	5.23	0.08	3,979.49	450.24	0.20
	East	0.19	0.04	96,636.97	5.33	0.11	7,243.90	268.36	0.25
	West	0.24	0.04	43,591.39	5.07	0.08	4,199.33	338.06	0.15
	South	0.21	0.05	37,140.07	4.81	0.07	6,611.64	180.24	0.25
	Total	0.20	0.05	45,137.10	4.94	0.08	9,498.70	279.07	0.21

Note: * The value is at a constant term of 1988.

Appendix 1.4: Political chronology

Fiscal year*	Month	Important events
2000	Oct. 1999 .. Jan. 2000 … Sep. 2000	
2001	Oct. 2000 .. January Feb, 2001 Mar. Apr. … Sep, 2001	 Thai Rak Thai party had a victory on the general election. UC has started in 6 out of 76 provinces.
2002	Oct, 2001 … Mar. 2002 Apr., 2002 .. Sep., 2002	UC was implemented to most provinces except Bangkok. UC was fully implemented.

Note: * Every fiscal year starts from October.

Appendix 1.5: Number of efficient hospitals (DEA = 1): by region and by type

Region	Type of hospitals	YEAR			Total
		2000	2001	2002	
North	Regional hospitals	1	-	3	4
	Large general hospitals	3	1	2	6
	Small general hospitals	-	1	1	2
	Total	4	2	6	12
Northeast	Regional hospitals	3	1	5	9
	Large general hospitals	5	3	7	15
	Small general hospitals	2	1	2	5
	Total	10	5	14	29
Central	Regional hospitals	-	-	1	1
	Large general hospitals	1	1	2	4
	Small general hospitals	1	-	1	2
	Total	2	1	4	7
East	Regional hospitals	-	-	-	-
	Large general hospitals	-	-	-	-
	Small general hospitals	1	-	1	2
	Total	1	-	1	2
West	Regional hospitals	-	-	-	-
	Large general hospitals	-	-	-	-
	Small general hospitals	-	-	1	1
	Total	-	-	1	1
South	Regional hospitals	-	-	2	2
	Large general hospitals	1	2	1	4
	Small general hospitals	4	3	5	12
	Total	5	5	8	18
TOTAL		22	13	34	69

Appendix 1.6: Bootstrapping DEA Algorithm (applied from Badin and Simar, 2003)

Step 1). Find the original efficiency estimates. For each observed producer $(x_i, y_i) \in \chi_n$, compute the DEA estimator of the efficiency score $\hat{\theta}_i = \hat{\theta}_{DEA}(x_i, y_i)$, i = 1, ..., n.

Step 2). If $(x_0, y_0) \notin \chi_n$ repeat step 1) for (x_0, y_0) to obtain $\hat{\theta}_i = \hat{\theta}_{DEA}(x_0, y_0)$.

Step 3). Define $S_m = \{\hat{\theta}_1, ..., \hat{\theta}_m\}$ where m = # $\{\hat{\theta}_i < 1\}_{1 \le i \le n}$, i.e. the number of inefficient producers.

Step 4). Giving n is the number of decision-making units (DMUs), estimate $f(\hat{\theta})$ from the remaining $\hat{\theta}$ and generate B samples of the boundary condition $\hat{\theta} < 1$ (the size n-1), which is $\{\hat{\theta}_1^{*b}, ..., \hat{\theta}_{n-1}^{*b} \cdot\}_{b=1}^{B}$. The steps are as follows:

4.1). Given a random sample $x_1, ..., x_n$ with a continuous, univariate density function, the kernel density estimator is defined by:[19]

$$\hat{f}(z) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right) \qquad \text{(a1)}$$

[19] *Kernel density estimation* is a nonparametric technique for density estimation in which a known density function (the *kernel*) is averaged across the observed data points to create a smooth approximation. Usually, the kernel function is a probability density function, symmetric around zero.

where $K(\cdot)$ is the kernel function and h is the bandwidth parameter. Under mild conditions (h must decrease with increasing n) the kernel estimate converges in probability to the true density. Performance of kernel is measured by MISE (mean integrated squared error).

Bandwidth selection is a crucial issue in the application of the smoothing procedure. Refer to Silverman (1986) for a completed review of several approaches of bandwidth selection. In this paper, the bandwidth function rule for univariate data recommended by silverman (1986, eq.3.31) is

$$h = 0.9\left(\min\begin{cases}\hat{\sigma}_{\hat{\theta}}\\ R_{13}/1.34\end{cases}\right) n^{-1/5}$$

where R_{13} denotes the inter-quartile range of the sample $\{\hat{\theta}_i\}$ and denotes the standard deviation estimate of the efficiency estimates $\{\hat{\theta}_i\}$, respectively.[20]

4.2). Using the reflection method (Silverman,1986), we estimate $f(\hat{\theta})$ under the boundary condition $\hat{\theta} < 1$. Suppose we have m inefficient producers, denoting $S_m = \{\hat{\theta}_1, ..., \hat{\theta}_m\}$. In order to find a consistent estimator of $f(\hat{\theta})$, let $\{\beta_1^*, ..., \beta_{n-1}^*\}$ be a bootstrap sample, obtained by sampling with replacement from S_m and $\{\varepsilon_1^*, ..., \varepsilon_{n-1}^*\}$ a random variable of standard normal deviates. By the convolution formula, we have

$$\tilde{\theta}_i^* = \beta_i^* + h\varepsilon_i^* \sim \frac{1}{m}\sum_{j=1}^{m}\frac{1}{h}\phi\left(\frac{z - \hat{\theta}_{(j)}}{h}\right)$$

for i = 1,…,n-1. Define now for i = 1,….,n-1 the bootstrap data:

$$\theta_i^* = \begin{cases}\tilde{\theta}_i^* & if\ \tilde{\theta}_i^* < 1\\ 2 - \tilde{\theta}_i^* & otherwise.\end{cases} \tag{a2}$$

20 The choice of smoothing variable is chosen because Silverman (1986) suggested that it copes very well for a wide range of densities; both unimodal densities and moderately bimodal densities.

where θ_i^* defined in (a2) is proved to be random variables distributed according to $\hat{f}_h(z)$. The final smoothed resample efficiencies are obtained by rescaling the bootstrap data making the variance is approximately the sample variance of $\hat{\theta}_i$. We employ the following transform:

$$\hat{\theta}_i^* = \bar{\hat{\theta}} + \frac{1}{\sqrt{(1 + h^2 / \hat{\sigma}^2)}} (\theta_i^* - \bar{\hat{\theta}}),$$

where $\bar{\hat{\theta}} = \frac{1}{n}\sum_{j=1}^{n} \hat{\theta}_j$ and $\hat{\theta}^2 = \frac{1}{n}\sum_{j=1}^{n} (\hat{\theta}_j - \bar{\hat{\theta}})^2$.

Step 5). Then, draw n-1 bootstrap values $\hat{\theta}_i^*$, i = 1,….,n-1 from the kernel density estimate of $f(\hat{\theta})$ and sort in ascending order: $\hat{\theta}_{(1)}^* \le ... \le \hat{\theta}_{(n-1)}^*$.

Step 6). Repeat step 5 (drawing n-1 bootstrap values $\hat{\theta}_i^*$) B times (in this study, 1000 times), to obtain a set of B bootstrap estimates $\left\{\hat{\theta}_{(n-j)}^{*b}\right\}_{b=1}^{B}$, for some $1 \le j \le n-1$.[21]

Step 7). Finally, approximate $\tilde{\theta}_{(n-j)}$ for some j $(1 \le j \le n-1)$ by the average of $\hat{\theta}_{(n-j)}^{*b}$ over the B simulations (in this study, 1000 times):

$$\tilde{\theta}_{(n-j)} = \frac{1}{B}\sum_{b=1}^{B} \hat{\theta}_{(n-j)}^{*b} \qquad \text{(a3)}$$

[21] Replications is set to B = 1000. Efron and Tibshirani (1993), p.275, recommend at least this number of simulation replicates in order to make the variability of the boundaries of the bootstrap confidence intervals "acceptably" low.

CHAPTER 2

EFFECTS OF HEALTHCARE ON THAI AGRICULTURAL HOUSEHOLD EARNINGS

Abstract

Health is an important type of human capital that can promote the efficiency of agricultural production. Insubstantial health may significantly reduce the productivity of household farm production, especially when its workforce mainly engages in physical labor. This paper investigates the effect of health status and healthcare utilization on agricultural household earnings in Thailand. The paper also evaluates the role demographic factors including household characteristics and the level of education of households, play in household production. Under the theoretical model, a utility-maximization production model is formulated in which health status and education affect household resource allocation.

The data from the Thai Socio Economic Survey (SES) in 2002 are categorized into five sectors; rice farming, other cropping, livestock, poultry, forestry and other agricultural services. Since the logarithms form is undefined, the Box-Cox transformation is applied. A key finding from the regression indicated that the human capital factor, which was education, appeared to increase farm income.

1. Introduction

The contribution of human capital to household earnings among farm households in developing countries has been well studied (see surveys by Jamison and Lau, 1982; Schultz, 1988). Yotopoulos (1967), and Patrick and Kehrberg (1973), among others, were among the first to evaluate the productive role of education for developing agriculture. However, human capital is broader than education. Health is another important type of human capital, and families often invest in human capital by improving health status. More skilled labor, as evidenced by greater education, can affect the efficiency of production. Healthier labor may have the same result (Strauss and Thomas, 1998).

In developing countries, poor health may substantially reduce the productivity of household farm production, especially when subsistence agriculture is a dominant sector and its workforce mainly engages in physical labor.[1] Although Thailand is no longer predominantly an agricultural economy, the agricultural sector, especially rice farming, is still one of the most important economic foundation. In terms of labor force, about half of the Thai population is engaged in agriculture, although farm agriculture accounts for only 10% of GDP.

The objective of this paper is to investigate the effect of health status and healthcare access and utilization on agricultural household earnings in Thailand. Besides examining if health status affects the household income, this paper also evaluate the role demographic factors, including household characteristics and the level of education of households, play in household production.

---

[1] Contoyannis and Rice (2001) provided some reasons that health may affect wages in a developed country as well. They argue that an increase in health increases individual productivity and eventually, wage rate.

The role of health in household decision making has been well-studied in the literature. Grossman (1972a) introduced a model of health capital using a utility maximization model as an extension of the theory of the allocation of time (Becker 1965). Applying the household production model, he assumed that family members can invest their time and inputs to create health. Health is thus treated as an endogenous choice. It is both a source of utility and affects the time spent in market and non-market activities. His seminal model of intertemporal choice concluded that the stock of health can be treated like a durable investment good, and that the demand for health care depended on the shadow price of health, the price of health care, education level, age, and current health status.

The theory of household production has played a significant role in the analysis of allocation, production and consumption of home activities. A number of studies focus on the role of intrahousehold resource decision on health. Pitt and Rosenzweig (1985) showed that nutritional intakes affect health in Indonesia. Deolalika (1988) supported this finding that the nutritional status, which was measured by weight-for-height, influences farm production and labor productivity. It is also possible that household members (husband and wife) determine and allocate health investment across family members. Jacobson (2000) extended the Grossman model explaining how family members may have some influence on an individual's health and family members related behavior. She found that a child with unhealthy parents is more likely to have worse health compared with a child with healthy parents. Applying a game theory model, Bolin et al. (2001 and 2002) presented a model of Nash-bargaining between spouses over resource allocation for the family distribution of health.

The agricultural household model has been extensively utilized in labor market and human capital studies, especially in developing countries. The basic idea behind this framework is that households allocate time and inputs to produce, consume, and sell their commodities, to maximize a utility function. Early examples of agricultural household models, focusing on farm price policies, can be found in Lao, Lin, and Yotopoulos (1978), Strauss (1982), and Singh, Squire, and Strauss (1986). Lao, Lin, and Yotopoulos (1978), using demand systems, examined the effect of full income on different kinds of household expenditures. Strauss (1982) applied a farm household model to investigate the effect of prices and income on household nutrient calorie ability in Sierra Leone, which had an underweight rural child problem. Later, the applications of farm household models extend to other topics such as off-farm labor supply, nutrition policy, labor supply, migration, income distribution, and family planning (Taylor, 2003). Recently, Matshe and Young (2004), adapting a model by Benjamin and Guyomard (1994), used a two-person household to analyze the off-farm labor decisions of small-scale agricultural household members. They found that gender, education, and farm characteristics affected the labor market and health decisions.

A number of studies analyze the effects of human capital on rural household earnings and efficiency (see Jamison and Lau, 1982 and Phillips, 1994). Many studies emphasize the impact education and schooling, as human capital, on rural household earnings. Yang and An (2002), using Chinese farm data, identified the role of human capital on the profit of agriculture and non-agriculture sectors. They showed that schooling and experience had significant impacts on rural household earnings, accounting for 27% of the total contribution to earnings. Several studies support the idea that the role of schooling improves production efficiency (e.g. Dey et al. 2000; Abdular and Eberlin 2001; and Yang, 2003).

The effect of human capital as measured by health on individual wage and/or earnings has also been studied (see Currie and Madrian, 1999 for a survey of literature). Most studies in both developed and developing countries found a positive impact of health on wage or earnings. In the U.S., Chirikos and Neltel (1985), estimating effects of poor health on wages and number of working hours, showed that poor health history has a negative impact on wages, especially with black populations. Johnson and Lambrinos (1985) used the 1972 Social Security Survey of Disabled and Non-Disabled Adults to estimate wage discrimination against handicapped workers. They found that discrimination had occurred in which handicapped women were subject to discrimination based on gender and handicap. Recently, Contoyannis and Rice (2001) examined the effect of health on hourly wages using longitudinal data from the British Household Panel Survey. They found that good self-assessed health increases the hourly wage for females and the health variables are positively correlated with the time-invariant individual effect.

In developing countries, Strauss and Thomas (1998) supported the positive relationship between health (represented by nutrient intake), on wages and productivity using both experimental and nonexperimental methods. They also reported that health had a larger marginal return at very low levels of health and in labor-intensive jobs. Schultz and Tansel (1997) examined the relationship of health and wage in Ghana and Cote d'Ivoire. Using the number of disabled days as a proxy of health status, they found that disabled days are negatively correlated to the wage rate.

Perhaps more importantly there seem to be a positive relationship between education and health (Grossman, 1975). Grossman (1972b) demonstrated a conceptual framework to examine why education may lead to better health. His model, assuming education is exogenously determined, predicted that the more educated demand a larger optimal stock of health. Grossman and Kaestner (1997) provided an extensive review on health and schooling in the U.S. and some developing countries. They suggested that education as measured by the years of formal schooling completed positively correlates with health. Studies on this topic differ by the measurement of health; either a direct measure of health status or an indirect measure of health inputs. Using mortality rate as a health variable, Duleep (1986), Menchik (1993) explained that schooling is a determinant of mortality. Moreover, Lleras-Muney (2005) argued that education might have a causal impact on health.

Several studies in developing countries also found a positive correlation between education and health. Glewwe and Jacoby (1995) showed that malnourished children tend to delay the entry into schools in Ghana. Moreover, Wolfe and Behrman (1983), by examining the role of schooling in nutrient intake of 15 food groups in India, showed that a woman's schooling is positively correlated to nutrients and health status.

Currie and Madrian (2004) discussed measurement errors and the endogeneity of health. Several methodologies have been used to deal with endogeneity bias, for example: Generalized Method of Moments (Havemen et al., 1994; Mullahy and Sindelar, 1995); three-stage least squares (Sundberg, 1996). Currie and Madrian (2004) reported that the majority of studies of an effect of health on earnings or wages used ordinary least squares and instrumental variable techniques. Studies that used two-stage least squares include

Grossman and Benham (1974) and Baldwin et al. (1994) Later, Yan and An (2002), applying the two-stage least squares (2SLS) procedure, estimated the structural profit functions to identify the impact of human capital on efficiency.

Section 2 discusses a theoretical section by formulating a utility-maximization model in which health status affects household resource allocation. Section 3 outlines the empirical model. Section 4 presents data and variable selection. Regression results are shown in section 5, while the conclusions and policy implications are drawn under the last section.

2. Theoretical model

Agricultural household models are widely used to investigate household labor allocation in agriculture sectors in developing countries. The economic decisions of the farm household consist of production, consumption, and labor supplies. These aspects of the choices are interrelated because production decisions determine the level of household income that affects consumption and demand for leisure. To establish a conceptual framework that identifies sources of human capital returns, we consider a subsistence household that maximizes a utility function with consumption, farm production, health status, and leisure as arguments.

Assume a household utility maximization problem is:

$$U = U(C, F, H, M) \qquad U^{'} > 0, U^{''} < 0 \qquad (1)$$

where F is a consumption of farm commodities, C represents a consumption of purchased goods, H is health status, and M represents leisure. This model includes household members' health in the utility function because good health is desirable (Pitt and Rosenzweig, 1985).

The stock of health status (*H*) is assumed to be determined by the consumption of health care (*h*) which includes goods such as exercises, consumption of medical services e.g. time spent going to the doctor, etc. that improve health status. The time for farm work (L) is also included in the health status function. The model assumes that farm work may be detrimental to health status because of its danger, exposure to chemical, and often aspects of the job compared to non-farm workers. Furthermore, the endowed health of an individual (μ), which is not a choice variable, captures factors beyond the control of the household, for example, genetic traits or environment factors.

$$H = h(h, L; \mu) \qquad \frac{\partial H}{\partial h} > 0\,,\ \frac{\partial^2 H}{\partial h^2} < 0\,,\ \frac{\partial H}{\partial L} < 0 \qquad (2)$$

Available time (Ω) is allocated to farm work (L), market (non-farm) work (*N*), health consumption (*h*), and leisure (*M*).

$$\Omega = M + L + N + h \qquad (3)$$

The money budget constraint is:

$$P_h h + P_C * C = P_F * (Q - F) + W * N \qquad (4)$$

where *W* represents wage for non-farm work, *Q* is farm production, and P_h, P_C and P_F represent the price of health consumption or the opportunity cost of going to see the doctor, the price of non-farm goods, and the price of farm goods, respectively.

Mincer (1958) provided the foundations for the schooling-earnings relationship model in which education influences the wage rate. The wage function is given by the following:

$$W = W(E) \qquad w'(E) > 0, \quad w''(E) < 0 \qquad (6)$$

Farm production is determined by:

$$Q = \theta(L, H; K, E) \qquad \theta_1 > 0, \quad \theta_2 > 0 \qquad (5)$$

where Q is a farm output; K is capital input, treated as quasi-fixed factor; L is labor input; and H is health status. E is a schooling level. Education in this model is conditionally predetermined.[2] The assumption of a positive effect of health status to farm production is an important point of the model for measuring efficiency. This model assumes a diminishing marginal return on L and H.

The household problem is given by the Lagrangian function:

$$\max_{C,F,h,L,N} \tilde{L} = U(C, F, H, M) + \lambda[P_F * Q + W(E) * N - P_C * C - P_h h - P_F F] \qquad (7)$$

where λ is the marginal utility of income. Choice variables include farm good consumption, purchased good consumption, health consumption, family farm labor, and family labor in market work. Using the time budget constraint, the first-order conditions for this maximization problem yield the following:

$$\frac{\partial \tilde{L}}{\partial C} = \frac{\partial U}{\partial C} - \lambda * P_C = 0 \qquad (8)$$

$$\frac{\partial \tilde{L}}{\partial F} = \frac{\partial U}{\partial F} - \lambda * P_F = 0 \qquad (9)$$

$$\frac{\partial \tilde{L}}{\partial h} = \frac{\partial U}{\partial H}\frac{\partial H}{\partial h} - \frac{\partial U}{\partial M} + \lambda\left(P_F \left(\frac{\partial Q}{\partial H}\frac{\partial H}{\partial h} \right) - P_h \right) = 0 \qquad (10)$$

$$\frac{\partial \tilde{L}}{\partial N} = -\frac{\partial U}{\partial M} + \lambda * W = 0 \qquad (11)$$

[2] Most Thai farmers, especially heads of households, who have finished schools, do not usually seek additional education or trainings.

$$\frac{\partial \tilde{L}}{\partial L} = -\frac{\partial U}{\partial M} + \lambda\left(\frac{\partial Q}{\partial L} + \frac{\partial Q}{\partial H}\frac{\partial H}{\partial L}\right)P_F = 0 \quad (12)$$

$$\frac{\partial \tilde{L}}{\partial \lambda} = P_F * (Q - F) + W * N - P_C * C - P_h * h = 0 \quad (13)$$

Dividing equation (8) by (9) gives the condition that the marginal rate of substitution between farm good consumption and purchased good consumption must equal their price ratio.

$$\frac{MU_F}{MU_C} = \frac{P_F}{P_C} \quad (14)$$

The model assumes that the budget constraint is always binding at an optimum, which implies that the marginal utility of income is positive. Equations (11) and (12) imply the optimum condition for wage of non-farm work below:

$$W = P_F \frac{\partial Q}{\partial L} + P_F \frac{\partial Q}{\partial H}\frac{\partial H}{\partial L} \quad (15)$$

The left hand side term, wage, represents the opportunity cost of non-farm work or the marginal farm cost of farm labor (MFC_L). The right hand side consists of two parts, which are the marginal revenue product of labor (MRP_L) and the health value effect. The sign of the health value effect is negative because the model assumes that $(\partial H / \partial L) < 0$. Equation (15) indicates the detrimental health cost of labor results a disconnection between the market wage and the *market* MRP_L, which implies that $W < P_F \frac{\partial Q}{\partial L}$. An important implication shows that the health effect may cause farm labor to stop working sooner than if there was no health effect.

Dividing (11) in (10) gives:

$$MRS_{hM} \equiv \frac{MU_h}{MU_M} = 1 + \left(\frac{P_h}{W} - \frac{P_F \frac{\partial Q}{\partial H} \frac{\partial H}{\partial h}}{W} \right) = 1 + \left(\frac{MC_h - MVP_h}{MC_M} \right) \quad (17)$$

Equation 17 shows the optimal marginal rate of substitution (MRS) between health consumption and leisure. It is expected that $MRS_{hM} > 0$. On the right hand side, P_h, is the market cost of health care to the worker (MC_h) and the latter term, $\left(P_F \frac{\partial Q}{\partial H} \frac{\partial H}{\partial h} \right)$, is the marginal value product of health care for farm work (MVP_h). Wage is the market cost or the opportunity cost of leisure (MC_M). The MRS is greater than zero because the indifference curve is convex. This implies that the MRS between health consumption and leisure (MU_h / MU_M) is always greater than $\left(\frac{MC_h - MVP_h}{MC_M} \right)$. It is usually expected that $\frac{MU_h}{MU_M} = \frac{MC_h}{MC_M}$ or the MRS between health consumption and leisure is equation to the real marginal cost of health consumption. Thus, (17) shows a disconnection from the optimum because of non-market effect of health consumption on farm output.[3]

Next, in order to examine an effect of education on health consumption, we substitute (15) and (11) in (17) which gives:

$$\frac{\partial U}{\partial H} \frac{\partial H}{\partial h} + \lambda \left(P_F \frac{\partial Q}{\partial H} \frac{\partial H}{\partial L} + P_F \frac{\partial Q}{\partial L} - P_F \frac{\partial Q}{\partial H} \frac{\partial H}{\partial h} + P_h \right) = 0 \quad (18)$$

Using a total differentiation on equation (18):

[3] Equation (17) also constrains how much *h* will be consumed; so that the MU_h, implicit in (10), is positive.

$$\frac{dh}{dE} = \frac{\lambda\left(P_F \frac{\overset{+}{\partial^2 Q}}{\partial L \partial E} + P_F \frac{\overset{-}{\partial H}}{\partial L}\frac{\overset{+}{\partial^2 Q}}{\partial H \partial E} - P_F \frac{\overset{+}{\partial H}}{\partial h}\frac{\overset{+}{\partial^2 Q}}{\partial H \partial E}\right)}{Z} \qquad (19)$$

where $Z = [H_h U_{HH} H_h + U_H H_{hh} - \lambda P_F Q_{LH} H_h - \lambda P_F Q_H H_{Lh} - \lambda P_F H_L Q_{HH} H_h + \lambda P_F H_h Q_{HH} H_h + \lambda P_F Q_H H_{hh}] < 0$

The above expressions such as H_h, U_{HH}, U_{hh}, and etc. denote partial derivatives.

Equation (19) represents the response of investment on health with changes in the exogenous level of education. Since the model assumes that education is a predetermined variable, this equation presents a comparative static for a change from the status quo, for example, in the compulsory education level.[4] It yields two possible conditions as follows:

Condition 1: If $\left(\frac{\overset{+}{\partial H}}{\partial h} - \frac{\overset{-}{\partial H}}{\partial L}\right)\left[P_F \frac{\partial^2 Q}{\partial H \partial E}\right] > P_F \frac{\partial^2 Q}{\partial L \partial E}$, then $\frac{dh}{dE} > 0$.

The term $P_F \frac{\partial^2 Q}{\partial H \partial E}$ on the left hand side represents how changes in education affect the marginal revenue product of health (before scaling by the net effect of health consumption (h) and farm labor (L) on health). $\partial H / \partial L$ is always negative because health status is negatively related to farm labor. Thus, the left hand side term represents the effect of education on the net marginal return to health care. The term $P_F \frac{\partial^2 Q}{\partial L \partial E}$ on the right hand side shows how changes in education affect the marginal revenue product of farm labor.

[4] In Thailand, the government mandated the education level to grade 4 and expanded to grade 6, and grade 9 in 1977 and 2003, respectively. Since 1977, Thailand's educational system was changed from a 4-3-3-2 structure to a 6-3-3 system in which six years of compulsory primary education is followed by three years of lower secondary school and by three years of upper secondary schooling.

Condition 1 shows that, when the change in education has a larger effect on the net marginal return to health care than it does on the marginal return to farm labor, then ($P_F \dfrac{\partial^2 Q}{\partial L \partial E}$), a change in health consumption *(h)* with respect to education *(E)* is positive. This condition implies that households with more education are likely to have higher health consumption compared to households with lower level of education. In other words, when the mandated level of public education is increased, households adjust their consumptions by increasing health consumption (*h)* when the effect of education on the net marginal return to health is greater than on the marginal return to farm labor. When condition 1) and 2) are equal to zero, it implies that $dh/dE = 0$, which is an optimality. Therefore, the society may adjust the amount of health consumption up to the equilibrium condition.

Condition 2: If $\left(\dfrac{\overset{+}{\partial H}}{\partial h} - \dfrac{\overset{-}{\partial H}}{\partial L}\right)\left[P_F \dfrac{\partial^2 Q}{\partial H \partial E}\right] < P_F \dfrac{\partial^2 Q}{\partial L \partial E}$, then $\dfrac{dh}{dE} < 0$.

From condition 2, when the change in education has a larger effect on the marginal return to farm labor than it does on the net marginal return to health care, then households with more education tend to have lower health consumption compared to households with lower level of education. That is, under this condition, households would invest less in health when the mandated level of education is increased. This optimal tradeoff level implies that when more education is mandated, health consumption is adjusted downward accordingly.

By solving the first order conditions (8) to (13), we obtain the household input demand equations for health and labor inputs at the optimum level. Gurgand (2003) suggested that we could define an implicit price of household farm labor that is a function of the exogenous, $\widetilde{W}(P_F, P_h, M, E, K)$, which is equal to the market price, W (as in equation 15). The demand functions are:

$$L = l(P_F, P_h, \widetilde{W}(P_F, P_h, M, E, K), M, E, K) \quad (20)$$

$$h = l(P_F, P_h, \widetilde{W}(P_F, P_h, M, E, K), M, E, K) \quad (21)$$

Assuming Π is household farm earnings, the earning function can be represented below:

$$\Pi = \pi(P_C, P_F, P_h, \widetilde{W}(P_F, P_h, M, E, K), M, E, K) \quad (22)$$

In addition, the utility maximization model in this paper is relevant to Thai agriculture because most Thai farmers are subsistence farmers who may not maximize enterprise profit. Labor input and health input are significant factors of production. The decision making of how much household inputs are allocated could affect income. Like Gurgand (2003), this paper applies a utility maximization model to determine the effect of human capital on farm income.

The utility maximization model can be connected with earning estimations when there is a trade-off between household net earnings and the consumption of farm products (*F*) in this study. When farm households consume more farm-grown food, this may increase utilities, but lower profits, and vice versa. It is important to note that *F* and profits cannot be perfect substitutes because the perfect substitution implies that *F* is equal to zero or in other words, there is no difference between store-bought food and farm-grown food. This analysis does not estimate a model of labor and health demand, but rather take a reduced form approach. In fact, health inputs determine labor inputs, and both health and labor inputs also jointly determine income. A profit (net income) maximization model, representing as a special case of our model under Appendix 2.1, shows fewer implications compared to the utility maximization model.

In the empirical analysis, in addition to an investigation of the health effect on earnings, we test the hypothesis that there is a relationship between health and education on the farm earnings regarding condition 1 and 2. The hypothesis testing is explained in section 3 in more detail.

3. Empirical model

Commonly in previous studies of household earnings used an unrestricted Cobb-Douglas production function because it is linear and homogenous in logarithms. However, the log transformation is only applicable when all the observations in the data set are positive. Since, in our data, both dependent variable and some explanatory variables contain some zero value, which are undefined in logarithms, this paper applies a Box-Cox transformation (Box and Cox, 1964) in which both the dependent and independent variables are subjected to the power transformation:

$$X^{(\lambda)} = \begin{cases} \dfrac{X^{\lambda} - 1}{\lambda} & \lambda \neq 0 \\ \log(X) & \lambda = 0 \end{cases} \tag{23}$$

where the X_i's are the independent variables observations that is subject to a nonlinear power transformation, λ is an unknown power parameter constrained to be strictly positive. At the same time, a dependent variable or Y is transformed to by $Y^{(\theta)}$, similarly to equation (23) where θ is an power parameter constrained to be positive. Then, the estimated regression becomes $Y^{(\theta)} = \alpha + \beta X^{(\lambda)} + \varepsilon$. The model is estimated by a maximum likelihood procedure allowing all variables to be Box-Cox transformed that may involve a non-linear relationship.

Notes that the coefficients in a nonlinear Box-cox model are not equal to the slopes with respect to independent variables. For the case that the power coefficient of the dependent variable and regressors are similar, Zarembka (1968) showed the elasticity of a given independent variable X_k, (k = 1,…, N) evaluated at the sample means as: $\eta_k = \beta_k (\overline{X}_k / \overline{Y})^{\lambda}$. In our analysis with different power coefficients of the independent variables and dependent variables; λ and θ, respectively, Osula and Adebisi (2001) noted that the elasticity of income (Y) with respect to a dependent variable (X_k) is:

$$\eta_k = \beta_k (\overline{X}_k^{\lambda} / \overline{Y}^{\theta}) \quad (24)$$

The aggregate profit function (net earnings) is approximated by the following specification.

$$\Pi_i^{(\theta)} = \alpha + \beta_1 Z_i^{(\lambda)} + \beta_2 H_i^{(\lambda)} + \beta_3 \left(H_i^{(\lambda)}\right)^2 + \beta_4 E_i^{(\lambda)} + \beta_5 \left(E_i^{(\lambda)}\right)^2 + \beta_6 C_i^{(\lambda)} + \beta_7 H_i^{(\lambda)} E_i^{(\lambda)} + \varepsilon_i \quad (25)$$

where i indexes the household, Π is a vector of net farm earnings, Z is a vector of input factors, H is a vector of health consumption variables, E is a vector of education variable, C is a vector of household characteristics including dummy variables, and ε is a vector of unobserved random errors. The quadratic terms in health and schooling account for potential non-linear effects of human capital factors on earnings.

From condition (1) and (2), in order to examine a relationship between health and education on earnings, the analysis in equation (25) is tested whether there is a statistical significance of the interaction variable. The null hypothesis is $H_0 : \beta_7 = 0$. When one can reject the coefficient, a positive sign of $\partial h / \partial E$ is consistent with Condition 1 while the negative sign matches Condition 2. The interaction variable between health and education is not rejected if it has no predictive power on the dependent variable.

In addition, since both education and health variables, related by unobserved variables, may correlate with the error term of the dependent variable, the parameter estimates may be inconsistent because of the endogeneity problem. It is quite common that estimated effects of education are biased because of omitted unobserved variables, such as unobserved ability, which are correlated both in years of schooling and with earnings. When ability is omitted from the dependent variable, the model overestimates education's true effect on earnings because it captures some of the earnings effects of ability. Moreover, health can lead to a bias similar to “ability bias” in human capital models if the model fails to accurately measure (Currie and Madrian, 1999).

As noted in Contoyannis and Rice (2001), previous literature use cross-sectional data with ordinary least squares and instrumental variable techniques such as two-stage least squares (2SLS) to mitigate the endogeneity problem. This paper uses this approach to take into account the possible endogeneity of health consumption and education variables on farm earnings. The analysis applies an instrumental variable technique of two-stage least squares; and compare the results to OLS.

The two-stage least squares are utilized as follow. On the first stage we run the endogenous variable on the *i* instruments. Second, this model uses the fitted value of the above regression instead of the health variable itself as an explanatory variable. The model shows that both education and health variables are endogenous regressors. When there are more than one endogenous regressor, more instrumental variables can be added. Each instrument cannot be perfectly correlated with the other instruments or the exogenous regressors.

In order to examine whether there is a discrepancy between OLS and IV estimates, a Wu-Hausman test is applied. The null hypothesis is that regressors are exogenous or $Cov(X, \varepsilon) = 0$. If we cannot reject the null hypothesis of weak exogeneity, there is no endogeneity problem implying that both OLS and 2SLS are consistent estimators. Thus, OLS is preferred because it is more efficient than 2SLS estimates. On the other hand, if the null hypothesis is rejected, 2SLS is necessary.

4. **Data and variable selection**

4.1 Data

The data used in this study is the Thai Socio-Economic Survey (SES) in 2002. The survey, covering 34,785 households, was collected by the National Statistics Office (NSO). The survey consists of two separate sets; the main SES set and the detail data of working household members[5]. The main survey contains household income and expenditures, household consumptions, changes in assets and liabilities, household medical care, and housing characteristics. The second set of data includes capital and labor input variables by type of household businesses (farm or non-farm), household earnings, and operating cost. Data are collected from every province in both municipal and non-municipal areas.

5 We would like to thank Dr. Vimut Vanitchareanthum and University of Chicago-UTCC Research Center for the contribution of both sets of data. The second set of data are the 2002 SES for Evaluation of the Impact of the Thailand Village Fund on Household Welfare and Community Development of human capital variables for each household, funded by the Worldbank.

Agricultural households include rice farming, other cropping, livestock, poultry, forestry and other agricultural services. Although there are approximately 9,500 agricultural households on the master SES survey, only 1,871 households were left after matching with the second set. As shown in Table 2.1 (column (2)), rice farming accounts for 47 % of all household farm enterprises following by other crop farms and permanent crops (17.9% and 17.7%, respectively). Although the distribution in type of farms in the sample is quite similar to the survey, the number of samples from rice farming is slightly lower than the survey data while samples from livestock and poultry are a little over-represented. For the sample by regions, the northern region, the northeast, the central, the south account for 23%, 44.6%, 15.7%, and 16.8%, respectively (Table 2.2, column (2)). The sample from the central and the northern regions are slightly under-represented while the sample from the northeast and the south are slightly over-represented. Furthermore, the distribution of head of household level of education is also about the same as in the survey data, in which almost 70% of head of agricultural households attain a lower elementary level (grade 4) -- the compulsory education level before 1977-- while 6% have no education. About 13.6% finish an elementary level (grade 6), 5% attain grade 9, and the other 5.5 % finish high school or higher education (Table 2.3).

The distribution of monthly farm income differs by each region. Table 2.4 provides more detail about the data used in the sample. While the average net earnings from the sample is 3,507 baht per month, households in the central region earn the highest average income by average (8,046 baht), followed by the southern region (5,165 baht) and the northern region (3,102). The northeastern households are the poorest, earning approximately 1,500 baht a month. Overall, fruits and permanent crops household earn the highest income (6,050 baht), followed by livestock's and other crops farms' (4,619 and 4,441 baht, respectively). Earnings from forestry and hunting households vary significantly, ranging from 480 baht in the northeast region to almost 7,000 baht in the central region. In term of regions, households in the central region make the highest income in livestock (17,685 baht) while earn the least in rice farming (5,849 baht). For the northern region, livestock households make the most earnings (5,481 baht), following by other crops and vegetables farm's (4,000 baht). The southern region's earn the most from permanent crops predominantly rubber (6,619 baht), while make the lowest income from other crops (3,437 baht).

4.2 Variables and descriptive statistics

The dependent variables is the net earnings or net profit (EARN), which is equal to the total household earnings minus the operating costs of all variable inputs, such as fertilizer, seeds, energy, and hired labor.

Farm land (LAND) is used as a proxy of capital, used in farming. Labor is measured as the number of effective workers in the household (LABOR). Hired labor includes number of hired workers in soil preparation (SOIL), planting (PLANT), harvesting (HARVEST), and others (OTHERS).

Human capital factors include both education and health factors. Schooling variables are included in the regression to capture their effects on earnings. The number of years of schooling completed is well recognized as an indicator of educational attainment (Strauss and Thomas, 1998; Schultz, 2004). It is expected to have a positive effect on a household ability to run a farm. Our hypothesis is tested by including the number of household head's years of schooling attained (SCHOOLH) and its quadratic term of education (SCHOOLH2), controlling for possible nonlinear relationships. The head of household education level may influence farm management decision and eventually affects income.

Unlike education, a measurement of health is difficult because there is no consensus in the literature to specify a good indicator for health. Although Strauss and Thomas (1998) indicated that the most widely used health indicator in the U.S. in the empirical literature is General Health Status, individual self-evaluation of health status suffers from subjectivity. In any case, the data provide no general measure of heath status. Instead this paper uses health consumption as measured by the number of days that any member stayed in a hospital during the past 12 months (HCON). One problem is a possible omitted variable bias because there is no direct measure of health status. That is healthier individuals will need less health consumption to achieve a given level of health, but more health consumption will improve health status. This study assumes farm output, thus household earnings depends on health status (*H*). However, since the model also assume health status depends on health consumption (*h*) and farm labor (*L),* we lose only the intermediate effect of health status and can still find the marginal contribution of *h* and *L* to earnings while the (average) effect of *H* is subsumed in the constant and error terms. The health consumption variable represents an investment in health care controlling a given level of health status. This paper includes a quadratic form of the health consumption variables for a possible non-linearity with respect to income (HCON2). Moreover, since labor spent by household members for farm work may affect their health status, the analysis takes into account for an interaction term combining the health consumption variable (number of hospital days) and household farm labor (HCONxLABOR).

Household characteristics variables include head of household age, gender, and number of children up to age 5. Age is a commonly used measure for general work experience. The age of head of households (AGEH) is included in the model to test whether that head of household households' experience positively affects household farm income. Head of household's gender (MALE) is included to reflect differences in decision making between men and women. The dummy variable indicates 1 if men, and zero otherwise. Having small children and infants requires care-taking time from household members indicating that even if the same number of adults are available, the time allocated to farm labor may be less.

Region dummy variables (NORTH, CENTRAL, NORTHEAST, and SOUTH) are also included to capture heterogeneity across regions. Since agricultural enterprises are classified to six types, this paper includes dummy variables for each type of farm: rice farming (RICE); other crops farming such as maize, sugarcane, etc. and vegetables (CROP): fruits, permanent crops, and shrubs farming (FRUIT); livestock (LIVESTOCK); poultry (POULTRY); and forestry, hunting, and agricultural services (FORESTRY). The paper excludes FORESTRY from the regression.

As noted above, to deal with possible endogeneity, the regression includes instrumental variables to capture the effect of endogenous variables. The variables that meet the criteria of good instruments based on the available survey data are chosen. Instrumental variables in this model include the average years of schooling of household (SCHOOLAVG). An instrument for health consumption variable is medical supplies and medical services (for outpatient and inpatient services) expenditure (MEDEXP). Currie and Madrian (1999) indicated that ideal instruments should proxy the price of obtaining the human capital. Since this paper assumes that most agricultural families, which generally earn lower income relatively to other profession, prefer public hospitals where prices are controlled by the Ministry of Public Health, medical expenditures is used as an instrument for the health consumption variable. Table 2.5 shows the correlation matrix among household earnings, endogenous variables, and instrumental variables.

Table 2.6 reports the descriptions and descriptive statistics of variables used in the analysis. Average agricultural household income is approximately 3,507 baht per month ranging from no income to 115,968 baht.[6] The average household size is 3.86 and average age is 36 years. The average size of farm land holding is 21.64 rai per household (approximately 8.55 acres), ranging from no land to 249 rai. Almost 80% of head of households are men. There are 3.13 household workers on average in each farm household, ranging from two persons to seven persons. Moreover, households hired more than five workers to work on harvesting, three to four planting workers, half labor for soil preparation, and one person for other farm work during last year. The average years of education attained of each household is 4.8 years (upper elementary school level), ranging from no education to 16 years.

Each agricultural household spends approximately 145 baht per month for medical care. While households usually sacrificed half a day for sickness each month, on average they spent one day in a hospital during the past year.

5. Empirical results

[6] It is possible that some households may experience zero income if they did not sell any products during the last month.

Table 2.7 reports the results of the ordinary least squares (OLS) and the two-stage least squares (2SLS) regressions. The study applied a Box-Cox transformation in which both the dependent and independent variables are subjected to different power transformations. The estimate of the θ and λ power parameters from (24) were 0.276 and 0.236, respectively. The regressions examined the effect of human capital such as health care and education, the interaction term of health and household labor, and capital and labor inputs, for example on farm net earnings.

In order to test for possible endogeneity, this paper applied a Wu-Hausman test with the null hypothesis that regressors are exogenous. As shown in Table 2.7, we cannot reject the null hypothesis implying that explanatory variables were not endogenous in the regressions. This implied that OLS estimates were more efficient than 2SLS. In fact, both 2SLS and OLS regressions were rather similar. The estimates of inputs, education, age, type of enterprise dummy variables, and region dummy variables were statistically significant at a 0.1 level or higher. Of these estimates, although all signs of coefficients were similar, most effects from the 2SLS were higher relative to the OLS's.

Next, the author analyzed whether health care may depend on returns to schooling by interacting the schooling completion variable (SCHOOLH) with the health consumption variable (HCON). Compared column (2) of Table 2.7 with Table 2.8, we tested this unrestricted model against a model with no interaction variable employing a likelihood ratio test. The results showed that the model was not significant with a LR test as the log likelihood ratio (0.8) was less than a chi-square statistic with one degree of freedom ($\chi^2_{(0.95;1)} = 3.84$). Thus, there was no interaction effect of health and education on earnings. This implied that the two types of human capital were not substitute for one another, nor did they complement one another. In fact, the regression results on both columns were almost identical.

The regression results showed that the estimated coefficients of household-owned capital and household labor input were statistically significant at a 0.01 level or higher. The elasticities of income with respect to capital and labor, calculated from equation (24), were 0.55 and 0.39, respectively. This implied that a 10% increase in capital of farm production raised earnings by 5.5%, while a 10% increase in labor increased income by 3.9%. It is known that Thai agricultural sector is labor-intensive, the findings supported that using more capital contributed to an increase in farm income. For hired labor variables, the regressions showed that most elasticities of earnings with respect to hired workers appeared to have a negative impact on farm earnings as expected at a 0.05 level of significance. Both elasticities with respect to hired labor hours in soil preparation and planting jobs were –0.03. However, the estimated elasticity for hired labor in harvesting was positive. Harvesting work always requires many workers other than family labor in a rapid harvesting season. It is possible that investing in hired labor in harvesting may contribute to higher income because it is more cost efficient.

A key finding from the regression indicated that the human capital factor, which was education, appeared to increase farm income since the estimated coefficient was significant at a 0.01 level. The elasticity of earnings with respect to number of head of households' years of schooling attained (SCHOOLH) was 0.11 suggesting that a 10% increase in investing in education may raise farm income by 1.1%. The quadratic term of education was negative subjecting to diminishing return, but not statistically significant. However, the estimates for the other human capital factor -- health variables, which were the length of hospital stay (HCON) and its quadratic term (HCON2), and the interaction term between HCON and LABOR were not statistically significant at a 0.05 level. The results implied that the health care consumption factor did not determine farm earnings. However, the sign of elasticities of the health variable and its quadratic term were positive and negative, respectively, as expected.

The head of household age (AGEH) appeared to have a positive impact on farm income at a 0.1 level since the elasticity of income with respect to the head of household age was 0.01. The results confirmed the hypothesis that head of households with higher experience were more likely to generate higher income. However, the estimated coefficient of gender dummy variable (MALE) was not statistically significant suggesting that the head of households' gender was not likely to affect farm earnings. Moreover, the results showed that having more small children and infants in a household tend to lower income as expected. The elasticity of income with respect to number of children was –0.02.

All types of enterprise dummy variables were statistically significant at 0.05 level or higher. Compared with the (excluded) forestry and agricultural services enterprise, rice farming household appeared to achieve the lowest level of income. Forestry and agricultural services households tended to make the most income, followed by Livestock households and Fruits and Permanent crops households.

Compared to the (excluded) southern region, the findings showed that farm households in the central region were more likely to generate the highest income, followed by the southern region. Households in the northeastern region appeared to make the least money. The results supported the fact that climate, soil conditions, and water conditions in the central region were rather superior compared to the rest of the country. On the other hand, the agricultural sector in the northeast, in particular, suffered with infertile soil and seasonal drought.

5.1 Disaggregated analysis

Each type of agricultural enterprise was different by its nature. The operational costs of various types of farms contributed to various levels of farm profit. Therefore, the opportunity costs of not doing market work for each farm type were diverse. In terms of farm labor, the number of workers needed for each farm group was also different depending on how labor-intensive the farm was. Moreover, health is also an important factor for households to decide how much time they should contribute to farm enterprises. It is possibly riskier for workers to be exposed to chemicals on one farm compared to others. Thus, in addition to the pooled regression, it was useful to estimate (25) by the each six enterprises separately.

Table 2.9 reported regression results by type of enterprise. Column (1) showed a regression of rice enterprise households. Compared to the pooled regression of Table 2.8, the result of the rice farming showed that estimated coefficients signs for most variables remain similar and statistically significant at a 0.1 level of significance or higher. The findings could come from the fact that rice farming households account the largest percentage of the sample (47%). Estimates of land and labor were positively significant while most of the hired labor coefficients are negative. The estimates showed that the elasticity of earnings with respect to capital and labor are 0.65 and 0.34 respectively. Moreover, the estimated coefficients of most hired labor variables were negatively statistically significant which was consistent with the pooled regression.

The coefficient estimated for the head of households' years of education attained was statistically significant at a 0.01 level. The elasticity of income with respect to education was 0.24 suggesting a positive relationship to household income. Although the estimates of health consumption variable (HCON) of all types of enterprise were not statistically significant at a 0.1 level, the estimated coefficient of rice farming was statistically positive at a 0.12 level. Since the quadratic term of health consumption (HCON2) and HCONxLABOR were not statistically significant, it had no predictive power on income. Thus, given a fixed level of health status, an increase in health care utilization improved health status, and household earnings. The elasticity of income with respect to health consumption was 0.03 implying that a 10% increase in health consumption raised household earnings by 0.3%. It should be noted that health consumption did not appear to be a determinant of income on other types of enterprises. For the rice farming enterprise, most households did not depend on high-tech equipments but on physical labor.

In addition, geographic location also had a significant impact on rice farm income. Farm households in the central region appeared to make the most money. In Thailand, rice farms were cultivated in the central region where environmental factors were the most appropriate for rice. On the other hand, rice households in the southern region tended to make the least money.

Column (2) showed the results of other crops and vegetables households. Although there was no statistical evidence that the education factor affects earnings, the quadratic term of health consumption and the interaction variable between health and household labor were statistically significant at a 0.1 level or higher. The negative sign of HCON2 indicated that health input was subjected to diminishing returns. The positive sign of the coefficient of HCONxLABOR indicated that healthy farmers tended to provide more farm labor because health consumption reflected health status. In other words, farm labor and health care were complements in production in that farm labor was more valuable if health consumption increased. In terms of locations, crops and vegetables farm households in the central region tended to achieve the highest level of income, following by the northern region.

The regression results of households that grow fruits, permanent crops, and shrubs are shown in column (3). Only the estimated coefficients of physical inputs and region dummy variables of these households were statistically significant. The elasticities of income with respect to capital and labor inputs were 0.53 and 0.88, respectively. Human capital factors did not appear to affect household income. Unlike other enterprises, farm households located in the southern region were more likely to earn the most followed by the central and the northeast regions, respectively.

The regressions from livestock households (column (4)) indicated that only capital input, hired labor (for soil preparation), and location affect farm earnings. The elasticity of income on the capital inputs (0.45) suggested that livestock households should invest more in capital. In addition to physical inputs, geographical areas appeared to affect farm earnings. Households located in the central were more likely to contribute the highest income while the northern tended to make the lowest.

In terms of poultry households, there was statistical evidence that education and health affected earnings since SCHOOLH was statistically significant at a 0.05 level. The elasticity of income with respect to the number of head of household years of education attained was 0.44, suggesting a positive relationship to earnings. The relationship of health and labor was calculated as $\frac{\partial h}{\partial L}$ = $\frac{\partial \Pi / \partial L}{\partial \Pi / \partial h}$. Suppose that $\Pi = ah + bL + c(hxL) + dh^2$, then

$\partial h / \partial L$ = $\frac{b + ch}{a + cL + 2dh}$ = 2.47 at the mean values for *h* and *L*. The change in *h* with respect to *L* was positive implying that farm labor was more valuable if health consumption increases because farm labor and health care were complements in production. Furthermore, using an implicit function rule, $\partial H / \partial L$ was negative which was as what the model assumed.

In contrast to the pooled regression results, the disaggregated regression indicated that a poultry head of households' age was negatively correlated with earnings since the age elasticity of income was negative. A conclusion could be drawn that in the poultry sector, formal education was more important than experience. In addition, none of the region dummy variables estimated coefficients were statistically significant implying that farm income did not rely on geographical areas.

The regression results of forestry, hunting, and agricultural households found that the elasticities of capital and labor inputs were –0.63 and 4.21, respectively. The results supported the hypothesis that labor was the major contribution of household income while land did not appear to be an important input for households in this enterprise. In fact, the sample showed that almost 80% of these households had no owned land. In terms of household characteristics, there was statistical evidence that households with more small children tended to acquire lower income. Unlike households with other enterprises, households in the central region tend to make the least income from forestry and hunting while households from the south tend to make the most.

6. Conclusions

The study examines the effect of healthcare utilization on agricultural household earnings using the data from Thai Socio-Economic Survey in 2002. The sample includes 1,871 agricultural households categorized by six enterprises. Since a log transformation was not possible with some zero values, this studies applied a Box-Cox transformation of both explanatory variables and the dependent variables. The Wu-Hausman test confirmed that explanatory variables were not endogenous, thus OLS estimates are consistent. The elasticities of income with respect to dependent variables were estimated following Zarempka (1968).

The empirical results showed that human capital factors, education and health, were neither substitutes nor complements in production.. While the findings supported the human capital theory that education increased income, the effect of healthcare consumption on income was less clear. The elasticity of earnings with respect to number of head of households' years of schooling attained was positive for the pooled regression as well as most regressions by farm types. The results showed that, overall, farm income may increase by 1.1% when any household raised an investment in household heads' education by 10%. In terms of healthcare, only the estimated coefficient of health variable (length of stay) for rice households was statistically significant at a 0.12 level with a health elasticity of income of 0.03.

The other main determinants for farm earnings included capital and labor inputs, head of households' age, number of small children in each household, geographic locations, and types of farm. The regression results showed that capital and household labor inputs were positively related to farm earnings on both pooled regression and regressions by types of enterprises as expected. The elasticities of farm earnings with respect to most hired labor were negative except the one from harvesting labor. The results also found that farm households in the central region appeared to make the most income, while those in the northeastern region tended to make the least.

In a disaggregated analysis, the regression results by rice farm were consistent with the pooled regression results. Rice farm income appeared to be the only sector that was affected by the healthcare consumption. The other key finding was that farm labor and health consumption in poultry households were complements in production. The findings showed a negative relationship between health status and farm labor as expected. However, the heads of households' gender did not contribute to changes in income in any regression.

REFERENCES

Abdulai, A. and Eberline, R. (2001), "Technical efficiency during economic reform in Nicaragua: Evidence from farm household survey data", *Economic Systems* 25 (2), 113-125.

Baldwin, M., Zeager, L., and Placco, P. (1994), Gender differences in wage losses from impairments", *Journal of Human Resources* 29, 865-887.

Becker, G. S. (1965), "A Theory of the Allocation of Time," *Economic Journal 75*, 493–517.

Benjamin, C. and Guyomard, H., 1994. Off-farm work decisions of French agricultural households. In: Caillavet, F., Guyomard, H., Lifran, R. (Eds.), Development in Agricultural Economics: *Agricultural Household Modeling and Family Economics*. Elsevier, Amsterdam.

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

Bolin, K., Jacobson, L., Lindgren, B. (2001), "The Family as the Health Producer--When Spouses Are Nash-Bargainers", *Journal of Health Economics* 20 (3), 349-62

Bolin, K., Jacobson, L., Lindgren, B. (2002), "The Family as the Health Producer--When Spouses Act Strategically", *Journal of Health Economics* 21 (3), 475-95

Box, G. and Cox, D. (1964), "An Analysis of Transformations", *Journal of the Royal Statistical Society*, Series B, 211-64.

Chaudhri, D. P. (1979) *Education, Innovations and Agricultural Development: A Study of*

Northern India (London: Croom Helm Ltd. for the International Labor Organization).

Chirikos, T. N. and Nestel, G. (1985), "Further evidence on the economic effects of poor health", *The Review of Economics and Statistics* 67, 61-69.

Contoyannis, P. and Jones, A. (2004), "Socio-economic status, health and lifestyle", *Journal of Health Economics* 23, 965-995

Contoyannis, P. and Rice, N. (2001), "The Impact of Health on Wages: Evidence from the British Household Panel Survey", *Empirical Economics* 26, 599-622.

Currie, J. and Madrian, B. (1999), "Health, Health Insurance and the Labor Market", in Ashenfelter, O. and Card, D. (1999) ed., *Handbook of Labor Economics*, Volume 3, 3309-3416.

Deolalika, A. B. (1988), "Nutrition and Labor Productivity in Agriculture: Estimates for Rural South India", *Review of Economics and Statistics* 70 (3), 406-413.

Dey, M., Paraguas, F. Bimbao, G., and Regaspi, P. (2000), "Technical efficiency of Tilapia Growout Pond Operations in The Philippines", *Aquaculture Economics and Management* 4(1/2), 33-47.

Duleep, H. O. (1986), "Measuring the Effect of Income on Adult Mortality Using Longitudinal Administrative Record Data", *Journal of Human Resources* 21, 238-51

Glewwe, P. and Jacoby, H. (1995), "An Economic Analysis of Delayed Primary School Enrollment in a Low Income Country: The Role of Early Childhood Nutrition", *The Review of Economics and Statistics* 77 (1), 156-169

Grossman, M., (1972a), "On the Concept of Health Capital and the Demand for Health", *Journal of Political Economy* 80, 223-255.

Grossman, M., (1972b), "The correlation between Health and Schooling", in Terleckyj, N.E. ed., *Household Production and Consumption*, (Columbia University Press, New York), 147-211.

Grossman, M. (2001), "Household Production and Health", *Review of Economics of the Household* 1, 331-342

Grossman, M. and Benham, L. (1974), Health, hours and wages. Macmillan and Co, London. In M. Perlman (ed), The economics of health and medical care.

Grossman, M. and Kaestner, R., (1997). Effects of education on health. In: Behrman, J.R., Stacey, N. (Eds.), The Social Benefits of Education, University of Michigan Press, Ann Arbor, MI, 69–123.

Gurgand, M. (2003), Farmer education and the weather evidence from Taiwan (1976-1992), *Journal of Development Economics* 71, 51-70

Havemen et al. (1994), "Market work, wages and men's health", Journal of Health Economics 13, 162-182.

Jacobson, L. (2000), "The family as producer of health: an extended Grossman model", *Journal of Health Economics* 19 (5), 611-37.

Jamison, D.T., Lau, L.J. (1982). *Farmer Education and Farm Efficiency*, Johns Hopkins Univ. Press, Baltimore.

Jensen, R. and Richter, K. (2001), "Understanding the Relationship between Poverty and Children's Health", *European Economic Review* 45 (4-6), 1031-39.

Johnson, W. and Lambrinos, J. (1985), "Wage discrimination against handicapped men and women", Journal of Human Resources 20, 264-277.

Lleras-Muney, A. (2005), "The Relationship Between Education and Adult Mortality in the United States", *Review of Economic Studies* 72, 189–221.

Matshe, I. and Young,T. (2004)," Off-farm labor allocation decisions in small-scale rural households in Zimbabwe", *Agricultural Economics* 30 , 175-186.

Menchik, P.L. (1993), "Economic Status as a Determinant of Mortality among Black and White Older Men: Does Poverty Kill?, *Population Studies* 47, 427-436.

Mincer, J. (1958), *Schooling, Experience, and Earnings*, New York: Columbia University Press.

Mullahy, J. and Sindelar, J. (1995), "Gender differences in labor market effects of alcoholism", *American Economic Review* 81, 162-165.

Osula, D. O. A. and Adebisi, O. (2001), "Testing the stability of travel expenditures in Nigeria", *Transportation Research* Part A 35, 269-287.

Patrick, G. and Kehrberg, E. (1973),"Costs and Returns of Education in Five Agricultural Areas of Eastern Brazil," American Journal of Agricultural Economics 55, 145-154.

Pitt, M. and Rosenzweig, M. (1985), "Health and Nutrient Consumption Across and Within Farm Households", *The Review of Economics and Statistics* 67 (2), 212-223

Pitt, M. and Rosenzweig, M. (1986), "Agricultural Prices, Food Consumption and the Health and Productivity of farmers," in Singh, I., Squire, L., Strauss, J. (1986), *Agricultural Household Models: Extensions and Applications*. John Hopkins University Press, Baltimore.

Pitt, M. and Rosenzweig, M. (1990), "Estimating the Intrahousehold incidence of illness:

Child Health and Gender-Inequality in the Allocation of time”, *International Economic Review* 31 (4), 969- 989.

Rosenzweig, M.R. (1995), “Where has all the education gone?”, *American Economic Review* 85 (2), 153-158.

Rosenzweig, M. and Schultz, T. P. (1983), “Estimating a Household Production Function: Heterogeneity, the Demand for Health Inputs, and Their Effects on Birth Weight”, *Journal of Political Economy* 91 (51), 723-746.

Savin, N.E. and White, K.J., (1978), “Estimation and testing for functional form and autocorrelation: a simultaneous approach”, *Journal of Econometrics* 8, 1-12.

Schultz, T. (1989), “Investing in People: Schooling in Low Income Countries”, *Economics of Education Review* 8 (3), 219-23.

Schultz, T. P. (2004), “Health Economics and Applications in Developing Countries”, *Journal of Health Economics* 23 , p. 637-641.

Singh, I., Squire, L., Strauss, J. (1986), *Agricultural Household Models: Extensions and Applications*. John Hopkins University Press, Baltimore.

Strauss, J. (1982), “Determinants of Food Consumption in Sierra Leone: Application of the Quadratic Expenditure System to the Consumption Leisure Component of a Household-Firm Model”, *Journal of Development Economics* 11, 327-353.

Strauss, J. and Thomas, D. (1998), Health, Nutrition, and Economic Development, Journal of Economic Literature 36 (2), 766-817.

Sundberg, G. (1996), “Essays on Health Economics”, PhD thesis, Department of Economics, Uppsala University.

Taylor, J. E. (2003), "Agricultural Household Models: Genesis, Evolution, and Extensions", *Review of Economics of the Household 1* (1), 33-58

Zarembka, P. (1968), "Functional Form in the Demand for Money", *Journal of American Statistics Association* 63, 502-11.

TABLES

Table 2.1: Distribution of data by type of household farm enterprise

TYPE OF ENTERPRISE	Full SES (1)		Sample (2)	
	Number of households	Percent	Number of households	Percent
Rice farming	4,663	49.1	872	46.61
Other crop farms and vegetables	1,702	17.9	334	17.85
Fruits, permanent, crops and shrubs	1,634	17.2	331	17.69
Livestock	1,177	12.4	262	14
Poultry	165	1.7	43	2.3
Forestry, hunting, agricultural service	156	1.6	29	1.55
Total	9,497	100	1,871	100

Table 2.2: Distribution of data by region

REGION	Full SES (1)		Sample (2)	
	Number of households	Percent	Number of households	Percent
Central	1,814	19.1	294	15.71
North	2,573	27.09	427	22.82
Northeast	3,706	39.02	835	44.63
South	1,404	14.78	315	16.84
Total	9,497	100	1,871	100

Table 2.3: Head of households' education level

Level of education	Full SES		Sample	
	Number of households	Percent	Number of households	Percent
No education	719	7.57	123	6.57
Grade 4	6,553	69	1,297	69.32
Grade 6 (upper elementary school)	1,245	13.11	254	13.58
Grade 9	499	5.25	93	4.97
Grade 12 (high school)	205	2.16	47	2.51
Bachelor degree or higher	276	2.91	57	3.05
Total	9,497	100	1,871	100

Table 2.4: Average income by region and household enterprise

REGION	Rice farming	Other crop farms and vegetables	Fruits, permanent, crops and shrub crops	Livestock	Poultry	Forestry, hunting, agricultural service	Total
Central	5,849.27	7,998.25	7,146.51	17,684.81	9,477.80	6,835.67	8,046.40
North	2,260.77	3,998.26	2,362.39	5,480.94	1,421.92	1,324.82	3,102.41
Northeast	947.44	3,436.91	2,903.11	1,955.56	2,496.14	480	1,490.49
South	1,545.63	2,352.29	6,618.58	4,107.70	3,241.79	4,279.20	5,165.05
Total	1,869.01	4,441.23	6,050.47	4,618.71	4,062.77	2,112.86	3,507.18

Table 2.5: Correlation matrix

Variables	EARN	HCON	SCHOOLAVG	MEDEXP	SCHOOLH
EARN	1				
HCON	0.0126	1			
SCHOOLAVG	0.1081	-0.0172	1		
MEDEXP	0.0658	0.7034	0.0294	1	
SCHOOLH	0.0534	-0.0279	0.5535	0.0539	1

Table 2.6: Variable description and descriptive statistics
(N=1871)

Variable	Description	Mean	Std. errors
Dependent variables			
EARN	Monthly net earnings (baht)*	3,507.18	6,890.221
Independent variables			
LAND	Farm land (rai)**	21.15	23.17
LABOR	Number of effective workers in the households	3.13	0.89
SOIL	Number of hired workers (for soil preparation)	0.61	1.19
PLANT	Number of hired workers (for planting)	3.69	5.87
HARVEST	Number of hired workers (for harvesting)	5.38	7.42
OTHERS	Number of hired workers (for others)	0.98	2.67
AGEH	Head of households' age in years	51.5	13.06
MALE	1 if head of household is male; 0 otherwise	0.78	0.41
CHILD	Number of infants and children (<5 years)	0.35	0.57
SCHOOLH	Number of household head's years of schooling attained	4.7	2.79
SCHOOLH2	Number of household head's years of schooling attained squared	29.84	44.46
HCON	Number of days members stayed in a hospital during the past 12 months	0.86	3.26
HCON2	Number of days members stayed in a hospital during the past 12 months squared	11.38	122.9
RICE	1 if rice farming; 0 otherwise	0.47	0.50
CROP	1 if other crops farming; 0 otherwise	0.18	0.38
FRUIT	1 if fruits, permanent crops, and shrub crops; 0 otherwise	0.18	0.38
LIVESTOCK	1 if livestock; 0 otherwise	0.14	0.35
POULTRY	1 if poultry; 0 otherwise	0.02	0.15
FORESTRY	1 if forestry and agricultural services; 0 otherwise	0.015	0.12
CENTRAL	1 if the central region; 0 otherwise	0.16	0.36
NORTH	1 if the northern region; 0 otherwise	0.23	0.42
NORTHEAST	1 if the northeastern region; 0 otherwise	0.45	0.50
SOUTH	1 if the southern region; 0 otherwise	0.17	0.37
Instrumental variables			
MEDEXP	Medical supplies and services expenditure (baht/month)	144.61	614.58
SCHOOLAVG	Average years of schooling of the household members	4.82	2.04

Note : * 1 baht is approximately 0.025 USD
** 1 rai is equivalent to 0.395 acres.

Table 2.7: Estimates of 2SLS and OLS regressions (Box-Cox transformation)

Dependent variable = EARN

Variables	2SLS (1)		OLS (2)	
	Coef.	t	Coef.	t
HCON	0.30	0.57	0.09	0.69
HCON2	-0.01	-0.66	-0.01	-0.55
SCHOOLH	0.75	1.85*	0.72	3.43***
SCHOOLH2	0.004	0.04	-0.01	-0.08
HCONxLABOR	0.01	0.68	0.005	0.56
HCONxSCHOOLH	-0.00003	-0.11	-0.0001	-0.25
MALE	0.12	0.19	0.12	0.19
AGEH	0.74	1.40	0.04	1.68*
CHILD	-0.30	-2.04**	-0.27	-2.08**
LAND	2.55	17.47***	2.55	17.62***
LABOR	2.79	3.31***	2.85	3.52***
SOIL	-0.34	-2.46**	-0.34	-2.51**
PLANT	-0.22	-2.02**	-0.23	-2.08**
HARVEST	0.22	2.10**	0.22	2.18**
OTHERS	-0.09	-0.75	-0.08	-0.70
CENTRAL	5.56	5.62***	5.51	5.61***
NORTH	-1.39	-1.38	-1.47	-1.48
NORTHEAST	-7.04	-7.25***	-7.11	-7.40***
RICE	-15.40	-6.84***	-15.39	-6.85***
CROPS	-6.37	-2.75***	-6.37	-2.75***
FRUITS	-5.51	-2.38**	-5.50	-2.38**
LIVESTOCK	-5.26	-2.31**	-5.26	-2.31**
POULTRY	-7.19	-2.63***	-7.06	-2.61***
Constant	16.09	2.92***	18.41	6.29***
Observations		1871		1871
R-squared		0.40		0.41
Wu Hausman chi-sq test (column (1) and (2))			Chi-sq(2)= 0.10	p-value = 0.9
Instrumental variables = SCHOOLAVG and MEDEXP				

Note: *** significant at 1%; ** significant at 5%; * significant at 10%

Table 2.8: Estimates of OLS regressions (Box-cox transformation)

Variables	Coef.	t	Elasticity
HCON	0.09	0.72	0.01
HCON2	-0.01	-0.54	-0.001
SCHOOLH	0.73	3.47***	0.11
SCHOOLH2	-0.003	-0.04	-0.002
HCON*LABOR	0.004	0.51	0.001
HCON*SCHOOLH			
MALE	0.11	0.18	0.01
AGEH	0.75	1.72*	0.01
CHILD	-0.27	-2.04**	-0.02
LAND	2.55	17.62***	0.55
LABOR	2.84	3.50***	0.39
SOIL	-0.34	-2.51**	-0.03
PLANT	-0.23	-2.1**	-0.03
HARVEST	0.22	2.19**	0.04
OTHERS	-0.08	-0.7	-0.01
CENTRAL	5.52	5.62***	
NORTH	-1.46	-1.48	
NORTHEAST	-7.10	-7.39***	
RICE	-15.38	-6.85***	
CROPS	-6.36	-2.75***	
FRUITS	-5.50	-2.38**	
LIVESTOCK	-5.25	-2.31**	
POULTRY	-7.04	-2.6***	
Constant	15.47	3.86***	
Observations		1871	
R-squared		0.41	
Likelihood ratio test (Table 2.7 (2) V.S. Table 2.8)	Prob > chi2 = 0.80		

Note: *** significant at 1%; ** significant at 5%; * significant at 10%

Table 2.9: Regression results and elasticities by type of enterprise

Variables	Rice (1)		Other crops/vegetables (2)		Fruits and Shrubs (3)		Livestock (4)		Poultry (5)		Forestry (6)	
	Coef.	elasticity	Coef.	elasticity	Coef.	elasticity	Coef.	elasticity	Coef.	elasticity	Coef.	elasticity
HCON	0.34†	0.03	-0.24	-0.02	-0.15	-0.02	0.37	0.04	-0.48	-0.05	-0.25	-0.03
HCON2	0.01	0.002	-0.09**	-0.02	0.08	0.02	-0.04	-0.01	0.48***	0.09	-0.34	-0.06
SCHOOLH	1.6***	0.24	0.23	0.04	-0.14	-0.02	-0.21	-0.03	2.91**	0.44	0.57	0.09
SCHOOLH2	0.1	0.02	-0.17	-0.04	-0.18	-0.04	-0.2	-0.05	0.53	0.12	-0.55	-0.13
HCONxLABOR	-0.02	-0.004	0.08*	0.02	-0.05	-0.01	0.02	0.005	-0.38***	-0.08	0.38	0.08
MALE	0.29		0.6		-0.36		1.42		7.98		0.54	
AGAH	1.77***	0.47	1.65*	0.44	-0.74	-0.20	0.6	0.16	-7.26*	-1.94	-2.49	-0.66
CHILD	-0.47***	-0.04	0.05	0.00	-0.06	-0.005	0.54	0.04	-1.04	-0.09	-3.35**	-0.27
LAND	3.02***	0.65	2.4***	0.52	2.44***	0.53	2.1***	0.45	3.53	0.76	-2.92**	-0.63
LABOR	2.45**	0.34	-0.39	-0.05	6.43***	0.88	1.98	0.27	27.15	3.73	30.63***	4.21
SOIL	-0.25	-0.02	-0.73***	-0.07	-0.25	-0.02	-0.76*	-0.07	4.43	0.42	2.52	0.24
PLANT	-0.32**	-0.05	0.17	0.02	-0.23	-0.03	-0.41	-0.06	-3.47**	-0.50	-0.46	-0.07
HARVEST	0.15	0.02	0.3	0.05	0.87***	0.14	0.06	0.01	-3.16	-0.49	2.18	0.34
OTHERS	-0.26*	-0.03	-0.27	-0.03	0.56*	0.06	0.09	0.01	-2.06	-0.22	-0.79	-0.08
CENTRAL	12.9***		9.26***		-2.73*		18.1***		0.16		-23.31*	
NORTH	4.93***		4.51*		-10.83***		6.22**		3.72		-16.49***	
NORTHEAST	-0.6		0.93		-8.68**		-2.81		3.12		-20.45***	
Constant	-15.95***		1.91		25.39***		11.78		8.79		15.88	
Observations		872		334		331		262		43		29
R-squared		0.37		0.36		0.37		0.3		0.54		0.65

Note: *** significant at 1%; ** significant at 5%; * significant at 10%; † significant at 12%

APPENDIX

Appendix 2.1: A profit maximization model

Assuming households are subsistence households, let a household be an income maximizer.

$$\text{Max } \Pi \equiv P*[Q(L,H;E,K)-F] + W(E)*N$$

$$S.T. \quad \Omega = M + L + N + h$$

Where $H = h(h, L; \mu)$, $Q = Q(L, H; E, K)$, and $W = W(E)$

Π = net income

Q = Farm production

L = Household farm labor

W = Wage for non-farm work

P = Price of farm goods

N = Market (non-farm) work

H = Health status

h = Health consumption

E = Education

Ω = Total time available

M = Leisure time

μ = (predetermined) health status and external factors

Using similar assumptions as in Section 2, the household problem is given by the Lagrangian function:

$$\tilde{L} \equiv P*[Q(L,H;E,K)-F] + W(E)*(\Omega - M - L - h)$$

$$\frac{\partial \tilde{L}}{\partial L} \qquad P_F\left(\frac{\partial Q}{\partial L}+\frac{\partial Q}{\partial H}\frac{\partial H}{\partial L}\right)-W(E) \;=\; 0 \tag{a1}$$

$$\frac{\partial \tilde{L}}{\partial h} \qquad P_F\left(\frac{\partial Q}{\partial H}\frac{\partial H}{\partial h}\right)-W(E) \;= 0 \tag{a2}$$

Using (a1) and (a2) gives the condition that the net marginal value product of labor is equal to their marginal value product of health consumption. The terms on the left hand side denote the direct effect of marginal value product of labor and its health value effect from labor, respectively.

$$P\;\frac{\partial Q}{\partial L}+P\frac{\partial Q}{\partial H}\frac{\partial \bar{H}}{\partial L} \quad = \quad P\left(\frac{\partial Q}{\partial H}\frac{\partial H}{\partial h}\right) \tag{a3}$$

Equation (a3) indicates that the marginal value product of health consumption may be smaller than the marginal return of labor because of the health value effect.

Then, the model examines an effect of education on health consumption by using a total differentiation on equation (a3), and obtain:

$$\frac{dh}{dE}: \quad \frac{\frac{\partial W}{\partial E}-P\frac{\partial H}{\partial L}\frac{\partial^2 Q}{\partial H\partial E}-P\frac{\partial^2 Q}{\partial L\partial E}}{Z} \tag{a4}$$

where Z = $PQ_{LH}H_h+PH_LQ_{HH}H_h+PQ_HH_{Lh} > 0$. Subscripts indicate (partial) derivatives.

Since Z is greater than 0, it yields two conditions:

Condition (1): $P\frac{\partial \bar{H}}{\partial L}\frac{\partial^2 Q}{\partial H\partial E}+P\frac{\partial^2 Q}{\partial L\partial E}<\frac{\partial W}{\partial E}$, then $\frac{dh}{dE}>0$

The term $P\frac{\partial^2 Q}{\partial L\partial E}$ on the left hand side shows how changes in education affect the marginal revenue product of farm labor. $P\frac{\partial \bar{H}}{\partial L}\frac{\partial^2 Q}{\partial H \partial E}$ indicates how changes in education and farm labor affect the marginal revenue product of health. The right hand side term represents how education affects wage or the opportunity cost of non-farm work. Condition (1) shows when the net effect of education on marginal returns of health and farm labor are less than the effect of education on wage, a change in health consumption *(h)* with respect to education *(E)* is positive.

Condition (2): $P\frac{\partial \bar{H}}{\partial L}\frac{\partial^2 Q}{\partial H \partial E} + P\frac{\partial^2 Q}{\partial L \partial E} > \frac{\partial W}{\partial E}$, then $\frac{dh}{dE} < 0$

Condition (2) indicates that when the net effect of education on marginal returns of health and farm labor are more than the effect of education on wage, then households with more education tend to have lower health consumption compared to households with lower level of education.

CHAPTER 3

IMPACTS OF HEALTH INSURANCE ON HOUSEHOLD EXPENDITURE PATTERNS IN THAILAND: A QUADRATIC ALMOST IDEAL DEMAND SYSTEM (QUAIDS) APPROACH

Abstract

This essay explores how the Universal Coverage (UC) program affects the allocation of household expenditures on consumption goods (i.e., housing, food, etc.) in Thailand. A Quadratic Almost Ideal Demand System (QUAIDS) model developed by Banks, Blundell, and Lewbel (1997) is used incorporating with a two-step approach introduced by Shonkwiler and Yen (1999). The programming was done on GAUSS 7.0 in order to solve the nonlinear least squares problems applying the Gauss-Newton optimization algorithm. The data consists of approximately 200,000 members of 24,586 and 34,607 households from Thai Socio Economic Survey (SES) in 2000 and 2002, respectively. The results of uncompensated own price elasticities indicated that the demand for food, clothing and miscellaneous goods, health and medical care, and housing became more elastic in 2002 after UC was introduced.

1. Introduction

Thailand adopted the Universal Health Coverage (UC) program in 2001. Because of this program, 75% of the populations, who were uninsured or were not covered by any formal insurance system, now have health insurance. In addition to equal access to health services, the new insurance system significantly decreases the price of medical care to consumers. Under the current system, the insured contribute a co-pay of only 30 baht (approximately 0.75 USD) per medical visit.

Changes in prices may affect the composition of household expenditures and demand patterns. A decrease in out-of-pocket expenditures on health as a result of the UC reform may also alter households' demand because of the substitution effect and the income effect. Household reactions may differ because their preferences over goods are not similar. Under the new national health insurance program, households are more likely to increase the quantity demanded of medical care services as a result of moral hazard. Additionally, they may spend some of extra wealth on other commodities in their budgets. Therefore, uncompensated and compensated price elasticities may vary across households.

The purpose of this paper is to explore the effects of the UC reform on the allocation of household expenditures on consumption goods (i.e. housing, food, etc.) by comparing expenditure patterns before and after the reform. The study examines the household expenditure patterns, using a Quadratic Almost Ideal Demand System (QUAIDS) developed by Banks, Blundell, and Lewbel (1997) incorporating with a two-step approach introduced by Shonkwiler and Yen (1999).

2. Literature review

Moral hazard, first referenced by Arrow (1963) and Pauly (1968), has been one of the most important issues in health insurance studies. It referred to the increased usage of services when the pooling of risks leads to decrease marginal costs for the services (Folland et al.,1998). Having insurance is a way of dealing with risk because risk can be shifted from insurers to insurance companies.

In terms of demand theory, in addition to the effect of health insurance on medical care consumption, acquiring health insurance is also likely to affect other commodities demand. With a reduction in out-of-pocket health care expenses, households tend to reduce consumption in other goods because of the substitution effect. At the same time, since UC is essentially free and mandatory, a decrease in budget for paying for medical care and insurance results in an income effect. That is an income effect may cause households to increase quantity demanded of other commodities, as a result of a rise in household real income. As a result, households may be expected to have impacts on their health care use and other consumption behaviors because of the UC reform.

Several studies show that a decrease in health care prices and medical care expenditure as a result of health insurance may affect household decisions and other household consumption. Chou et al. (2003), studying the impact of the introduction of National Health Insurance (NHI) in Taiwan on household consumption, found that households tend to decrease their precautionary savings and increase consumption under NHI. Wagstaff and Pradhan (2005) examined the effect of Vietnam Health Insurance, which covers people who work in formal public and private sectors on health outcomes and household consumption. They found that the program increased non-medical household consumption especially on non-food items. Jalan and Rallion (2001), using data from rural China, found that risk associated with medical expenses affect household decisions by decreasing or increasing the amount of wealth that households hold.

In addition to a change in medical care expenditures, factors such as other expenditures, preferences, health perception, and demographic factors can alter household consumption patterns. Busch et al. (2004), estimating the patterns of substitution and complementarity effects between tobacco products and other goods, showed that expenditures on tobacco decrease other household consumption such as housing. In Thailand, Paxson (1993) found seasonal consumption patterns caused by seasonal changes in preferences or prices. In term of perception on health, Taube (1989) indicated that household expenditure patterns and its elasticities change because of improved perceptions of health by the head of household. Demographics and other household factors may also alter consumption patterns. Kalwij et al. (1998) found that demographic factors, household expenditure and female employment alter the allocation of household expenditure to consumer goods in the Netherlands. Handa (1996) also showed that the sex of household head tends to influence members expenditure behavior.

Several studies indicate that household consumptions may vary with income. Park et al. (1996) explored a complete systems demand analysis in U.S. household consumption pattern at various poverty levels. They reported that although income elasticities are higher for the poor households, both poor and non-poor households showed similar own-price elasticities for 12 food groups. Sengul and Tuncer (2005), using the Linearly Approximated Almost Ideal Demand System (LA-AIDS), estimated the food demand for poor households in Turkey. The results showed that extremely poor households' food demand is more sensitive to prices and income than are those of poor households. Raper et al. (2002) estimated demand for nine aggregated food commodity groups by segmenting households to poverty status and non-poverty status groups. They suggested that adding demographic variables in the demand system yields better predictions of household food expenditure behavior.

Numerous studies of household consumption expenditures have been applied to various kinds of model system estimation. Several studies in household expenditures apply the single-equation estimation such as the Linear Expenditure Model (Burney and Akmal, 1991; Raper, 2002) or the Quadratic Expenditure Model (Barnes and Gillingham, 1984; Kohn and Missong, 2003).[1] Due to several limitations of this system, a flexible form demand system such as Almost Ideal Demand System (AIDS) proposed by Deaton and Muellbauer (1980) has been the most widely used framework for estimating the demand system.[2]

Various areas of studies, especially on food demand and other agricultural demand have applied the AIDS models in many countries such as Greece (Mergos and Donatos, 1989); Japan (Hayes et al., 1990); China, (Gao et al., 1994); India, Abdulai, et al. (1999); and Lithuania (Hossain and Jensen, 2000). Moreover, previous work that focus on estimating household consumption demand often use the AIDS model or the extended AIDS models. These studies include Farooq et al. (1999), Michelini (1999), Han and Wahl (1998), and Wang and Chen (1992).

---

[1] Two main limitations of the single-equation approach are 1) it cannot calculate the influence of cross-price elasticities, and 2) it cannot be used to test the symmetry and adding-up hypotheses associated with demand theories.

[2] Although the flexible demand system like AIDS is more flexible, the major limitation is that under the demand system, all income elasticities and all cross price elasticities are restricted to be positive.

Recently, literature on household demand patterns has questioned whether a linear flexible functional form such as the AIDS models can fit data adequately (Blundell et al., 1993; and Banks et al., 1997). Cranfield et al. (2003) compared the predictive ability of five demand models namely; An Implicitly, Directly Additive Demand System (AIDADS), Quadratic Almost Ideal Demand System (QUAIDS), Quadratic Expenditure Model, Linear Expenditure Model, and AIDS models, by analyzing international consumption patterns in 64 countries. Comparing the predictive ability, they found that the first three models are superior to the linear demand systems; the latter two models. Furthermore, Karagiannis and Velentzas (2004) studied a decomposition analysis of Greek consumption patterns by applying a habit persistence version of the QUAIDS model. The results confirmed that the methodology deals well with simultaneous changes in prices, expenditures, and tastes.

The QUAIDS model has been receiving more attention during the past few years. Especially in food demand, Abdulai and Aubert (2004), applying the QUAIDS model, estimated price and expenditure elasticities and explore how socio-economic characteristics affect food demand in Tanzania. Gould and Villarreal (2006) used a Logit model and QUAIDS to study how Chinese households allocate their food expenditures. To evaluate the effect of external factors on household consumptions, Tiezzi (2002) examined the effect of Environmental Defensive Expenditures on Italian household consumption behavior. However, he found no evidence to support the hypothesis that the external factors determine household consumption. Estimating the demand system in Switzerland, Abdular (2002) found that food groups are necessities, while the non-food group is a luxury. In addition, Molina and Gil (2005), investigating consumer demand in

Peru, showed that tobacco, health and miscellaneous goods are found to be necessities, while transport and leisure are luxury goods.

The remainder of the paper is organized as follows. The following section outlines the theoretical model in a nonparametric regression and a demand system. The non-parametric analysis is used in order to determine the functional form of the Engel curves. Section 4 presents a description of the survey data and the variable selection. The empirical results of household expenditures are shown in section 6. Section 7 presents the conclusions.

3. Theoretical Model

Although the AIDS model has been the most popular framework for estimating the demand system, Bank et al. (1997) indicated that many empirical studies show that expenditure share equations of some commodities appear to be non-linear in the logarithm. Thus, the AIDS model requires an additional term in income in order to capture the nature of household data.

Bank et al. (1997) proposed a Quadratic Almost Ideal Demand System (QUAIDS), in which expenditure shares are quadratic in the logarithm of income. The model allow for Engel curves to be non-linear in the log of expenditure. QUAIDS maintains all the relevant properties of the AIDS model because it was derived as a generalization of Price-independent generalized logarithmic (PIGLOG) preferences. Similar to AIDS, the QUAIDS model is based on the consumer demand theory so that advantages include linear restrictions on the estimation and its flexible functional form. Moreover, results from QUAIDS are consistent when aggregated over consumers. Since QUAIDS incorporate additional terms's of income, the model provides more flexibility. This quadratic logarithmic model also allows goods to be luxuries or necessities at different income levels. Moreover, the demand system allow for the inclusion of several demographic characteristics of each household (Molina and Gil, 2005). Therefore, this paper examines expenditure share equations in an attempt to determine a right functional form by employing a non-parametric regression.

3.1 A non-parametric regression: Engel curves

A non-parametric regression specifies an ad hoc relationship between a dependent variable and a set of independent variables when there is little knowledge of its functional relationship. It estimates the regression function, $\mu(y) = E(y|x)$, by computing the location of *y* within a specific bandwidth-scaled of *x* (Abdulai, 2004). Hardle (1990) provides a good review of non-parametric regression techniques.

This paper applies the nonparametric Generalized Cross Validation function (GCV) in a Gaussian regression, which was proposed by Wahba (1990). GCV can also be used to select a smoothing parameter (bandwidth) by the leave-one-out method to minimize the prediction risk. A survey for the bandwidth selection in the density estimation setting can be found in Jones et al. (1996). Mittelhammer et al. (2000) suggested that the cross-validation approach is most widely used to obtain an appropriated bandwidth.

The generalized cross-validation function, to be minimized as a function of the parameters, is given by

$$GCV(\lambda) = \frac{\sum_{i=1}^{n}(y_i - \hat{\eta}_\lambda(x_i))^2}{(n - tr(A(\lambda)))^2} \quad (1)$$

where y_i and x_i are a dependent variable and an explanatory variable, respectively, $\hat{\eta}_\lambda(x_i)$ is defined as $A(\lambda)Y$ indicating fits at x_i computed by leaving out the *i*th data point.

The models were implemented using a GAM procedure of a SAS program (SAS/STAT) to explore the shape of the nonparametric Engel curves. Six household commodity shares are used as independent variables in each regression while logarithm of total expenditure of all commodities is a dependent variable.

Figure 3.1 shows the non-parametric Engel curves estimations of six commodity groups on the log of total expenditure. Because the shapes of Engel curves of clothing/miscellaneous goods, medical care and healthcare, tobacco and alcohol and housing are quadratic, this shows clear evidence of a non-linear relationship on most Engel curves of budget share and logarithm of total expenditure. Although the Engel curves of the food demand and the transportation group do not appear to have distinct non-linear behavior, they do not show linear relationships. Since the Engel curves illustrate quadratic-logarithmic shape, the AIDS model may not sufficiently capture behavior of the analysis.

In addition, a joint Wald test described by Greene (2003) is conducted for testing QUAIDS against AIDS in order to evaluate whether the quadratic term is necessary. The null hypothesis is: $\lambda_1 = \lambda_2 = ... = \lambda_6 = 0$ where λ_i is a vector of parameters of the quadratic term (equation (9)) . From Table 3.8 and 3.9, the Wald test of joint significance for these parameters shows that with χ^2 =15.302 and its associated P-value = 0.000, we reject the null hypothesis at the 5% level of significant in both years. Therefore, this analysis applies QUAIDS since the QUAIDS model is superior to AIDS.

3.2 Demand systems and QUAIDS

The AIDS model in budget share terms is defined as:

$$w_i = \alpha_i + \sum_{i=1}^{n} \gamma_{ij} \ln p_j + \beta_i \ln\left(\frac{m}{a(p)}\right) \tag{2}$$

where w_i is the household budget share of the i^{th} good, *m* is the household expenditure of all goods in the demand system, p_j is the price of good *j*, *p* is a N-vector of prices, and *ln a(p)* is a price index which is defined as:

$$\ln a(p) = \alpha_0 + \sum_{i}^{n} \alpha_i \ln p_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_{ij} \ln p_i \ln p_j \tag{3}$$

The expenditure share of equation (2) can be written in the following form:

$$w_i = A_i(p) + B_i(p) \ln x \tag{4}$$

where *x* = *m/a(p)* and *a(p)* is a price index as suggested in equation (3). Equation (4) shows that the expenditure share equation is in a form of the “price independent generalized linear” demand equations, which is known as a PIGLOG demand system. Muellauer (1975) showed that PIGLOG demands have aggregation characteristics suggesting that the expenditure term *x* is a function only of expenditures, not prices.

Gorman (1981) indicated that the maximum rank of any exactly aggregable demand system is three, where rank refers to the number of linearly independent terms on the right hand-side of an expenditure share equation. Rank two demand systems include Translog, Linear AIDS, PIGLOG systems, and Price-Independent Generalized Linear

(PIGL). Since exactly aggregable demands are linear in functions of expenditure, Banks et al. (1997) showed expenditure share equations of the form:

$$w_i = A_i(p) + B_i(p)\ln x + C(p)g(x) \tag{5}$$

for goods $i = 1,\ldots, n$ where $A_i(p), B_i(p), C(p)$ and g(x) are differentiable function.

Banks et al. (1997) showed that the indirect utility function of the QUAIDS demand system, which is consistent with equation (5) is:

$$\ln V(p,m) = \left\{ \left(\frac{\ln m - \ln a(p)}{b(p)} \right)^{-1} + \lambda(p) \right\}^{-1} \tag{6}$$

where the price index, $\ln a(p)$, is defined by equation (3).

$$b(p) = \prod_i^n p_i^{\beta_i} \text{ is the Cobb-Douglas price aggregator} \tag{7}$$

$$\lambda(p) = \sum_i^n \lambda_i \ln p_i \tag{8}$$

a(p), b(p), and $\lambda(p)$ are differentiable, homogeneous functions of degree zero in prices, and the expenditure share equation (5) requires *g(x)* = $(\ln x)^2$. The expression $\dfrac{\ln m - \ln a(p)}{b(p)}$ is the indirect utility function of a demand system with a linear form of budget shares in logarithm of total expenditure, which is known as a PIGLOG demand system.

Substituting equations (3) (6) (7) and (8) together, the Marshallian demand of the QUAIDS model as introduced in Banks et al. (1997) is obtained. Using Roy's identity, a budget share equation can be derived as:

$$w_i = \alpha_i + \sum_{i=1}^n \gamma_{ij} \ln p_j + \beta_i \ln\left(\frac{m}{a(p)} \right) + \frac{\lambda_i}{b(p)} \left[\ln\left(\frac{m}{a(p)} \right) \right]^2 + \varepsilon_i . \tag{9}$$

From equation (9), the leading terms in QUAIDS are linear in log expenditure which is similar to the AIDS model in equation (2). In an empirical analysis, one can determine whether AIDS or QUAIDS are more appropriate by testing a joint null hypothesis of $H_0 : \lambda_i = 0 \; \forall i$. When $\lambda_i = 0, \forall i$, QUAIDS collapses to Deaton and Muellbauer's AIDS model. In addition, since the adding-up theoretical restriction requires that $\sum_i^n w_i = 0$, this implies that $\sum_i^n \alpha_i = 1$, $\sum_i^n \gamma_{ij} = 0$, $\sum_i^n \beta_i = 0$, and $\sum_i^n \lambda_i = 0, \forall j$. A symmetric property of the demand theory is satisfied if $\gamma_{ij} = \gamma_{ji}, \forall i \neq j$. Moreover, Homogeneity requires the expenditure shares to be homogenous of degree zero in prices $\sum_{j=1}^n \gamma_{ij} = 0, \forall i$.

The expenditure shares in equation (7) show a quadratic form in the logarithm of income. Since (7) was derived from the PIGLOG preference of the AIDS model, it allows significant properties of the AIDS model as well as the flexibility of a non-linear part.

3.3 Two-step procedure

One problem on household surveys is that a significant number of households encounter with zero consumption purchase during the month of survey caused by infrequent purchases. When any of the demand systems consist of households with many zero consumption values, a standard maximum likelihood (ML) estimator may not be appropriate because its estimates are biased. Table 3.1 shows details of zero consumption purchase of household data in 2000 and 2002. For example, in 2002, 49% of total households report zero consumption purchase for Tobacco and Alcohol, 35% for Medical care and Health, 18.7% for clothing and miscellaneous goods, and 8.7% for the transportation expenditure. It is surprising that the number of zero purchase in health/medical care increased from approximately 30% in 2000 to 35% in 2002 although the price of medical visit decreased because of UC. To explain, the majority of the expenditure transactions in health and medical care were from medical supplies and Over-The-Counter (OTC) medicine, not from doctor visits. Although the number of households using outpatient and inpatient care slightly increased by two and four percent, respectively when the price of medical care decreased, the number of households with zero purchase on medical supplies/medicine increased by almost six percent. Therefore, the net number of households with zero purchase in the medical care and health expenditure increased in 2002.

To deal with the censored demand system on QUAIDS, several studies such as Yen et al. (2002), Kang (2003), Shiptsova et al. (2004), and Lambert et al (2006) incorporated a two-step estimation procedure proposed by Shonkwiler and Yen (1999). They showed that the procedure produces consistent estimators for all estimated parameters.[3] A two-step estimation procedure can be shown as follows:

A system of equations with censoring of expenditure *i* of household *h* is governed by a separate stochastic process $z_{ih}^{'}\tau_i + \upsilon_{ih}$ such that

$$w_{ih} = \begin{cases} w_{ih}(p, X \mid \theta) + \varepsilon_{ih} & \text{if } z_{ih}^{'}\tau_i + \upsilon_{ih} > 0 \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

$(i = 1, \ldots\ldots, n;\; h = 1, \ldots\ldots, H)$

[3] Heien and Wessells (1990) proposed a methodology of estimating Probit models in the first stage and an AIDS model in the next stage using an inverse Mills ratio, however Shonkwiler and Yen (1999) showed that this methodology performed poorly in Monte Carlo simulations.

where w_{ih} denotes the observed expenditure share, θ represents all parameters in a certain demand system, z_{it} is a vector of exogenous variables, τ_i is a conformable parameter vector, and ε_{ih} and υ_{ih} are random errors.

A system of demand equation is expressed as:

$$w_{ih} = E(w_{ih}) + \xi_{ih} = \Phi(z_{ih}^{'}\hat{\tau}_i)w_{ih}(p, X; \theta) + \delta_i \phi(z_{ih}^{'}\hat{\tau}_i) + \xi_{ih} \qquad i = 1, 2, \ldots, n \qquad (11)$$

where $\xi_{ih} = w_{ih} - E(w_{ih})$, with $E(\xi_{ih}) = 0$, and ξ_{ih} is heteroscedastic with variance (Shonkwiler and Yen, 1999), Φ is the standard normal CDF for each expenditure equation *i*, ϕ is the standard normal PDF for each equation *i*, $z_{ih}^{'}$ is a vector of explanatory variables for household *h* from Probit model estimations in (11), and $\hat{\tau}_i$ is a vector of estimated parameters from Probit model estimations in (11). The two-step estimation procedure yields consistent estimators for all parameters (Yen et al., 2002; Lambert et al., 2006).[4]

The demand system can be estimated with a two-step procedure. First, we obtain maximum-likelihood (ML) Probit estimates of $\hat{\tau}_i$ for each of the *n* equations by using binary outcomes $w_{ih} = 0$ and $w_{ih} > 0$. The exogenous variables used in these Probit estimations were household characteristics that may influence purchasing decisions, such as household size and income; dummy variables for geographic location; number of young children, and number of retired members (> 65 years old). Second, by using the cumulative distribution functions (CDFs) and standard normal probability density

[4] Although the estimator for this procedure is not efficient, Shonkwiler and Yen (1999), using simulations, showed that this methodology performs well.

functions (PDFs) derived from Probit estimations, $\Phi(z_{ih}^{'}\hat{\tau}_i)$ and $\phi(z_{ih}^{'}\hat{\tau}_i)$ for all i are calculated. Then, $\theta, \delta_1, \delta_2, ..., \delta_n$ can be estimated in the augmented system.

Therefore, the estimated equations for the QUAIDS model for each household is shown as follows:

$$w_i = \Phi_i(z_{ih}^{'}\hat{\tau}_i)\left\{\alpha_{i0} + \sum_{i=1}^{n}\gamma_{ij}\ln p_j + \beta_i \ln\left(\frac{m}{a(p)}\right) + \frac{\lambda_i}{b(p)}\left[\ln\left(\frac{m}{a(p)}\right)\right]^2\right\} + \delta_i\phi_i(z_{ih}^{'}\hat{\tau}_i) + \varepsilon_i \quad (12)$$

where w_i is budget share of expenditure type *i* for *i* = 1, …, 7.

Seemingly Unrelated Regressions (SUR) is used to estimate the demand system. Yen et al. (2002) noted that the second-step estimation of the system should be based on the full n-vector since the right-hand side of the system does not add up to one implying that the adding-up condition is not satisfied.

The QUAIDS model expenditure elasticities, uncompensated price elasticities can be formulated by differentiating equation (12) with respect to *ln m* and $\ln p_j$, respectively, which give:

$$\mu_i \equiv \frac{\partial w_i}{\partial \ln m} = \Phi(z_{ih}^{'}\hat{\tau}_i)\cdot\left(\beta_i + \frac{2\lambda_i}{b(p)}\left\{\ln\left(\frac{m}{a(p)}\right)\right\}\right) \quad (13)$$

$$\mu_{ij} \equiv \frac{\partial w_i}{\partial \ln p_j} = \Phi(z_{ih}^{'}\hat{\tau}_i)\cdot\left(\gamma_{ij} - \mu_i\left(\alpha_j + \sum_{k=1}^{n}\gamma_{jk}\ln p_k\right) - \frac{\lambda_i\beta_j}{b(p)}\left\{\ln\left(\frac{m}{a(p)}\right)\right\}^2\right) \quad (14)$$

From this procedure, the expenditure (budget) elasticities (e_i) is shown as $\frac{\mu_i}{w_i} + 1$ and the uncompensated price elasticities are derived as $e_{ij}^{u} = \frac{\mu_{ij}}{w_i} - \varphi_{ij}$ where φ_{ij} is the

Kronecker delta, with $\varphi_{ij} = 1$ for $i = j$, and 0 otherwise. The description of this non-linear iterative approach can be found in Bank et al. (1997).

4. Estimation methodology

The demand system of nonlinear regression equations can be written in vertically stacked form as:

$$w = g(x,\theta) + \varepsilon \tag{15}$$

where $w = \begin{pmatrix} w_1 \\ \vdots \\ w_6 \end{pmatrix}$, $g(x,\theta) = \begin{pmatrix} g_1(x,\theta_1) \\ \vdots \\ g_6(x,\theta_6) \end{pmatrix}$, $\varepsilon = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_6 \end{pmatrix}$,

x = a vector of endogenous variables for the all six demand equations

= $\begin{pmatrix} 1 & \ln(m) & \ln(p) \end{pmatrix}$.

Since, this model apply a Shonkwiler and Yen two-step approach according to equation (12), θ represents the model coefficient parameters for i^{th} demand, which are

$$\begin{pmatrix} \alpha_1 & \gamma_{11} & \cdots & \gamma_{16} & \beta_1 & \lambda_1 & \delta_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ \alpha_6 & \gamma_{61} & \cdots & \gamma_{66} & \beta_6 & \lambda_6 & \delta_6 \end{pmatrix}_{6\times 10} \tag{16}$$

This study applies SUR by utilizing a non-linear Generalized Least Squares (GLS) procedure to estimate sets of regression equations. The Gauss-Newton optimization algorithm is applied to solve nonlinear least squares problems. For an initial value, α_0, Deaton and Muellbauer (1980) suggested to use the lowest value of *ln(m)* in the data. The increment to the starting values is the least squares regression of the residuals y - $h(x,\theta_0)$ on the Jacobian matrix of the nonlinear regression function, $h(x,\theta_0)$.

For a single nonlinear equation, the unrestricted estimator of the model parameters is given by

$$\theta_{GLS} = \theta_0 + \left[\frac{\partial h}{\partial \theta}' \Psi^{-1} \frac{\partial h}{\partial \theta}\right]^{-1} \left[\frac{\partial h}{\partial \theta}' \Psi^{-1}\left(y - h(x,\theta_0)\right)\right] \quad (17)$$

where the covariance matrix structure is $E(\varepsilon\varepsilon') = \sigma^2 \Psi = \Sigma \otimes I_n$.

Σ is a covariance matrix generated by the product of error vectors from all equations.

$$\Sigma = ee' = \begin{bmatrix} e_1 \\ \vdots \\ e_7 \end{bmatrix} \begin{bmatrix} e_1 & \cdots & e_7 \end{bmatrix} = \begin{bmatrix} e_1 e_1 & \cdots & e_1 e_7 \\ \vdots & \ddots & \vdots \\ e_7 e_1 & \cdots & e_7 e_7 \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{17} \\ \vdots & \ddots & \vdots \\ \sigma_{71} & \cdots & \sigma_{77} \end{bmatrix} \quad (18)$$

Since the system is non-linear, the estimation of the nonlinear system takes the form of (19)

$$\begin{bmatrix} W_1 \\ \vdots \\ W_7 \end{bmatrix} = \begin{bmatrix} h_1(x,\theta_0) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & h_7(x,\theta_0) \end{bmatrix} + \begin{bmatrix} u_1 \\ \vdots \\ u_7 \end{bmatrix} \quad (19)$$

where $h_i(x,\theta_0)$ denotes the QUAIDS functional form, requires a modification of (17).

With unknown noise covariance (Ψ), Mittelhammer, Judge, and Miller (2000) suggest to replace Ψ^{-1} with $(\Sigma \otimes I)^{-1}$ when estimates the system of nonlinear equations. Alternatively, Eq. (17) can be expressed as:

$$\theta_{GLS} = \theta_0 + \left[\frac{\partial h}{\partial \theta}' (\Sigma \otimes I)^{-1} \frac{\partial h}{\partial \theta}\right]^{-1} \left[\frac{\partial h}{\partial \theta}' (\Sigma \otimes I)^{-1}\left(y - h(x,\theta_0)\right)\right] \quad (20)$$

In the case of Feasible GLS (FGLS), Σ is replaced by $\hat{\Sigma}$:

$$\theta_{GLS} = \theta_0 + \left[\frac{\partial h}{\partial \theta}' \left(\hat{\Sigma} \otimes I \right)^{-1} \frac{\partial h}{\partial \theta} \right]^{-1} \left[\frac{\partial h}{\partial \theta}' \left(\hat{\Sigma} \otimes I \right)^{-1} \left(y - h(x, \theta_0) \right) \right] \tag{21}$$

where $$\frac{\partial h}{\partial \theta} = \begin{bmatrix} \frac{\partial h_1}{\partial \alpha_1} & \frac{\partial h_1}{\partial \gamma_{1j}} & \frac{\partial h_1}{\partial \beta_1} & \frac{\partial h_1}{\partial \lambda_1} & \cdots & \frac{\partial h_1}{\partial \alpha_7} & \frac{\partial h_1}{\partial \gamma_{7j}} & \frac{\partial h_1}{\partial \beta_7} & \frac{\partial h_1}{\partial \lambda_7} \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial h_7}{\partial \alpha_1} & \frac{\partial h_7}{\partial \gamma_{1j}} & \frac{\partial h_7}{\partial \beta_1} & \frac{\partial h_7}{\partial \lambda_1} & \cdots & \frac{\partial h_7}{\partial \alpha_7} & \frac{\partial h_7}{\partial \gamma_{7j}} & \frac{\partial h_7}{\partial \beta_7} & \frac{\partial h_7}{\partial \lambda_7} \end{bmatrix}$$

Alternatively, (10) can be derived to

$$\theta_{GLS} = \theta_0 + \left[\frac{\partial h}{\partial \theta}' \left(\hat{\Sigma}^{-1} \otimes I \right)^{-1} \frac{\partial h}{\partial \theta} \right]^{-1} \left[\frac{\partial h}{\partial \theta}' \left(\hat{\Sigma}^{-1} \otimes I \right) \left(y - h(x, \theta_0) \right) \right] \tag{22}$$

where $$\left[\frac{\partial h}{\partial \theta}' \left(\hat{\Sigma}^{-1} \otimes I \right) \frac{\partial h}{\partial \theta} \right]^{-1} = \begin{bmatrix} \sigma^{11} \frac{\partial h_1}{\partial \theta}' \frac{\partial h_1}{\partial \theta} & \cdots & \sigma^{17} \frac{\partial h_1}{\partial \theta}' \frac{\partial h_7}{\partial \theta} \\ \vdots & \ddots & \vdots \\ \sigma^{71} \frac{\partial h_7}{\partial \theta}' \frac{\partial h_1}{\partial \theta} & \cdots & \sigma^{77} \frac{\partial h_7}{\partial \theta}' \frac{\partial h_7}{\partial \theta} \end{bmatrix}^{-1} \tag{22}$$

and $$\left[\frac{\partial h}{\partial \theta}' \left(\hat{\Sigma}^{-1} \otimes I \right) e \right] = \begin{bmatrix} \sigma^{11} \frac{\partial h_1}{\partial \theta}' e_1 & \cdots & \sigma^{17} \frac{\partial h_1}{\partial \theta}' e_7 \\ \vdots & \ddots & \vdots \\ \sigma^{71} \frac{\partial h_7}{\partial \theta}' e_1 & \cdots & \sigma^{77} \frac{\partial h_7}{\partial \theta}' e_7 \end{bmatrix} \tag{23}$$

5. Data and Variable selection

5.1 Data

The main data set used in this study is the 2000 and 2002 Socio Economic Survey (SES) from the National Statistics Office, Thailand. The 2002 survey includes 118,763 members or 34,820 households from 76 provinces in Thailand. The survey collects data of household expenditures, household income, and household members characteristics such as sex, age, education, and occupation. After discarding observations with negative budget shares, missing data, and outliers, 34,607 households are left. The 2000 SES survey contains 24,747 households with similar variables. After dropping some observations that lack crucial variables, 24,586 households are remained. Consumption expenditures are categorized in six broad commodity groups; food, housing, clothing and miscellaneous goods, health and medical care, transportation, and tobacco and alcohol.

The SES survey provides neither price nor quantity information on expenditure items. Raper et al. (2002) suggested to match households with monthly regional price index for each goods category. Since Heien and Durham (1991) argued that using area price indices may capture regional effects and may be linearly dependent with the regional dummy variables, this paper utilizes the household data from the each 76 provinces in which were interviewed at a different month over the year. Thus, prices of households in a given province would not be the same. The author matches the share of expenditure categories of each household with its provincial consumer price index (CPI) using the information of their provincial affiliation and interview month as a proxy of price. Since the price indices may capture regional or provincial effects, regional/provincial dummy variables are not used in the QUAIDS model. Annual consumer price indices of each consumption category are collected from Bureau of Trade and Economic Indices (BTEI), the Ministry of Commerce, Thailand. Since there are six commodity groups, 76 provinces and 12 interviewed months, the analysis involves 5,472 prices in 2002.

One problem is that, in 2000, BTEI provides the CPI data from only 19 (mostly large) provinces from all regions in Thailand. In this study, monthly prices of any 19 provinces are matched to their surrounded provinces. Since Thailand is a small country, consumers living in provinces nearby may experience similar consumer prices.

5.2 Variable selection and descriptive statistics

All variables are explained as follows. For dependent variables the budget shares of six aggregated consumption commodities consist of; 1) food, 2) housing, 3) clothing and miscellaneous goods, 4) health and medical care, 5) transportation and 6) tobacco and alcohol. The definition of all types of expenditures is reported in Appendix 1. Explanatory variables consist of the consumer price indices of each commodity group in logarithmic terms and the monthly household total expenditure.

Demographic variables used in the Probit estimation include the household size (HSIZE), number of dependents in a household (number of small children age 0-5 years old, the number of children age 6-15 years old (CHILD), and number of young children, and number of retired members (>65 years old)), education years of the household head, dummy variables indicating regions (Bangkok, Central, North, Northeast, and South), gender of the head of household (Female; 1 if a household head is female). This paper also include household income which was divided within three group into: (1) the bottom 40 percent (BOTTOM), (2) the middle 40 percent (MIDDLE) and (3) the top 20 percent (TOP) based on per capita expenditures.

Table 3.2 and Table 3.3 show the variable description and descriptive statistics of variables used in the first-step Probit estimation and the second-step QUAIDS analysis in the year 2000 and 2002. From Table 3.2, the average total expenditure of the demand system in 2000 and 2002 were approximately 9,500 baht per month. While in 2000 housing commodities accounted the most (42%), the majority of the budget share in 2002 was food (46%). The average amount of spending in food increased from 2,626 in 2000 to 3,743 baht in 2002. Although the health and medical care expenditure of both years were rather similar which were approximately 260-270 baht. Interestingly, the transportation expenditure significantly increased over two years by approximately 60%. On both years, the average expenditure shares of health and medical care and tobacco/alcohol were the lowest, ranging from two to three percent of all expenditures.

In terms of household characteristics, on both years, the average size of households in the sample was three to four members mostly from the central, northern, and north eastern regions (Table 3.3). On average, head of households completed six to seven years of education, which attained the elementary level or higher. From the sample, the average household income increased from approximately 13,630 baht in 2000 to 14,250 baht in 2002. In addition, about 30% of household heads are female. In addition, the pairwise comparisons of equal means are conducted on the null hypothesis that there is no mean difference on household heads' level of education, the size of household and household income across the two years (2000 and 2002). The two-sample t test indicated that we cannot reject the hypothesis that the population mean of the education level cross the two years are zero implying that the average of household heads education level are similar. However, the results showed that the population mean of the size of household and household income are different.

Table 3.4 reports means of household expenditures by household characteristics in 2002. On average, households in Bangkok spent the most on all types of commodities relative to other regions with an average expenditure of approximately 18,160 baht per month, followed by households in the central region (10,468 baht) and the south (10,077 baht), respectively. The average amounts of expenditures in Bangkok are considerably higher than those in other regions. In particular, households in Bangkok spend twice or three times more than other regions on medical care and health, housing, and transportation.

Households that have female heads generally spent less than households with male household heads. The average spending on housing and food of matriarchal households was approximately 13% lower than patriarchal households'. Although the difference in other expenditures was not significant, the amount of spending in tobacco and alcohol by households with female household heads was 100% lower than households with male head of households.

In addition, the amounts of expenditure rise with the level of education attained by heads of household and household income. The average of total expenditure by households, in which a head of household had no education, was approximately 6,800 baht per month; in which households whose a head acquired a bachelor degree or higher spent roughly 20,600 baht. As expected, the average total expenditure was the highest on high-income households (16,840 baht), and was the lowest on low-income households (5,757 baht).

The expenditure patterns by household characteristics in 2000 were moderately similar to the 2002's (Table 3.5). On average, households in Bangkok spent the highest amount of money (16,708 baht/month) while households from the northern and the north eastern regions spend the least (approximately 8,060 and 7,990 baht, respectively). Households with higher education level and higher income contributed to higher amounts on all types of expenditures.

6. Empirical results

In the first-step Probit regression, demographic variables consist of several variables indicated in section 5.1. STATA 9 is used to estimate the Probit models of each demand equation that consist of serious zero consumption problems. For the second step, Seemingly Unrelated Regression (SUR) is estimated by utilizing a non-linear GLS procedure to estimate sets of regression equations. The programming was done on GAUSS 7.0 by Aptech Systems in order to solve the nonlinear least squares problems.

6.1 The first-step results

Table 3.6 and 3.7 show the Probit estimation of the Thai demand system in 2000 and 2002, respectively. The Probit estimations of each demand group were estimated for the demand groups without the serious zero consumption problems. Thus, the study does not include the food demand and the housing demand in this step.

The results indicated that most estimated coefficients of both years are significant at the 5% level of significance or higher. The effects of income were positively correlated to all consumption groups in 2000. The top earners (TOP) households tended to consume commodities the most, while the poorest households (BOTTOM) appeared to consume the least. It is interesting that in 2002, the richest households became the least likely group to consume tobacco and alcohol, while the poorest group consumed them the most in relation to their budgets. The occurence may come from two possible reasons. A new cigarette labeling law has been in effect since 2001. The law requires message and picture-based warnings on every cigarette package. The health messages must comprise, on average, at least 50% of the front and back of the package. A number of middle to upper class consumers may become more aware of the health effects of smoking. However, the introduction of UC may induce moral hazard when consumers understand that most medical care is accessible at almost no cost.

For geographic location, Central households tended to consume the most healthcare compared to other regions while households from Bangkok spent the least. After UC, Bangkok still appeared to spend the least on healthcare while households from the Northern region were likely to consume it the most. In both years, Northeastern households tended to consume transportation the least while southern households were more likely to spend the most. For tobacco and alcohol, households from the southern region appeared to consume them the least relative to other regions in 2000, however, Bangkok replaced the last rank in 2002. Also, households from the southern region tended to consume clothing and miscellaneous goods the most in 2000 in relation to their budgets, while households from the north appeared to consume clothing the most in 2002.

In addition, the results show the number of children (up to 15 years) was negatively related to household consumptions on tobacco/alcohol and transportation, but showed a positive relationship on clothing and miscellaneous goods. Although households with higher number of small children spent more in health and medical care, households having older children (6-15 years) appeared to consumes less healthcare. In contrast, households having more retirees were more likely to consume less on every group of commodities except healthcare which appeared to become more important for the elderly. This demand pattern stays unchanged in 2002.

In terms of household head's years of education attained, households headed with high-educated members were more likely to consume clothing and miscellaneous goods and transportation while they tended to spend less in health& medical care and tobacco/alcohol. The size of household also played a positive role in determining all types of consumption as expected in that larger households tended to consume more on every commodity compared to the smaller ones. In addition, the results also indicated that households with female household heads were more likely to consume clothing and miscellaneous goods.

6.2 The second-step results

The second step estimates equation (12) with the QUAIDS model by using Seemingly Unrelated Regressions (SUR) in six consumption groups. The QUAIDS estimation results are reported in Table 3.8 and 3.9.

Most parameter estimates for the six consumption groups were statistically different from zero at the 5% level. The coefficients for the censoring correction, δ, were statistically significant in both years in every group except food and housing which having no Probit estimates. The results justified the problem of zeros in the observed budget shares. Comparing various parameter estimation results between the year 2000 and 2002, the magnitude of most demand groups in the year 2000 were less than the ones in 2002. Moreover, the results indicated that the standard errors reported for the results from year 2000 are smaller than those for the year 2002.

Table 3.10 presents the uncompensated price elasticity and expenditure elasticity estimates of the Thai household data from the two-step QUAIDS procedure. The formulae used to calculate the elasticities are from Banks, Blundell and Lewbel (1997). In both years, all expenditure elasticity estimates were positively significant at the 5% level as expected. In 2000, the expenditure elasticity on clothing and miscellaneous goods was positive with the largest magnitude (1.35) while the food expenditure elasticity was the lowest and less than one in both years. In 2002, the expenditure elasticities decreased in most groups except food and transportation. The expenditure elasticity for transportation was the highest in 2002 (1.45), followed by clothing and miscellaneous goods (1.22), health and medical care (1.07), and tobacco and alcohol (1.05), respectively.

All uncompensated price elasticities showed negative signs and were statistically significant at the 10% level which were consistent with the demand theory (Table 3.10).[5] Most elasticity estimates for the year 2000 own-price elasticities were lower in absolute value than the ones in 2002 which indicated that the demand for most consumption groups were more sensitive to changes in own price in 2002. Only the 2000 point estimates for the own price elasticities of clothing/miscellaneous goods and transportation were less elastic compared to those in 2002.

Additionaly, the food own-price elasticity was smaller in absolute value on both years confirming that foods are necessities. It appeared that food consumers in 2000 were less sensitive to price increases than those in 2002. The elasticity estimates for health and medical care (-1.33) and housing (-1.3) were above one in absolute value in 2002 indicating that the demand for these commodities were more elastic. Interestingly,

[5] The estimated parameter of the health and medical care in 2000 were statistically significant at the 11% level.

the elasticity estimate for tobacco/alcohol was almost unitary elastic in 2000 (-0.968), but appeared to be very elastic in 2002 (-4.587).

In terms of cross-price elasticities, approximately half of the estimated parameters were statistically significant at the 10% level or higher. The results showed that food appeared to be a complement for health and medical care in both years. For transportation goods, showing positive signs, remain a substitute for health and medical care goods in both years that is the demand for these goods responded negatively to a decrease in the price of medical care. Although in 2000 housing appeared to be a complement, in 2002 a decrease in the price of healthcare appeared to decrease the demand for housing since it became a substitute. The cross-price elasticities of all groups of commodities except housing's showed the same sign, but , the magnitude of most demand groups in the year 2002 was higher than the one in 2000 indicating that the demand for most goods was more sensitive to a decrease in medical care cost as a results of the UC reform.

7. Conclusions

This study evaluates the effects of the Thai Universal Health Coverage (UC) program, introduced in 2001, in the allocation of household expenditures on their consumption goods. The Quadratic Almost Ideal Demand System (QUAIDS) model is applied in this analysis because 1) the Wald test rejects the hypothesis of the quadratic term quadratic term parameters are zero and 2) the Engel curves illustrate quadratic-logarithmic shape. Since the existence of zero values in observed consumption is always observed in survey data, demand estimation procedures that ignore the censoring problem in the dependent variables produce biased and inconsistent parameter estimates. This study incorporates a two-step approach introduced by Shonkwiler and Yen (1999) to estimate the non-linear demand systems in 2000 and 2002. The programming was done on GAUSS 7.0 in order to solve the nonlinear least squares problems applying the Gauss-Newton optimization algorithm.

The first-step Probit estimations showed that household income was positively correlated to all consumption groups in 2000. However, the richest households became the least likely group to consume tobacco and alcohol, while the poorest group consumed them the most in relation to their budgets in 2002. In addition, Central households tended to spend the most in healthcare compared to other regions while households from Bangkok spent the least. After UC, households from the Northern region were more likely to consume it the most. The results also indicated that households with children tended to have less expenditure in tobacco/alcohol and transportation, but spent more on clothing and miscellaneous goods. Education also played an important role in determining the demand pattern in that households headed with high education were more likely to consume clothing and miscellaneous goods and transportation while they tended to spend less in health& medical care and tobacco/alcohol.

On the second step, the results of uncompensated own price elasticities indicated that the demand for food, clothing and miscellanous goods, health and medical care, and housing became more elastic in 2002 after UC was introduced. In addition, households tended to spend more on food and tobacco/alcohol, but spent less in clothing/miscellaneous goods, housing and transportation as a result of a reduction in the price of medical care. The cross-price elasticities of all groups of commodities except housing's showed the same sign, but having larger the magnitude in most demand groups. Although food appeared to be a complement in both years, the demand for food became more sensitive to a drop in medical care price. In addition, UC decreased the demand for transportation because of the income effect of having higher real income and the substitution effect with respect to lower health care costs.

REFERENCES

Abdulai, A. and Aubert, D. (2004), "A Cross-Section Analysis of Household Demand for Food and Nutrients in Tanzania", *Agricultural Economics* 31 (1), 67-79

Abdulai, A. (2002), "Household Demand for Food in Switzerland: A Quadratic Almost Ideal Demand System", *Journal of Economics and Statistics* 138 (1), 1-18

Abdular, A., Jain, D. K., Sharma, A. K. (1999), "Household Food Demand Analysis in India", *Journal of Agricultural Economics* 50 (2), 316-27.

Altman, D.G. (1991), *Practical statistics for medical research*, Chapman and Hall, London.

Arrow, K. J. (1963), "Uncertainty and the Welfare Economics of Medical Care", *American Economic Review* 53, 941-973.

Bai, J. (2005), "Estimating a Quadratic AIDS Model Using Chinese Urban Household Data and Implications for Dairy Products", *China's Agri-food value chain: opportunities and challenges*, WERA-101 proceedings, Reno.

Banks, J. , Blundell, R., and Lewbel, A. (1997), "Quadratic Engel curves and consumer demand", *Review of Economics and Statistics*, LXXIX (4), 527-539.

Barnes, R. and Gillingham, R. (1984), "Demographic effects in demand analysis: Estimation of the Quadratic Expenditure System Using Microdata", *The Review of Economics and Statistics* 66 (4), 591-601.

Blundell, R., Pashardes, P., and Weber, G. (1993), "What do We Learn About Consumer Demand Patterns from Micro Data?", *American Economic Reviews* 83(3), 570-597.

Burney, N. A., and Akmal, M. (1991), “Food Demand in Pakistan: An Application of the Extended Linear Expenditure System”, *Journal of Agricultural Economics* 42 (2), 185-195.

Busch, S. H., Jofre-Bonet, M., Falba, T. A, and Sindelar, J. L. (2004), “Burning a Hole in the Budget: Tobacco Spending and Its Crowd-Out of Other Goods”, *Applied Health Economics and Health Policy* 3 (4), pp. 263-72.

Chou, S. Y., Liu, J. T., and Hammitt, J. K. (2003). "National Health Insurance and Precautionary Saving: Evidence from Taiwan." *Journal of Public Economics* 87 (9-10), 1873-94.

Cranfield, J. A. L., Eales, J. S., Hertel, T. W., and and Preckel, P.V. (2003), “Model selection when estimating and predicting consumer demands using international, cross section data”, *Empirical Economics* 28, 353-64.

Deaton A, Muellbauer J. (1980), “An Almost Ideal Demand System”, *American Economic Review* 70, 312-26.

Folland. S., Goodman, A. C., and Stano, M. (1997), The Economics of Health and Health Care, 2nd ed. Prentice Hall, NJ.

Farooq, U., Young, T., Iqbal, M. (1999), “An Investigation into the Farm Households Consumption Patterns in Punjab, Pakistan”, Pakistan Development Review 38 (3), 293-305

Gao, X. M., Wailes, E. J., and Cramer, G. L. (1994), “A Synthetic Demand System: An Application to U.S. Consumer Demand for Rice and Selected Rice Substitutes”, *Review of Agricultural Economics* 16 (1), 27-38

Gould, B. W and Villarreal, H. J. (2006), “An Assessment of the Current Structure of Food Demand in Urban China”, *Agricultural Economics* 34 (1), 1-16

Greene, W. H. (2003), *Econometric Analysis*, 5th edi., Prentice Hall, NY.

Han, T., Wahl, T. (1998), "China's Rural Household Demand for Fruit and Vegetables", *Journal of Agricultural and Applied Economics* 30 (1), 141-50

Handa, S. (1996), "Expenditure behavior and children' s welfare: An analysis of female headed households in Jamaica", *Journal of Development Economics* 50, 165-187

Hardle, W. (1990), *Applied Nonparametric Regression*, Cambridge University Press, NY.

Hayes, D. J., Wahl, T. I, Williams, G. W. (1990), "Testing Restrictions on a Model of Japanese Meat Demand**",** *American Journal of Agricultural Economics* 72 (3), 556-66

Hossain, F. and Jensen, H. (2000), "Lithuania's Food Demand during Economic Transition", *Agricultural Economics* 23, 31-40

Jalan, J. and Rallion, M. (2001), "Behavioral responses to risk in rural China", *Journal of Development Economics* 66, 23-49.

Jones, M. C., Matron, J. S. and SheatheL S. J. (1996), "A brief survey of bandwidth selection for density estimation", *Journal of the American Statistical Association*, 91, 401-07.

Kalwij et al. (1998), "Household Commodity Demand and Demographics in the Netherlands: A microeconometric analysis", *Journal of Population Economics,* 11, 551-577

Kang, E. L.(2003), Food demand in urban China: an empirical analysis using micro household data, PhD dissertation. Ohio State University

Karagiannis, G. and Velentzas, K. (2004), "Decomposition Analysis of Consumers' demand changes: an application to Greek consumption data", *Applied Economics*

36,

Kohn, K. and Missong, M. (2003), “Estimation of Quadratic Expenditure Systems Using German Household Budget Data”, *Jahrbücher* 223 (4), 422

Lambert, R., Larue, B., Yelou, C. , and Criner, G. (2006), “Fish and Meat Demand in Canada: Regional Differences and Weak Separability”, *Agribusiness*, 22 (2), 175–199.

Lewbel, A., “A Unified Approach to Incorporating Demographic or Other Effects into Demand Systems,” Review of Economic Studies, Vol. 52, 1985: 1-18.

Lewbel, A. (1991), “The Rank of Demand Systems: Theory and Nonparametric Estimation”, *Econometrica* 59, 711–30.

Merz, J. (1983), “The Functionalized Extended Linear Expenditure System: Theory, Estimation Procedures and Application to Individual Household Consumption Expenditures Involving Socioeconomic and Sociodemographic Characteristics”, *European Economic Review* 23 (3), 359

Mergos, G. J. and Donatos, G. S., (1989), “Demand for Food in Greece: An Almost Ideal Demand System Analysis”, *Journal of Agricultural Economics* 40 (2), 178-84

Michelini, C. (1999), “New Zealand Household Consumption Patterns 1983-1992: An Application of the Almost-Ideal-Demand-System”, New Zealand Economic Papers 33 (2), 15-26.

Mittelhammer, R. C., Judge, G. G., and Miller, D. J. (2000), *Econometric Foundation*, Cambridge University Press, NY.

Molina, J. A. and Gil, A. I. (2005), “The Demand Behaviour of Consumers in Peru: A

Demographic Analysis Using the QUAIDS", *Journal of Developing Areas* 39 (1), 191-206

Pauly, M.V. (1968), "The Economics of Moral Hazard. Comment", *American Economic Review* 58, 531-537.

Park, J. L., Holcomb, R. B., Raper, K. C. and Capps, O. Jr (1996), "A demand systems analysis of food commodities by US households segmented by income", *American Journal of Agricultural Economics* 78, 290-300.

Paxson, C. H. (1993), "Consumption and Income Seasonality in Thailand", *Journal of Political Economy* 101 (1), 39-72.

Poi, B. (2002), "From the help desk: Demand system estimation", Stata Journal 2 (4), 403-410.

Raper, K. C., Wanzala, M. N., and Nayga, R. M., Jr (2002), "Food expenditures and Households demographic composition in the U.S.: a demand systems approach", *The review of Economics and Statistics LXXIII* (2), 189-199.

Sengul and Tuncer (2005), "Poverty Levels and Food Demand of the Poor in Turkey", *Agribusiness,* 21 (3), 289–311

Shiptsova, R., Goodwin, H.L., and Holcomb, R. B. (2004), "Household expenditure patterns for carbohydrate sources in Russia", *Journal of Agricultural and Resource Economics* 29(2), 296-307

Shonkwiler, J. Scott, and Yen, Steven T., (1999), "Two-step estimation of a censored system of equations", American Journal of Agricultural Economics 81 (4), 972–982.

Taube, P. M. (1989), "An Analysis of the Influence of Self-assessed Health on

Expenditure in a Complete Demand System", *Journal of Economic Psychology* 10 (1), 101-16.

Tiezzi, S. (2002), "Environmental Defensive Expenditures and Households Behavior in Italy", *Applied Economics* 34 (16), 2053-61.

Wagstaff, A. and Pradhan, M (2005), "Insurance Health impacts on Health and Non-medical Consumption in a Developing Country", *World Bank Policy Research Working Paper* 3563.

Wang, Z. and Chen, W. S. (1992), "Effects of Rationing on the Consumption Behavior of Chinese Urban Households during 1981-1987", *Journal of comparative economics* 16 (1), 1 –26

Wahba, G. (1990), *Spline Models for Observational Data,* Philadelphia: Society for Industrial and Applied Mathematics.

Yen, S.T., Kan, K, and Su (2002), "Household Demand for Fats and Oils: Two-step Estimation of a Censored Demand System", *Applied Economics* 14, 1799-1806.

TABLES

Table 3.1: Number of households with zero purchase categorized by type of expenditure

Variable	Number of observations with zero purchase	Number of observed households	Percentage
Year 2000			
Food expenditure	22	24,586	0.10%
Clothing/Miscellaneous	4,511	24,586	18.30%
Housing expenditure	0	24,586	0.00%
Medical care and health expenditure	7,477	24,586	30.40%
Transportation expenditure	3,276	24,586	13.30%
Tobacco and alcohol expenditure	16,921	24,586	68.80%
Year 2002			
Food expenditure	0	34,607	0%
Clothing and miscellaneous	6,484	34,607	18.70%
Housing expenditure	0	34,607	0%
Medical care and health expenditure	12,162	34,607	35.10%
Transportation expenditure	3,005	34,607	8.70%
Tobacco and alcohol expenditure	16,957	34,607	49%

Table 3.2: Descriptive statistics of variables used in QUAIDS in 2000 and 2002

Variables	Description	2000 (N= 24,586)		2002 (N=34,607)	
		Mean	Std. Dev.	Mean	Std. Dev.
FOOD	Food expenditure	2,625.72	1,521.81	3,742.80	2,303.41
CLOTHMISC	Clothing and Miscellaneous goods expenditure	869.87	1,838.69	789.74	1,509.10
HOUSE	Housing expenditure	3,888.17	3,943.89	2,319.03	2,355.08
HEALTH	Medical care and health expenditure	267.24	884.33	263.24	1,166.91
TRANSPORT	Transportation expenditure	1,317.61	3,678.66	2,121.06	5,728.35
TOBACCO	Tobacco and alcohol expenditure	215.89	587.53	267.51	584.93
EXP	Total expenditure	9,499.82	8,282.43	9,503.37	9,366.92
SFOOD	Share of food expenditure	0.33	0.14	0.46	0.14
SCLOTHMISC	Share of clothing and miscelеneous goods expenditure	0.08	0.11	0.04	0.06
SHOUSE	Share of housing expenditure	0.42	0.16	0.27	0.11
SHEALTH	Share of medical care and health expenditure	0.03	0.05	0.02	0.05
STRANSPORT	Share of transportation expenditure	0.1	0.11	0.15	0.13
STOBACCO	Share of tobacco and alcohol expenditure	0.02	0.05	0.03	0.05

Note: 1 baht is approximately 0.025 USD

Table 3.3: Descriptive statistics of variables used in the first-step Probit model in 2000 and 2002

Variable	Description	Mean	Std. Dev.	Mean	Std. Dev.
		2000 (N = 24,586)		2002 (N = 34,607)	
EDU	Number of head of household's years of education	6.58	4.54	6.57	4.57
SIZE	Number of household members	3.54	1.73	3.42	1.66
BOTTOM	Per capita household income (the Bottom 40%)	5,119.752	2,773.24	5,679.31	3,025.86
MIDDLE	Per capita household income (the Middle 40%)	12,509.67	7,318.14	13,032.64	7,352.64
TOP	Per capita household income (the Top 20%)	33,041.51	34,069.32	33,833.41	30,999.23
INCOME	Household income (baht)	13,629.41	18,984.19	14,251.78	18,004.36
CHILD5	Number of small children (up to 5 years)	0.33	0.6	0.3	0.57
CHILD15	Number of children (6-15 years)	0.65	0.88	0.61	0.83
RETIREE	Number of retirees (>65 years)	0.28	0.57	0.26	0.55
FEMALE	1, if household head is female	0.28	0.45	0.3	0.46
BANGKOK	1, if locate in Bangkok	0.07	0.25	0.06	0.23
CENTRAL	1, if locate in the central region	0.27	0.44	0.3	0.46
NORTH	1, if locate in the northern region	0.22	0.42	0.23	0.42
NORTH EAST	1, if locate in the north eastern region	0.27	0.44	0.26	0.44
SOUTH	1, if locate in the southern region	0.17	0.38	0.16	0.36

Table 3.4: Household expenditures by household and head of household characteristics in 2002

N = 34,607

	Food	Clothing/ misc. goods	Housing	Health& Medical care	Transportation	Tobacco/alcohol	Total expenditure
By region							
Bangkok	6,012.98	1,759.24	5,075.27	629.12	4,309.93	374.10	18,160.63
Central (Exclude Bangkok)	4,032.69	819.09	2,616.20	308.49	2,339.98	351.65	10,468.11
North	2,903.01	639.46	1,772.65	194.98	1,549.55	213.43	7,273.09
Northeast	3,401.90	658.69	1,890.46	199.29	1,801.19	188.68	8,140.21
South	4,167.55	823.08	2,272.55	251.47	2,283.77	278.26	10,076.68
By gender							
Female	3,452.81	726.77	2,316.03	255.06	1,737.71	152.91	8,641.29
Male	3,867.09	816.72	2,320.31	266.76	2,285.37	316.64	9,872.89
By age							
15-65	3,852.10	425.68	2,341.29	244.78	2,295.52	288.13	11,955.59
65 and above	3,201.10	267.83	2,193.45	351.18	1,290.42	167.51	9,086.96
By level of education							
No education	2,820.75	338.88	1,598.80	191.61	788.55	158.84	5,897.44
Grade 6	3,417.43	532.714	1,940.68	215.33	1423.37	231.92	7,761.44
Grade 12	4,188.07	913.63	2,637.29	296.73	2,582.93	341.30	10,959.95
University or higher	4,985.45	1,851.76	3789.84	447.30	4941.51	385.30	16,401.16
By income/capita							
Less than 2,417	2,940.65	370.36	1,411.84	122.87	748.32	163.09	5,757.14
Between 2,417 and 6,708	3,942.77	740.08	2,349.75	249.48	2,003.32	295.22	9,580.62
Above 6,708	4,947.07	1,727.72	4,071.84	571.50	5,101.77	420.93	16,840.83
Total	3,742.80	789.74	2,319.03	263.24	267.51	9,503.37	3,742.80

Note: 1 baht is approximately 0.025 USD

Table 3.5: Household expenditures by household and head of household characteristics in 2000

N = 24,586

	Food	Clothing/ misc. goods	Housing	Health& Medical care	Transpotation	Tobacco & alcohol	Total expenditure
By region							
Bangkok	3,491.61	2,144.50	8,053.31	509.25	2,162.86	347.37	16,708.91
Central (exclude Bangkok)	2,862.11	1,146.04	4,251.66	286.21	1,459.70	254.14	10,259.85
North	2,273.09	1,088.71	3,123.22	239.84	1,137.01	196.68	8,058.55
Northeast	2,407.40	957.64	3,140.94	207.62	1,088.17	186.94	7,988.72
South	2,731.29	1,364.05	3,898.21	274.5	1,369.85	176.1	9,813.99
By gender							
Female	2,402.12	1,104.33	3,980.55	255.12	990.54	115.95	8,848.60
Male	2,714.31	1,217.24	3,851.56	272.04	1,447.20	255.49	9,757.85
By age							
15-65	2,654.25	1,251.34	3,894.46	250.01	1,393.36	232.27	9,675.69
65 and above	2,462.92	807.75	3,852.27	365.54	885.36	122.44	8,496.28
By level of education							
No education	2,187.71	655.33	2,912.05	213.78	579.19	103.82	6,651.87
Grade 6	2,464.94	827.48	3,251.60	234.82	933.95	163.68	7,876.47
Grade 12	2,771.16	1,270.51	4,138.91	268.26	1,449.87	278.17	10,176.88
University or higher	3,211.81	2,739.13	6,467.30	424.88	2,992.58	358.76	16,194.45
By income/capita							
Less than 2,161 baht	2,190.10	629.93	2,367.24	152.14	575.37	117.45	6,032.23
Between 2,161 and 6,385	2,778.33	1,126.83	4,002.04	264.49	1,253.82	222.65	9,648.16
Above 6,385	3,196.32	2,421.94	6,724.32	504.76	2,942.14	400.68	16,190.15
Total	2,625.72	1,185.20	3,888.17	267.24	1,317.61	215.89	9,499.83

Note: 1 baht is approximately 0.025 USD

Table 3.6: First-step Probit estimation results (2000)

N = 24,586

Parameter	Clothing/ misc. goods		Health & Medical care		Transport-ation		Tobacco & alcohol	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
EDU	0.052**	(0.003)	-0.024**	(0.002)	0.075**	(0.004)	-0.003	(0.002)
SIZE	0.267**	(0.010)	0.082**	(0.007)	0.459**	(0.012)	0.102**	(0.007)
BOTTOM	-0.517**	(0.038)	-0.089**	(0.029)	-0.937**	(0.045)	-0.161**	(0.030)
MIDDLE	-0.256**	(0.032)	0.007	(0.025)	-0.345**	(0.042)	-0.057**	(0.026)
CHILD5	0.014	(0.025)	0.074**	(0.017)	-0.289**	(0.025)	-0.080**	(0.017)
CHILD15	1.099**	(0.030)	-0.027**	(0.012)	-0.182**	(0.018)	-0.064**	(0.012)
RETIREE	-0.204**	(0.019)	0.103**	(0.016)	-0.332**	(0.019)	-0.245**	(0.017)
FEMALE	0.094**	(0.024)	-0.040**	(0.019)	-0.173**	(0.025)	-0.590**	(0.021)
Bangkok	-0.275**	(0.050)	-0.157**	(0.039)	-0.186**	(0.065)	0.320**	(0.041)
Central	-0.134**	(0.035)	-0.050*	(0.026)	-0.143**	(0.039)	0.396**	(0.028)
North	-0.039	(0.036)	0.091**	(0.028)	-0.193**	(0.039)	0.452**	(0.029)
Northeast	-0.200**	(0.036)	-0.026	(0.026)	-0.391**	(0.038)	0.438**	(0.028)
constant	-0.124**	(0.054)	0.416**	(0.042)	0.393**	(0.063)	-0.825**	(0.043)
Log likelihood	-8,669.57		-14,768.2		-7,519.92		-14,399.97	

Note: Asymptotic standard errors in parentheses. ** significant at 5%; * significant at 10% levels.

Table 3.7: First-step Probit estimation results (2002)

N = 34,607

Parameter	Clothing/ misc. goods Coef.	Std. Err.	Health & medical care Coef.	Std. Err.	Transport-ation Coef.	Std. Err.	Tobacco & alcohol Coef.	Std. Err.
EDU	0.060**	(0.003)	-0.018**	(0.002)	0.101**	(0.004)	-0.031**	(0.002)
SIZE	0.313**	(0.009)	0.095**	(0.006)	0.507**	(0.013)	0.121**	(0.006)
BOTTOM	-0.546**	(0.031)	-0.084**	(0.024)	-1.243**	(0.054)	0.254**	(0.024)
MIDDLE	-0.259**	(0.027)	0.029	(0.021)	-0.557**	(0.053)	0.108**	(0.021)
CHILD5	0.063**	(0.022)	0.106**	(0.015)	-0.387**	(0.026)	-0.077**	(0.015)
CHILD15	0.998**	(0.024)	-0.039**	(0.011)	-0.218**	(0.019)	-0.106**	(0.011)
RETIREE	-0.208**	(0.016)	0.125**	(0.014)	-0.347**	(0.019)	-0.088**	(0.013)
FEMALE	0.143**	(0.020)	-0.017	(0.016)	-0.055**	(0.025)	-0.616**	(0.016)
Bangkok	-0.257**	(0.045)	-0.027	(0.035)	-0.017	(0.083)	-0.158**	(0.035)
Central	-0.041	(0.029)	0.095**	(0.022)	-0.106**	(0.040)	0.121**	(0.022)
North	0.060*	(0.031)	0.062**	(0.023)	-0.204**	(0.039)	0.016	(0.023)
Northeast	-0.061**	(0.031)	0.086**	(0.022)	-0.332**	(0.039)	0.094**	(0.022)
constant	-0.355**	(0.046)	0.106**	(0.035)	0.726**	(0.070)	-0.096**	(0.035)
Log likelihood	-12,286.83		-21,956.58		-7,516.89		-22,282.30	

Note: Asymptotic standard errors in parentheses. ** significant at 5%; * significant at 10% levels.

Table 3.8: Second-step QUAIDS estimations in 2000

Parameter	Food (1)	Clothing/ miscellaneous goods (2)	Housing (3)	Health & medical care (4)	Transportation (5)	Tobacco & alcohol (6)
Alpha	0.584**	-0.227**	0.454**	0.012	0.098**	-0.032**
	(0.008)	(0.016)	(0.008)	(0.015)	(0.020)	(0.015)
Gamma1	0.010	-0.398**	0.363**	-0.009	-0.191**	0.005
	(0.026)	(0.041)	(0.026)	(0.042)	(0.041)	(0.044)
gamma2	0.383**	-0.063	-0.989**	0.087	0.444**	0.166**
	(0.056)	(0.077)	(0.056)	(0.074)	(0.078)	(0.080)
Gamma3	-0.242**	0.176**	0.145**	-0.024	-0.068	-0.013
	(0.030)	(0.046)	(0.030)	(0.048)	(0.048)	(0.046)
Gamma4	-0.189**	0.798**	-0.411**	0.009*	0.243**	-0.014
	(0.037)	(0.057)	(0.036)	(0.056)	(0.056)	(0.056)
Gamma5	0.025	-0.525**	0.339**	-0.036	0.026**	-0.068**
	(0.020)	(0.032)	(0.020)	(0.032)	(0.033)	(0.033)
Gamma6	-0.165**	-0.031	0.174**	0.042	-0.069	0.000
	(0.027)	(0.043)	(0.026)	(0.044)	(0.044)	(0.045)
Beta	-0.109**	0.191**	0.009	-0.004	-0.056**	0.032**
	(0.005)	(0.011)	(0.005)	(0.010)	(0.013)	(0.010)
Lambda	-0.001	-0.023**	-0.005**	0.003	0.021**	-0.006**
	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)
Delta	0.000	0.180**	0.000	0.058**	0.154**	0.043**
	(0.106)	(0.004)	(0.106)	(0.004)	(0.005)	(0.003)
Wald test		Chi2(6, 0.05)	= 15.302	Prob>Chi2 =	0.000	

Note: Asymptotic standard errors in parentheses. ** significant at 5%; * significant at 10% levels.

Table 3.9: Second-step QUAIDS estimations in 2002

N = 34,607

Parameter	Food (1)	Clothing/ miscellaneous goods (2)	Housing (3)	Health & medical care (4)	Transportation (5)	Tobacco & alcohol (6)
Alpha	0.706**	-0.165**	0.473**	0.771**	-1.067**	-0.018
	(0.009)	(0.029)	(0.009)	(0.028)	(0.032)	(0.025)
Gamma1	-0.001	-0.226*	-0.056	1.030**	0.427**	0.071
	(0.062)	(0.130)	(0.062)	(0.124)	(0.133)	(0.121)
Gamma2	0.687**	0.072	-0.459**	-2.138**	-1.042**	-0.486**
	(0.129)	(0.155)	(0.129)	(0.155)	(0.156)	(0.154)
Gamma3	0.113*	-0.059	-0.121*	0.020	0.221*	0.040
	(0.061)	(0.130)	(0.061)	(0.125)	(0.134)	(0.122)
Gamma4	-0.629**	0.150	0.390**	-0.392**	2.010**	-0.167
	(0.127)	(0.155)	(0.127)	(0.154)	(0.156)	(0.153)
Gamma5	-0.222**	-0.642**	0.234**	-3.153**	-0.826**	-0.145
	(0.047)	(0.113)	(0.047)	(0.106)	(0.117)	(0.102)
Gamma6	-0.049	0.102	0.147	-1.589**	0.230	-0.306**
	(0.122)	(0.154)	(0.122)	(0.153)	(0.155)	(0.152)
Beta	0.013*	0.137**	-0.064**	0.070**	1.310**	0.075**
	(0.007)	(0.022)	(0.007)	(0.020)	(0.024)	(0.018)
Lambda	-0.022**	-0.006	0.003**	0.003	-0.129**	-0.013**
	(0.001)	(0.004)	(0.001)	(0.004)	(0.004)	(0.003)
Delta	0.000	0.048**	0.000	0.209**	0.179**	0.003
	(0.159)	(0.003)	(0.159)	(0.003)	(0.004)	(0.003)
Wald test	Chi2(6, 0.05) = 15.302 Prob > Chi2 = 0.000					

Note: Asymptotic standard errors in parentheses. ** significant at 5%; * significant at 10% levels.

Table 3.10: Uncompensated Price and Expenditure elasticities for the Thai household demand system

Parameter	Food (1)	Clothing/ miscellaneous goods (2)	Housing (3)	Health & medical care (4)	Transportation (5)	Tobacco & alcohol (6)	Expenditure
Year 2000							
Food	-0.737**	1.184**	-0.633**	-0.621**	0.118*	-0.554**	0.617**
	(0.099)	(0.230)	(0.107)	(0.134)	(0.071)	(0.094)	(0.004)
Clothing & Miscellaneous goods	-2.033	-1.110**	0.720	3.655	-2.466	-0.113	1.352**
	(574.15)	(0.370)	(405.61)	(80.276)	(96.02)	(309.27)	(0.007)
Housing	0.886**	-2.367**	-0.633**	-0.986**	0.814**	0.419**	0.959**
	(0.088)	(0.206)	(0.096)	(0.120)	(0.063)	(0.084)	(0.003)
Health & Medical care	-0.189	1.544*	-0.488	-0.843$^{\tau}$	-0.651**	0.754**	1.157**
	(0.383)	(0.844)	(0.424)	(0.516)	(0.272)	(0.369)	(0.016)
Transportation	-0.932**	2.011**	-0.401**	1.153**	-0.857**	-0.344**	1.246**
	(0.158)	(0.375)	(0.176)	(0.217)	(0.115)	(0.153)	(0.008)
Tobacco & alcohol	-0.404	3.853**	-0.570	-0.304	-1.567**	-0.968*	1.095**
	(0.459)	(1.072)	(0.472)	(0.588)	(0.315)	(0.428)	(0.017)
Year 2002							
Food	-0.859**	1.242**	0.275**	-0.950**	-0.178*	-0.068	0.804**
	(0.123)	(1.242)	(0.275)	(-0.95)	(-0.178)	(-0.068)	(0.002)
Clothing & Miscellaneous goods	-0.558	-0.802	-0.170	0.334	-1.236	0.253	1.215**
	(280.65))	(0.805)	(185.75)	(157.94))	(192.64)	(8.050)	(0.010)
Housing	-0.069	-1.461**	-1.300**	1.323**	0.477**	0.448	0.851**
	(0.162)	(0.524)	(0.154)	(0.492)	(0.121)	(0.447)	(0.003)
Health & medical care	0.796*	-1.683	0.007**	-1.326	-2.481**	-1.252	1.066**
	(0.417)	(1.329)	(0.396)	(1.267)	(0.309)	(1.146)	(0.008)
Transportation	0.194	-0.582	0.051	1.238	-0.724	0.189	1.448**
	(0.348)	(1.096)	(0.332)	(1.073)	(0.259)	(0.967)	(0.008)
Tobacco & alcohol	-0.560	-5.335**	-0.554	-3.475	2.024**	-4.587*	1.050**
	(0.810)	(2.594)	(0.764)	(2.449)	(0.614)	(2.219)	(0.015)

Note: Asymptotic standard errors in parentheses. ** significant at 5%; * significant at 10% , $^{\tau}$ significant at 11% levels.

FIGURES

Figure 3.1: Non-parametric Engel curves for seven household consumption groups in 2002

Figure 3.1.1: Engel curve: food

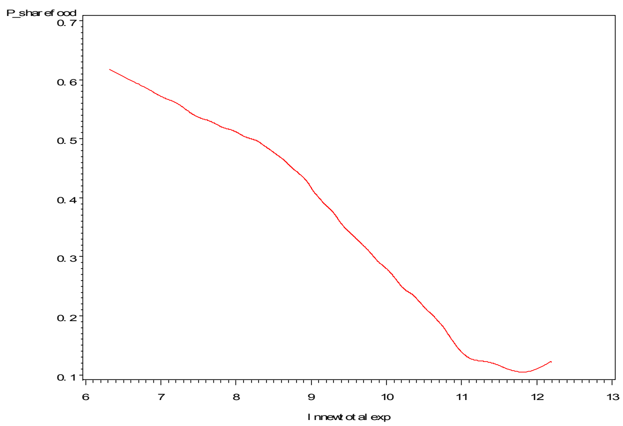


Figure 3.1.2: Engel curve: clothing and miscellaneous goods

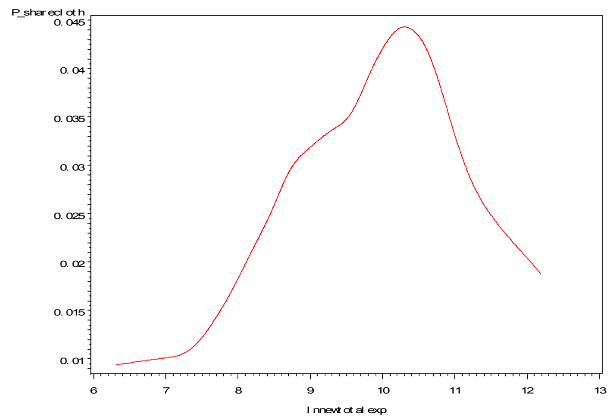


Figure 3.1.3: Engel curve: housing

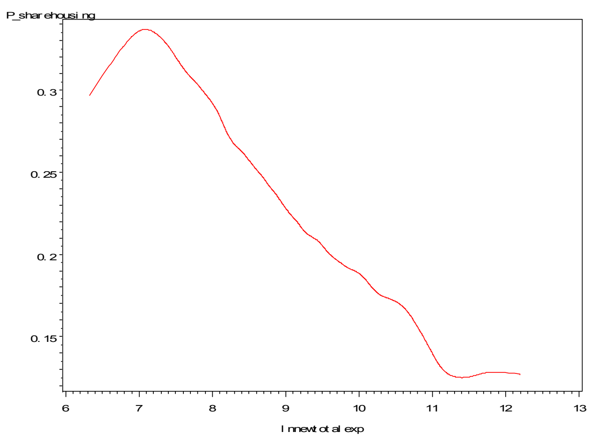


Figure 3.1.4: Engel curve: health /medical care

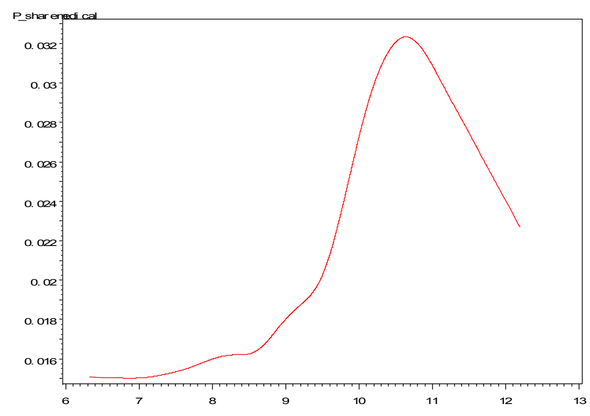


Figure 3.1: Non-parametric Engel curves for seven household consumption groups (cont.)

Figure 3.1.5: Engel curve: transportation

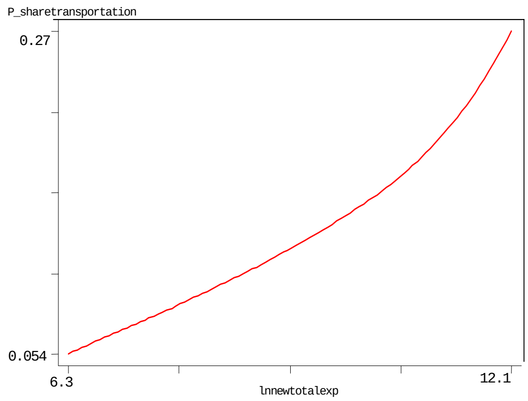
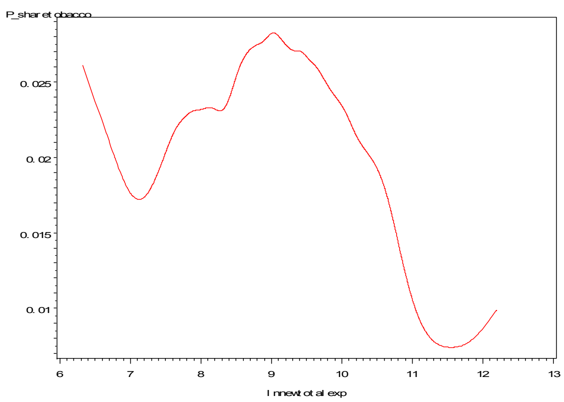


Figure 3.1.6: Engel curve: tobacco and alcohol



APPENDIX

Appendix 3.1: Definitions of expenditure

Expenditure	Descriptions
1. Food	All types of food and non-alcohol beverage
2. Clothing and Miscellaneous goods	Clothing, footwear, leisure (i.e. recreation and sport equipment, musical equipment, and admissions), and education
3. Housing	Final consumption expenditure of resident households taken up for gross rent, fuel and utilities, household textiles, furniture, appliances, and equipments.
4. Health and medical care	Medical supplies, medicine and medical services (outpatients and inpatients services)
5. Transportation	Travel cost, vehicle operation, vehicle purchase, communication services and equipments
6. Tobacco and alcohol	Tobacco and alcohol beverages

APPENDIX

A. GAUSS PROGRAM USED FOR CHAPTER 1

```
/*
THIS PROGRAM DEMONSTRATES BOOTSTRAPPING  DEA ACCORDING TO
APPENDIX 1.6
*/

new;
cls;
graphset;

library  pgraph;
format /rd 12,8;

load data[276,2]=C:\DEAvrs.csv;

Obs = data[.,1];
DEA = data[.,2];

/* After acquiring DEA scores from the DEAP program, the coding starts from STEP
3 */

/*****************************************************************/
/* Define Sm where m is the number of DEA scores that less than 1  */
/*****************************************************************/

/*        Sorting all scores in ascending order              */
data = sortc(data,2);

/*        Count number of row in data matrix with a value less than 1          */
m = 0;
i = 1;

do while i <= rows(DEA);
if abs(DEA[i]) < 1;
   m = m+1;
endif;
   i = i+1;
endo;

print " number of scores with its value less than 1 ";
print m;

/* Select only scores with the value less than 1 and put them in to Sm matrix          */
Sm = data[1:m,.];
print " Set of Sm includes only scores with its value less than 1 ";
print Sm;
```

```
/*******************************************************************/
/*  Determine the bandwidth h based on Sm    */
/*******************************************************************/

/*         4.1 Calculate standard deviation               */

stderr = stdc(Sm);
sigma = stderr[2,1];
print "Standard Error of Sm_DEA scores";
print " ";
print sigma;


/*         4.2 Calculate interquartile              */

P25 = quantile(Sm,.25);
P75 = quantile(Sm,.75);
R13 = P75[.,2] - P25[.,2];
print "         Percentile    ";
print "     25th            75th         Interquantile  ";
print P25[.,2] "    " P75[.,2] "     " R13;


/*         4.3 Determine the bandwidth h              */
/*      h = 0.90*min{sigma, R13/1.34}*n^(-1/5)         */

z = sigma|(R13/1.34);
n = rows(data);
h = 0.9*minc(z)*n^(-1/5);

print " Optimal Bandwidth; h ";
print h;

b = 100;
BBoot = zeros(n-1,b);

jboot = 1;

do while jboot <= b;

/*******************************************************************/
/*    Draw n-1 Bootstrap values         */
/*******************************************************************/

/*         5.1 Draw n-1 Bootstrap values              */
All = zeros(n-1,2);
ThetaBoot = zeros(n-1,1);
i = 1;
idx = round(0.5+ m*rndu(m,1));
print " ";
```

```
do while i <= n-1;
   idx = round(0.5 + m*rndu(1,1));
   Select = Sm[idx,2];
   ThetaBoot[i,1] = ThetaBoot[i,1]+Select;
   All[i,1] = All[i,1]+idx;
   All[i,2] = All[i,2]+Select;
   i = i+1;

endo;

/*      5.2 Create ThetaTilda and (2-ThetaTilda)        */

estar = rndn(n-1,1);
ThetaTildaStar = ThetaBoot + h*estar;
print "  ThetaBoot    estar       ThetaTilda   (2-ThetaTilda) ";
print " ";
print All~ThetaBoot~estar~ThetaTildaStar~2-ThetaTildaStar;

/*      5.3 Define ThetaStar        */

Theta = ThetaTildaStar~(2-ThetaTildaStar);
ThetaStar = minc(Theta');
print " ";
print " ";
print ;

/*      5.4 Rescale Bootstrap data by the following transformation        */
/*      ThetaHatStar = ThetaHatBar + ((1+(h^2)/(sigmahatsqr))^(-0.5))*(ThetaStar-
ThetaHatBar)      */

ThetaHatBar = meanc(DEA);
diff = DEA-ThetaHatBar;
print " ";
sigmahatsqr = sumc(diff^2)/n;

ThetaHatStar = ThetaHatBar + ((1+(h^2)/(sigmahatsqr))^(-0.5))*(ThetaStar-
ThetaHatBar);
ThetaHatStar = sortc(ThetaHatStar,1);
print;

/*  BBoot is the result of applying bootstrap B times   */

/*  as indicate in the STEP 6                    */

BBoot[.,jboot] = BBoot[.,jboot]+ThetaHatStar;

jboot = jboot+1;
endo;
```

```
/*********************************************************************/
/*   Compute ThetaTildabar(n-j) by bagging the B estimates obtained on (BBoot)     */
/*********************************************************************/

/*    ThetaTildaBar is an average across B columns of BBoot
*/
ThetaTildaBar = meanc(BBoot');
print ThetaTildaBar;

end;
```

B. GAUSS PROGRAM USED FOR CHAPTER 3

```
/*----------------------------------------------------------------------------------------------------
        THIS GAUSS PROGRAM PERFORMS A QUAIDS MODEL WITH
CENSORED DEMAND PROBLEMS OF THAI HOUSEHOLD DEMAND.

        THE TWO-STEP QUAIDS ESTIMATION WAS INTRODUCED BY
SHONKWILER AND YEN, 2002.

----------------------------------------------------------------------------------------------------*/

/* Note: the program utilize the 2002 SES dataset   */

new;
cls;

mat = xlsreadm("data02drop26.xls", "A2:Y34608", 1, 0);
n = 34607; @ Y34608 @

/* ANALYTICAL ANALYSIS  */
/*  PART 1: INITIALIZED VARIABLES  */

p1 = mat[.,1];
p26 = mat[.,2];
p3 = mat[.,3];
p4 = mat[.,4];
p5 = mat[.,5];
p7 = mat[.,6];
ms = mat[.,7];
w1 = mat[.,8];  w26 = mat[.,9];
w3 = mat[.,10]; w4 = mat[.,11];
w5 = mat[.,12]; w7 = mat[.,13];

w = w1~w26~w3~w4~w5~w7;
wvec = w1|w26|w3|w4|w5|w7;

pdf1 = mat[.,14];   pdf26 = mat[.,15];
pdf3 = mat[.,16];   pdf4 = mat[.,17];
pdf5 = mat[.,18];   pdf7 = mat[.,19];
pdf = pdf1~pdf26~pdf3~pdf4~pdf5~pdf7;
cdf1 = ones(n,1);   cdf26 = mat[.,21];
cdf3 = ones(n,1);   cdf4 = mat[.,23];
cdf5 = mat[.,24];   cdf7 = mat[.,25];
cdf = cdf1~cdf26~cdf3~cdf4~cdf5~cdf7;

p = p1~p26~p3~p4~p5~p7;
p = p/100;

print;
a0 = minc(ln(Ms));
a = ones(6,1);
gam = -0.05*eye(6);
```

```
beta = 0.05*ones(6,1);
lam = -0.05*ones(6,1);
phi = zeros(6,1);


gam = 0.05*zeros(6,6);
beta = 0.05*ones(6,1);
lam = 0.05*ones(6,1);
phi = zeros(6,1);

params = 0*a~gam~beta~lam~phi;
lnp = ln(p);
cls;


/* PART 2: STRUCTURAL EQUATIONS */

pp = lnp[.,1].*lnp[.,1];
pp = pp~lnp[.,1].*lnp[.,2];
pp = pp~lnp[.,1].*lnp[.,3];
pp = pp~lnp[.,1].*lnp[.,4];
pp = pp~lnp[.,1].*lnp[.,5];
pp = pp~lnp[.,1].*lnp[.,6];

pp = pp~lnp[.,2].*lnp[.,1];
pp = pp~lnp[.,2].*lnp[.,2];
pp = pp~lnp[.,2].*lnp[.,3];
pp = pp~lnp[.,2].*lnp[.,4];
pp = pp~lnp[.,2].*lnp[.,5];
pp = pp~lnp[.,2].*lnp[.,6];

pp = pp~lnp[.,3].*lnp[.,1];
pp = pp~lnp[.,3].*lnp[.,2];
pp = pp~lnp[.,3].*lnp[.,3];
pp = pp~lnp[.,3].*lnp[.,4];
pp = pp~lnp[.,3].*lnp[.,5];
pp = pp~lnp[.,3].*lnp[.,6];

pp = pp~lnp[.,4].*lnp[.,1];
pp = pp~lnp[.,4].*lnp[.,2];
pp = pp~lnp[.,4].*lnp[.,3];
pp = pp~lnp[.,4].*lnp[.,4];
pp = pp~lnp[.,4].*lnp[.,5];
pp = pp~lnp[.,4].*lnp[.,6];

pp = pp~lnp[.,5].*lnp[.,1];
pp = pp~lnp[.,5].*lnp[.,2];
pp = pp~lnp[.,5].*lnp[.,3];
pp = pp~lnp[.,5].*lnp[.,4];
pp = pp~lnp[.,5].*lnp[.,5];
```

```
pp = pp~lnp[.,5].*lnp[.,6];

pp = pp~lnp[.,6].*lnp[.,1];
pp = pp~lnp[.,6].*lnp[.,2];
pp = pp~lnp[.,6].*lnp[.,3];
pp = pp~lnp[.,6].*lnp[.,4];
pp = pp~lnp[.,6].*lnp[.,5];
pp = pp~lnp[.,6].*lnp[.,6];

pgp = pp[.,1:6]*gam[1,.]';
pgp = pgp+pp[.,7:12]*gam[2,.]';
pgp = pgp+pp[.,13:18]*gam[3,.]';
pgp = pgp+pp[.,19:24]*gam[4,.]';
pgp = pgp+pp[.,25:30]*gam[5,.]';
pgp = pgp+pp[.,31:36]*gam[6,.]';

lnap = a0 + lnp*params[.,1] + 0.5*pgp;
lnbp = lnp*params[.,8];
bp = exp(lnbp);
lnmp = (ln(Ms) - lnap);
lnmp2 = (lnmp^2)./bp;

lnpipj =
lnp[.,1].*lnp~lnp[.,2].*lnp~lnp[.,3].*lnp~lnp[.,4].*lnp~lnp[.,5].*lnp~lnp[.,6].*lnp;

wvec = w1|w26|w3|w4|w5|w7;


                              /* 1st EQUATION DERIVATIVE */
h1a = (ones(n,1)~zeros(n,5)) - (params[1,8].*(lnp)) -
2*params[1,9]./bp.*(lnmp).*(lnp);
            /*    n x 6 matrix           beta1              lamda1      */
h1g = (lnp~zeros(n,30)) - ((0.5*params[1,8])+(params[1,9]./bp.*(lnmp))).*(lnpipj);
            /*   n x 36 matrix          beta1                 lamda1            */
h1b = (lnmp~zeros(n,5))+ params[1,9].*(lnmp2).*(-lnp);
            /*                  lamda1       */
h1phi = pdf[.,1];


                              /* 2nd EQUATION DERIVATIVE */
h2a = (zeros(n,1)~ones(n,1)~zeros(n,4)) - (params[2,8].*(lnp)) -
2*params[2,9]./bp.*(lnmp).*(lnp);
         /*    n x 6 matrix                beta2              lamda2      */
h2g = (zeros(n,6)~lnp~zeros(n,24)) - ((0.5*params[2,8]) +
         (params[2,9]./bp.*(lnmp))).*(lnpipj);
         /*   beta2      lamda2           */
h2b = (zeros(n,1)~lnmp~zeros(n,4))+ params[2,9].*(lnmp2).*(-lnp);
                /*                          lamda2       */
h2phi = pdf[.,2];
```

```
                    /* 3rd EQUATION DERIVATIVE */
h3a = (zeros(n,2)~ones(n,1)~zeros(n,3)) - (params[3,8].*(lnp)) -
2*params[3,9]./bp.*(lnmp).*(lnp);
      /*    n x 6 matrix            beta3           lamda3     */
h3g = (zeros(n,12)~lnp~zeros(n,18))- ((0.5*params[3,8])+ (params[3,9]./bp.
      *(lnmp))).*(lnpipj);
      /*                         beta3            lamda3           */
h3b = (zeros(n,2)~lnmp~zeros(n,3))+ params[3,9].*(lnmp2).*(-lnp);
            /*                    lamda3      */
h3phi = pdf[.,3];


                    /* 4th EQUATION DERIVATIVE */
h4a = (zeros(n,3)~ones(n,1)~zeros(n,2)) - (params[4,8].*(lnp)) –
      2*params[4,9]./bp.*(lnmp).*(lnp);
      /*    n x 6 matrix            beta4           lamda4     */
h4g = (zeros(n,18)~lnp~zeros(n,12)) - ((0.5*params[4,8])+
      (params[4,9]./bp.*(lnmp))).*(lnpipj);
/*              beta4     lamda4          */
h4b = (zeros(n,3)~lnmp~zeros(n,2))+ params[4,9].*(lnmp2).*(-lnp);
          /*                   lamda4      */
h4phi = pdf[.,4];


                    /* 5th EQUATION DERIVATIVE */
h5a = (zeros(n,4)~ones(n,1)~zeros(n,1)) - (params[5,8].*(lnp)) -
2*params[5,9]./bp.*(lnmp).*(lnp);
/*    n x 6 matrix             beta5          lamda5     */
h5g = (zeros(n,24)~lnp~zeros(n,6)) -
((0.5*params[5,8])+(params[5,9]./bp.*(lnmp))).*(lnpipj);
/*                         beta5     lamda5          */
h5b = (zeros(n,4)~lnmp~zeros(n,1))+ params[5,9].*(lnmp2).*(-lnp);
/*                      lamda5      */
h5phi = pdf[.,5];

                    /* 6th EQUATION DERIVATIVE */
h6a = (zeros(n,5)~ones(n,1)) - (params[6,8].*(lnp)) -
2*params[6,9]./bp.*(lnmp).*(lnp);
            /*    n x 6 matrix       beta6           lamda6     */
h6g = (zeros(n,30)~lnp) - ((0.5*params[6,8])+(params[6,9]./bp.*(lnmp))).*(lnpipj);
            /*               beta6      lamda6          */
h6b = (zeros(n,5)~lnmp)+ params[6,9].*(lnmp2).*(-lnp);
            /*             lamda6      */
h6phi = pdf[.,6];


/* WITH RESPECT TO LAMDA */
hlamda = lnmp2;
```

```
                                /* 1st EQUATION GRADIENT */
grad1 = h1a[.,1]~h1g[.,1:6]~h1b[.,1]~hlamda~zeros(n,1)~
/*      dalpha1  dgam11-16  dbeta1   dlamda        */
        h1a[.,2]~h1g[.,7:12]~h1b[.,2]~zeros(n,2)~
/*      dalpha2  dgam21-26  dbeta2                 */
        h1a[.,3]~h1g[.,13:18]~h1b[.,3]~zeros(n,2)~
/*      dalpha3  dgam31-36  dbeta3                 */
        h1a[.,4]~h1g[.,19:24]~h1b[.,4]~zeros(n,2)~
/*      dalpha4  dgam41-46  dbeta4                 */
        h1a[.,5]~h1g[.,25:30]~h1b[.,5]~zeros(n,2)~
/*      dalpha5  dgam51-56  dbeta5                 */
        h1a[.,6]~h1g[.,31:36]~h1b[.,6]~zeros(n,2);
/*      dalpha6  dgam61-66  dbeta6                 */


                                /* 2nd EQUATION GRADIENT */
grad26 = h2a[.,1]~h2g[.,1:6]~h2b[.,1]~zeros(n,2)~
/*      dalpha1  dgam11-16  dbeta1   dlamda        */
        h2a[.,2]~h2g[.,7:12]~h2b[.,2]~hlamda~h2phi~
/*      dalpha2  dgam21-26  dbeta2                 */
        h2a[.,3]~h2g[.,13:18]~h2b[.,3]~zeros(n,2)~
/*      dalpha3  dgam31-36  dbeta3                 */
        h2a[.,4]~h2g[.,19:24]~h2b[.,4]~zeros(n,2)~
/*      dalpha4  dgam41-46  dbeta4                 */
        h2a[.,5]~h2g[.,25:30]~h2b[.,5]~zeros(n,2)~
/*      dalpha5  dgam51-56  dbeta5                 */
        h2a[.,6]~h2g[.,31:36]~h2b[.,6]~zeros(n,2);
/*      dalpha6  dgam61-66  dbeta6                 */


                                /* 3rd EQUATION GRADIENT */
grad3 = h3a[.,1]~h3g[.,1:6]~h3b[.,1]~zeros(n,2)~
/*      dalpha1  dgam11-16  dbeta1   dlamda        */
        h3a[.,2]~h3g[.,7:12]~h3b[.,2]~zeros(n,2)~
/*      dalpha2  dgam21-26  dbeta2                 */
        h3a[.,3]~h3g[.,13:18]~h3b[.,3]~hlamda~zeros(n,1)~
/*      dalpha3  dgam31-36  dbeta3                 */
        h3a[.,4]~h3g[.,19:24]~h3b[.,4]~zeros(n,2)~
/*      dalpha4  dgam41-46  dbeta4                 */
        h3a[.,5]~h3g[.,25:30]~h3b[.,5]~zeros(n,2)~
/*      dalpha5  dgam51-56  dbeta5                 */
        h3a[.,6]~h3g[.,31:36]~h3b[.,6]~zeros(n,2);
/*      dalpha6  dgam61-66  dbeta6                 */

                                /* 4th EQUATION GRADIENT */
grad4 = h4a[.,1]~h4g[.,1:6]~h4b[.,1]~zeros(n,2)~
/*      dalpha1  dgam11-16  dbeta1   dlamda        */
        h4a[.,2]~h4g[.,7:12]~h4b[.,2]~zeros(n,2)~
```

```
/*      dalpha2  dgam21-26  dbeta2                 */
        h4a[.,3]~h4g[.,13:18]~h4b[.,3]~zeros(n,2)~
/*      dalpha3  dgam31-36  dbeta3                 */
        h4a[.,4]~h4g[.,19:24]~h4b[.,4]~hlamda~h4phi~
/*      dalpha4  dgam41-46  dbeta4                 */
        h4a[.,5]~h4g[.,25:30]~h4b[.,5]~zeros(n,2)~
/*      dalpha5  dgam51-56  dbeta5                 */
        h4a[.,6]~h4g[.,31:36]~h4b[.,6]~zeros(n,2);
/*      dalpha6  dgam61-66  dbeta6                 */

                                /* 5th EQUATION GRADIENT */
grad5 = h5a[.,1]~h5g[.,1:6]~h5b[.,1]~zeros(n,2)~
/*      dalpha1  dgam11-16  dbeta1   dlamda        */
        h5a[.,2]~h5g[.,7:12]~h5b[.,2]~zeros(n,2)~
/*      dalpha2  dgam21-26  dbeta2                 */
        h5a[.,3]~h5g[.,13:18]~h5b[.,3]~zeros(n,2)~
/*      dalpha3  dgam31-36  dbeta3                 */
        h5a[.,4]~h5g[.,19:24]~h5b[.,4]~zeros(n,2)~
/*      dalpha4  dgam41-46  dbeta4                 */
        h5a[.,5]~h5g[.,25:30]~h5b[.,5]~hlamda~h5phi~
/*      dalpha5  dgam51-56  dbeta5                 */
        h5a[.,6]~h5g[.,31:36]~h5b[.,6]~zeros(n,2);
/*      dalpha6  dgam61-66  dbeta6                 */


                                /* 6th EQUATION GRADIENT */
grad6 = h6a[.,1]~h6g[.,1:6]~h6b[.,1]~zeros(n,2)~
/*      dalpha1  dgam11-16  dbeta1   dlamda        */
        h6a[.,2]~h6g[.,7:12]~h6b[.,2]~zeros(n,2)~
/*      dalpha2  dgam21-26  dbeta2                 */
        h6a[.,3]~h6g[.,13:18]~h6b[.,3]~zeros(n,2)~
/*      dalpha3  dgam31-36  dbeta3                 */
        h6a[.,4]~h6g[.,19:24]~h6b[.,4]~zeros(n,2)~
/*      dalpha4  dgam41-46  dbeta4                 */
        h6a[.,5]~h6g[.,25:30]~h6b[.,5]~zeros(n,2)~
/*      dalpha5  dgam51-56  dbeta5                 */
        h6a[.,6]~h6g[.,31:36]~h6b[.,6]~hlamda~h6phi;
/*      dalpha6  dgam61-66  dbeta6                 */

grad26 = (cdf26).*grad26;
grad4 = (cdf4).*grad4;
grad5 = (cdf5).*grad5;
grad6 = (cdf7).*grad6;

grad26[.,10] = grad26[.,10]./(cdf26); grad26[.,20] = grad26[.,20]./(cdf26);
grad26[.,30] = grad26[.,30]./(cdf26); grad26[.,40] = grad26[.,40]./(cdf26);
grad26[.,50] = grad26[.,50]./(cdf26); grad26[.,60] = grad26[.,60]./(cdf26);
grad4[.,10] = grad4[.,10]./(cdf4); grad4[.,20] = grad4[.,20]./(cdf4); grad4[.,30] =
grad4[.,30]./(cdf4); grad4[.,40] = grad4[.,40]./(cdf4); grad4[.,50] = grad4[.,50]./(cdf4);
grad4[.,60] = grad4[.,60]./(cdf4);
```

```
grad5[.,10] = grad5[.,10]./(cdf5); grad5[.,20] = grad5[.,20]./(cdf5); grad5[.,30] =
grad5[.,30]./(cdf5); grad5[.,40] = grad5[.,40]./(cdf5); grad5[.,50] = grad5[.,50]./(cdf5);
grad5[.,60] = grad5[.,60]./(cdf5);
grad6[.,10] = grad6[.,10]./(cdf7); grad6[.,20] = grad6[.,20]./(cdf7); grad6[.,30] =
grad6[.,30]./(cdf7); grad6[.,40] = grad6[.,40]./(cdf7); grad6[.,50] = grad6[.,50]./(cdf7);
grad6[.,60] = grad6[.,60]./(cdf7);

/* ---------------------------------------------------------------------------------------------------
        PROCEDURE TO CALCULATE INCOME, OWN-PRICE, AND CROSS-
PRICE ELASTICITIES
---------------------------------------------------------------------------------------------------*/
new;
cls;

library pgraph;

mat = xlsreadm("data02drop26.xls", "A2:Y34608", 1, 0);
bhat = xlsreadm("data02drop26.xls", "B2:B61", 2, 0);
sd = xlsreadm("data02drop26.xls", "C2:C61", 2, 0);

/*  PART 1: INITIALIZED VARIABLES  */

n = 34607;
p1 = mat[.,1];
p26 = mat[.,2];
p3 = mat[.,3];
p4 = mat[.,4];
p5 = mat[.,5];
p7 = mat[.,6];
Ms = mat[.,7];
w1 = mat[.,8];  w26 = mat[.,9];
w3 = mat[.,10]; w4 = mat[.,11];
w5 = mat[.,12]; w7 = mat[.,13];

w = w1~w26~w3~w4~w5~w7;
wvec = w1|w26|w3|w4|w5|w7;

pdf1 = mat[.,14];   pdf26 = mat[.,15];
pdf3 = mat[.,16];   pdf4 = mat[.,17];
pdf5 = mat[.,18];   pdf7 = mat[.,19];
pdf = pdf1~pdf26~pdf3~pdf4~pdf5~pdf7;
cdf1 = ones(n,1);   cdf26 = mat[.,21];
cdf3 = ones(n,1);   cdf4 = mat[.,23];
cdf5 = mat[.,24];   cdf7 = mat[.,25];
cdf = cdf1~cdf26~cdf3~cdf4~cdf5~cdf7;

p = p1~p26~p3~p4~p5~p7;
p = p/100;

print;
```

```
a0 = minc(ln(Ms));
a = ones(6,1);
gam = -0.05*eye(6);
beta = 0.05*ones(6,1);
lam = -0.05*ones(6,1);
phi = zeros(6,1);

params = 0*a~gam~beta~lam~phi;
lnp = ln(p);

print;
print;
print;
print;

/* PART 2: STRUCTURAL EQUATIONS */

pp = lnp[.,1].*lnp[.,1];
pp = pp~lnp[.,1].*lnp[.,2];
pp = pp~lnp[.,1].*lnp[.,3];
pp = pp~lnp[.,1].*lnp[.,4];
pp = pp~lnp[.,1].*lnp[.,5];
pp = pp~lnp[.,1].*lnp[.,6];

pp = pp~lnp[.,2].*lnp[.,1];
pp = pp~lnp[.,2].*lnp[.,2];
pp = pp~lnp[.,2].*lnp[.,3];
pp = pp~lnp[.,2].*lnp[.,4];
pp = pp~lnp[.,2].*lnp[.,5];
pp = pp~lnp[.,2].*lnp[.,6];

pp = pp~lnp[.,3].*lnp[.,1];
pp = pp~lnp[.,3].*lnp[.,2];
pp = pp~lnp[.,3].*lnp[.,3];
pp = pp~lnp[.,3].*lnp[.,4];
pp = pp~lnp[.,3].*lnp[.,5];
pp = pp~lnp[.,3].*lnp[.,6];

pp = pp~lnp[.,4].*lnp[.,1];
pp = pp~lnp[.,4].*lnp[.,2];
pp = pp~lnp[.,4].*lnp[.,3];
pp = pp~lnp[.,4].*lnp[.,4];
pp = pp~lnp[.,4].*lnp[.,5];
pp = pp~lnp[.,4].*lnp[.,6];

pp = pp~lnp[.,5].*lnp[.,1];
pp = pp~lnp[.,5].*lnp[.,2];
pp = pp~lnp[.,5].*lnp[.,3];
pp = pp~lnp[.,5].*lnp[.,4];
pp = pp~lnp[.,5].*lnp[.,5];
```

```
pp = pp~lnp[.,5].*lnp[.,6];

pp = pp~lnp[.,6].*lnp[.,1];
pp = pp~lnp[.,6].*lnp[.,2];
pp = pp~lnp[.,6].*lnp[.,3];
pp = pp~lnp[.,6].*lnp[.,4];
pp = pp~lnp[.,6].*lnp[.,5];
pp = pp~lnp[.,6].*lnp[.,6];

print;

pgp = pp[.,1:6]*gam[1,.]';
pgp = pgp~pp[.,7:12]*gam[2,.]';
pgp = pgp~pp[.,13:18]*gam[3,.]';
pgp = pgp~pp[.,19:24]*gam[4,.]';
pgp = pgp~pp[.,25:30]*gam[5,.]';
pgp = pgp~pp[.,31:36]*gam[6,.]';

lnap = a0 + lnp*params[.,1] + 0.5*pgp;
lnbp = lnp*params[.,8];
bp = exp(lnbp);
print;
lnmp = (ln(Ms) - lnap);
lnmp2 = (lnmp^2)./bp;

lnpipj =
lnp[.,1].*lnp~lnp[.,2].*lnp~lnp[.,3].*lnp~lnp[.,4].*lnp~lnp[.,5].*lnp~lnp[.,6].*lnp;

wvec = w1|w26|w3|w4|w5|w7;

/* Data prepared for calculating elasticities */
alphat = bhat[1]|bhat[11]|bhat[21]|bhat[31]|bhat[41]|bhat[51];
gamhat = bhat[2:7]'|bhat[12:17]'|bhat[22:27]'|bhat[32:37]'|bhat[42:47]'|bhat[52:57]';
lnap = a0*ones(n,1) + lnp*alphat + 0.5*quadform(lnp,gamhat);
betahat = bhat[8]|bhat[18]|bhat[28]|bhat[38]|bhat[48]|bhat[58];
lnbp = lnp*betahat;
bp = exp(lnbp);
lnmp = (ln(Ms) - lnap);
lamhat = bhat[9]|bhat[19]|bhat[29]|bhat[39]|bhat[49]|bhat[59];

/* STRUCTURAL EQUATIONS FOR CALCUALTING ELASTICITIES */

/* Structural equations for income elasticities */
mu1 = cdf1.*(betahat[1]*ones(n,1)+2*(lamhat[1]./bp).*lnmp);
mu26 = cdf26.*(betahat[2]*ones(n,1)+2*(lamhat[2]./bp).*lnmp);
mu3 = cdf3.*(betahat[3]*ones(n,1)+2*(lamhat[3]./bp).*lnmp);
mu4 = cdf4.*(betahat[4]*ones(n,1)+2*(lamhat[4]./bp).*lnmp);
mu5 = cdf5.*(betahat[5]*ones(n,1)+2*(lamhat[5]./bp).*lnmp);
mu7 = cdf7.*(betahat[6]*ones(n,1)+2*(lamhat[6]./bp).*lnmp);
```

```
e1 = mu1./what1 + 1; e2 = mu26./what26 + 1; e3 = mu3./what3 + 1;
e4 = mu4./what4 + 1; e5 = mu5./what5 + 1; e6 = mu7./what7 + 1;

e1 = meanc(e1); e2 = meanc(e2); e3 = meanc(e3);
e4 = meanc(e4); e5 = meanc(e5); e6 = meanc(e6);


QUAIDSgrad = gradp(&SYQUAIDS,bhat);
qgrad1 = quaidsgrad[1:n,.]; qgrad2 = quaidsgrad[n+1:2*n,.]; qgrad3 =
quaidsgrad[(2*n)+1:3*n,.];
qgrad4 = quaidsgrad[(3*n)+1:4*n,.]; qgrad5 = quaidsgrad[(4*n)+1:5*n,.]; qgrad6 =
quaidsgrad[(5*n)+1:6*n,.];

cov = invsigma[1,1].*qgrad1 + invsigma[1,2].*qgrad2 + invsigma[1,3].*qgrad3 +
invsigma[1,4].*qgrad4 + invsigma[1,5].*qgrad5 + invsigma[1,6].*qgrad6;
cov = cov|invsigma[2,1].*qgrad1 + invsigma[2,2].*qgrad2 + invsigma[2,3].*qgrad3 +
invsigma[2,4].*qgrad4 + invsigma[2,5].*qgrad5 + invsigma[2,6].*qgrad6;
cov = cov|invsigma[3,1].*qgrad1 + invsigma[3,2].*qgrad2 + invsigma[3,3].*qgrad3 +
invsigma[3,4].*qgrad4 + invsigma[3,5].*qgrad5 + invsigma[3,6].*qgrad6;
cov = cov|invsigma[4,1].*qgrad1 + invsigma[4,2].*qgrad2 + invsigma[4,3].*qgrad3 +
invsigma[4,4].*qgrad4 + invsigma[4,5].*qgrad5 + invsigma[4,6].*qgrad6;
cov = cov|invsigma[5,1].*qgrad1 + invsigma[5,2].*qgrad2 + invsigma[5,3].*qgrad3 +
invsigma[5,4].*qgrad4 + invsigma[5,5].*qgrad5 + invsigma[5,6].*qgrad6;
cov = cov|invsigma[6,1].*qgrad1 + invsigma[6,2].*qgrad2 + invsigma[6,3].*qgrad3 +
invsigma[6,4].*qgrad4 + invsigma[6,5].*qgrad5 + invsigma[6,6].*qgrad6;

cov =
qgrad1'cov[1:n,.]+qgrad2'cov[n+1:2*n,.]+qgrad3'cov[(2*n)+1:3*n,.]+qgrad4'cov[(3*n
)+1:4*n,.]+qgrad5'cov[(4*n)+1:5*n,.]+qgrad6'cov[(5*n)+1:6*n,.];
cov = cov[.,1:9]~cov[.,11:29]~cov[.,31:60];
cov = cov[1:9,.]|cov[11:29,.]|cov[31:60,.];
cov = invpd(cov);

/* cov = invpd(QUAIDSgrad'cov*QUAIDSgrad); */

incelast = Inc_elast(bhat);
incgrad = gradp(&Inc_elast,bhat);
incgrad = incgrad[.,1:9]~incgrad[.,11:29]~incgrad[.,31:60];
incstd = sqrt(diag(incgrad*cov*incgrad'));

print;
print " Income Elasticity ";
print "     e1        e2       e3        e4         e5       e6       e7 ";
print incelast[1]~incelast[2]~incelast[3]~incelast[4]~incelast[5]~incelast[6];
print;
print incstd[1]~incstd[2]~incstd[3]~incstd[4]~incstd[5]~incstd[6];
print;
print e1/incstd[1]~e2/incstd[2]~e3/incstd[3]~e4/incstd[4]~e6/incstd[6]~e6/incstd[6];

print;
```

```
/*****************************************************************/
diaggam = diag(gamhat);

/* Structural equations for own-price elasticities */

mu11 = cdf1.*(diaggam[1]*ones(n,1)-mu1.*(alphat[1]*ones(n,1)+lnp*gamhat[1,.]')-
((lamhat[1]*betahat[1]./bp).*(lnmp)^2));
mu22 = cdf26.*(diaggam[2]*ones(n,1)-
mu26.*(alphat[2]*ones(n,1)+lnp*gamhat[2,.]')-
((lamhat[2]*betahat[2]./bp).*(lnmp)^2));
mu33 = cdf3.*(diaggam[3]*ones(n,1)-mu3.*(alphat[3]*ones(n,1)+lnp*gamhat[3,.]')-
((lamhat[3]*betahat[3]./bp).*(lnmp)^2));
mu44 = cdf4.*(diaggam[4]*ones(n,1)-mu4.*(alphat[4]*ones(n,1)+lnp*gamhat[4,.]')-
((lamhat[4]*betahat[4]./bp).*(lnmp)^2));
mu55 = cdf5.*(diaggam[5]*ones(n,1)-mu5.*(alphat[5]*ones(n,1)+lnp*gamhat[5,.]')-
((lamhat[5]*betahat[5]./bp).*(lnmp)^2));
mu66 = cdf7.*(diaggam[6]*ones(n,1)-mu7.*(alphat[6]*ones(n,1)+lnp*gamhat[6,.]')-
((lamhat[6]*betahat[6]./bp).*(lnmp)^2));

e11 = (mu11./what1)-1; e22 = (mu22./what26)-1; e33 = (mu33./what3)-1;
e44 = (mu44./what4)-1; e55 = (mu55./what5)-1; e66 = (mu66./what7)-1;

obs = seqa(1,1,n);
dw1 = w1-what1;
dw2 = w26-what26;
dw3 = w3-what3;
dw4 = w4-what4;
dw5 = w5-what5;
dw6 = w7-what7;

e11 = meanc(e11); e22 = meanc(e22); e33 = meanc(e33); e44 = meanc(e44); e55 =
meanc(e55); e66 = meanc(e66);

/* Calculating t-stat   */
OPelast = Ownprice(bhat);
OPgrad = gradp(&Ownprice,bhat);
OPgrad = OPgrad[.,1:9]~OPgrad[.,11:29]~OPgrad[.,31:60];
OPstd = sqrt(diag(OPgrad*cov*OPgrad'));

print " Own-price Elasticity   ";
print "     e11          e2626          e33          e44          e55          e77 ";
print;
print Opelast[1]~Opelast[2]~Opelast[3]~Opelast[4]~Opelast[5]~Opelast[6];
print;
print OPstd[1]~OPstd[2]~OPstd[3]~OPstd[4]~OPstd[5]~OPstd[6];
print (OPelast./OPstd)';

/*****************************************************************/
```

```
/* Structural equations for cross-price elasticities */

   mu12 = cdf1.*(gamhat[1,2]*ones(n,1)-
mu1.*(alphat[2]*ones(n,1)+lnp*gamhat[2,.]')-(lamhat[1]*betahat[2]./bp.*(lnmp)^2));
   mu13 = cdf1.*(gamhat[1,3]*ones(n,1)-
mu1.*(alphat[3]*ones(n,1)+lnp*gamhat[3,.]')-(lamhat[1]*betahat[3]./bp.*(lnmp)^2));
   mu14 = cdf1.*(gamhat[1,4]*ones(n,1)-
mu1.*(alphat[4]*ones(n,1)+lnp*gamhat[4,.]')-(lamhat[1]*betahat[4]./bp.*(lnmp)^2));
   mu15 = cdf1.*(gamhat[1,5]*ones(n,1)-
mu1.*(alphat[5]*ones(n,1)+lnp*gamhat[5,.]')-(lamhat[1]*betahat[5]./bp.*(lnmp)^2));
   mu16 = cdf1.*(gamhat[1,6]*ones(n,1)-
mu1.*(alphat[6]*ones(n,1)+lnp*gamhat[6,.]')-(lamhat[1]*betahat[6]./bp.*(lnmp)^2));

   e12 = mu12./what1; e13 = mu13./what1; e14 = mu14./what1; e15 = mu15./what1;
e16 = mu16./what1;

   e12 = meanc(e12); e13 = meanc(e13); e14 = meanc(e14); e15 = meanc(e15); e16 =
meanc(e16);

   mu21 = cdf26.*(gamhat[2,1]*ones(n,1)-
mu26.*(alphat[1]*ones(n,1)+lnp*gamhat[1,.]')-
(lamhat[2]*betahat[1]./bp.*(lnmp)^2));
   mu23 = cdf26.*(gamhat[2,3]*ones(n,1)-
mu26.*(alphat[3]*ones(n,1)+lnp*gamhat[3,.]')-
(lamhat[2]*betahat[3]./bp.*(lnmp)^2));
   mu24 = cdf26.*(gamhat[2,4]*ones(n,1)-
mu26.*(alphat[4]*ones(n,1)+lnp*gamhat[4,.]')-
(lamhat[2]*betahat[4]./bp.*(lnmp)^2));
   mu25 = cdf26.*(gamhat[2,5]*ones(n,1)-
mu26.*(alphat[5]*ones(n,1)+lnp*gamhat[5,.]')-
(lamhat[2]*betahat[5]./bp.*(lnmp)^2));
   mu26 = cdf26.*(gamhat[2,6]*ones(n,1)-
mu26.*(alphat[6]*ones(n,1)+lnp*gamhat[6,.]')-
(lamhat[2]*betahat[6]./bp.*(lnmp)^2));

   e21 = mu21./what26; e23 = mu23./what26; e24 = mu24./what26; e25 =
mu25./what26; e26 = mu26./what26;

   e21 = meanc(e21); e23 = meanc(e23); e24 = meanc(e24); e25 = meanc(e25); e26 =
meanc(e26);

   mu31 = cdf3.*(gamhat[3,1]*ones(n,1)-
mu3.*(alphat[1]*ones(n,1)+lnp*gamhat[1,.]')-(lamhat[3]*betahat[1]./bp.*(lnmp)^2));
   mu32 = cdf3.*(gamhat[3,2]*ones(n,1)-
mu3.*(alphat[2]*ones(n,1)+lnp*gamhat[2,.]')-(lamhat[3]*betahat[2]./bp.*(lnmp)^2));
   mu34 = cdf3.*(gamhat[3,4]*ones(n,1)-
mu3.*(alphat[4]*ones(n,1)+lnp*gamhat[4,.]')-(lamhat[3]*betahat[4]./bp.*(lnmp)^2));
   mu35 = cdf3.*(gamhat[3,5]*ones(n,1)-
mu3.*(alphat[5]*ones(n,1)+lnp*gamhat[5,.]')-(lamhat[3]*betahat[5]./bp.*(lnmp)^2));
```

mu36 = cdf3.*(gamhat[3,6]*ones(n,1)-mu3.*(alphat[6]*ones(n,1)+lnp*gamhat[6,.]')-(lamhat[3]*betahat[6]./bp.*(lnmp)^2));

e31 = mu31./what3; e32 = mu32./what3; e34 = mu34./what3; e35 = mu35./what3; e36 = mu36./what3;

e31 = meanc(e31); e32 = meanc(e32); e34 = meanc(e34); e35 = meanc(e35); e36 = meanc(e36);

mu41 = cdf4.*(gamhat[4,1]*ones(n,1)-mu4.*(alphat[1]*ones(n,1)+lnp*gamhat[1,.]')-(lamhat[4]*betahat[1]./bp.*(lnmp)^2));
mu42 = cdf4.*(gamhat[4,2]*ones(n,1)-mu4.*(alphat[2]*ones(n,1)+lnp*gamhat[2,.]')-(lamhat[4]*betahat[2]./bp.*(lnmp)^2));
mu43 = cdf4.*(gamhat[4,3]*ones(n,1)-mu4.*(alphat[3]*ones(n,1)+lnp*gamhat[3,.]')-(lamhat[4]*betahat[3]./bp.*(lnmp)^2));
mu45 = cdf4.*(gamhat[4,5]*ones(n,1)-mu4.*(alphat[5]*ones(n,1)+lnp*gamhat[5,.]')-(lamhat[4]*betahat[5]./bp.*(lnmp)^2));
mu46 = cdf4.*(gamhat[4,6]*ones(n,1)-mu4.*(alphat[6]*ones(n,1)+lnp*gamhat[6,.]')-(lamhat[4]*betahat[6]./bp.*(lnmp)^2));

e41 = mu41./what4; e42 = mu42./what4; e43 = mu43./what4; e45 = mu45./what4; e46 = mu46./what4;

e41 = meanc(e41); e42 = meanc(e42); e43 = meanc(e43); e45 = meanc(e45); e46 = meanc(e46);

mu51 = cdf5.*(gamhat[5,1]*ones(n,1)-mu5.*(alphat[1]*ones(n,1)+lnp*gamhat[1,.]')-(lamhat[5]*betahat[1]./bp.*(lnmp)^2));
mu52 = cdf5.*(gamhat[5,2]*ones(n,1)-mu5.*(alphat[2]*ones(n,1)+lnp*gamhat[2,.]')-(lamhat[5]*betahat[2]./bp.*(lnmp)^2));
mu53 = cdf5.*(gamhat[5,3]*ones(n,1)-mu5.*(alphat[3]*ones(n,1)+lnp*gamhat[3,.]')-(lamhat[5]*betahat[3]./bp.*(lnmp)^2));
mu54 = cdf5.*(gamhat[5,4]*ones(n,1)-mu5.*(alphat[4]*ones(n,1)+lnp*gamhat[4,.]')-(lamhat[5]*betahat[4]./bp.*(lnmp)^2));
mu56 = cdf5.*(gamhat[5,6]*ones(n,1)-mu5.*(alphat[6]*ones(n,1)+lnp*gamhat[6,.]')-(lamhat[5]*betahat[6]./bp.*(lnmp)^2));

e51 = mu51./what5; e52 = mu52./what5; e53 = mu53./what5; e54 = mu54./what5; e56 = mu56./what5;

e51 = meanc(e51); e52 = meanc(e52); e53 = meanc(e53); e54 = meanc(e54); e56 = meanc(e56);

mu61 = cdf7.*(gamhat[6,1]*ones(n,1)-mu5.*(alphat[1]*ones(n,1)+lnp*gamhat[1,.]')-(lamhat[6]*betahat[1]./bp.*(lnmp)^2));
mu62 = cdf7.*(gamhat[6,2]*ones(n,1)-mu5.*(alphat[2]*ones(n,1)+lnp*gamhat[2,.]')-(lamhat[6]*betahat[2]./bp.*(lnmp)^2));
mu63 = cdf7.*(gamhat[6,3]*ones(n,1)-mu5.*(alphat[3]*ones(n,1)+lnp*gamhat[3,.]')-(lamhat[6]*betahat[3]./bp.*(lnmp)^2));

```
   mu64 = cdf7.*(gamhat[6,4]*ones(n,1)-
mu5.*(alphat[4]*ones(n,1)+lnp*gamhat[4,.]')-(lamhat[6]*betahat[4]./bp.*(lnmp)^2));
   mu65 = cdf7.*(gamhat[6,5]*ones(n,1)-
mu5.*(alphat[5]*ones(n,1)+lnp*gamhat[5,.]')-(lamhat[6]*betahat[5]./bp.*(lnmp)^2));

   e61 = mu61./what7; e62 = mu62./what7; e63 = mu63./what7; e64 = mu64./what7;
e65 = mu65./what7;

   e61 = meanc(e61); e62 = meanc(e62); e63 = meanc(e63); e64 = meanc(e64); e65 =
meanc(e65);

/* Calculating t-stat   */

CPelast = Crossprice(bhat);
CPelast = reshape(CPelast,6,5);
CPgrad = gradp(&Crossprice,bhat);
CPgrad = CPgrad[.,1:9]~CPgrad[.,11:29]~CPgrad[.,31:60];
CPstd = sqrt(diag(CPgrad*cov*CPgrad'));
index =
seqa(12,1,5)|21|seqa(23,1,4)|31|32|seqa(34,1,3)|seqa(41,1,3)|45|46|seqa(51,1,4)|56|seq
a(61,1,5);

print " Cross-price Elasticity ";
print "     Index       Elast.        Std.         t-value    ";
print index~CPelast~CPstd~CPelast./CPstd;
end;

proc QUADFORM(P,G);
   @This proc has an n by m matrix P and a m by m matrix of the quadratic
   form, G, as input and returns an n by 1 matrix of P[i,.]*G*P[i,.]'
   quadratic form values@
retp(sumc((P*G.*P)'));
endp;

proc SYQUAIDS(theta);
   local alpha, gam, gammat, beta, lamda, phi, w, w1hat, w2hat, w3hat, w4hat, w5hat,
w6hat, w7hat, lnp, lnap, lnbp, bp;
   @ Global variables used in this proc are prices (p), expenditure (Ms), and
probability density value (pdf). @
   @ From the input of the 7*11 coefficients matrix, we can identify alphas, gammas,
betas, lamdas, and phis. @

   /* Specify a vector of alpha coefficients (7 by 1) */
   alpha = theta[1]|theta[11]|theta[21]|theta[31]|theta[41]|theta[51];
   /* Specify a vector of gamma coefficients (49 by 1) */
   gam = theta[2:7]|theta[12:17]|theta[22:27]|theta[32:37]|theta[42:47]|theta[52:57];

   /* Re-specify a matrix of gamma coefficients (7 by 7), for the use in
proc_quadform */
   gammat = gam[1:6]'|gam[7:12]'|gam[13:18]'|gam[19:24]'|gam[25:30]'|gam[31:36]';
```

```
    /* Specify a vector of beta coefficients (7 by 1) */
    beta = theta[8]|theta[18]|theta[28]|theta[38]|theta[48]|theta[58];

    /* Specify a vector of lamda coefficients (7 by 1) */
    lamda = theta[9]|theta[19]|theta[29]|theta[39]|theta[49]|theta[59];

    /* Specify a vector of phi coefficients (7 by 1), for the use with the pdf variable */
    phi = theta[10]|theta[20]|theta[30]|theta[40]|theta[50]|theta[60];

    lnp = ln(p);
    lnap = a0 + lnp*alpha + 0.5*quadform(lnp,gammat);
    lnbp = lnp*beta;
    bp = exp(lnbp);
    lnmp = (ln(Ms) - lnap);
    lnmp2 = (lnmp^2)./bp;
    w1hat = alpha[1]*ones(n,1) + lnp*(gam[1:6]) + lnmp*beta[1] + (lamda[1]*lnmp2);
    w2hat = alpha[2]*ones(n,1) + lnp*(gam[7:12]) + lnmp*beta[2] +
(lamda[2]*lnmp2);
    w3hat = alpha[3]*ones(n,1) + lnp*(gam[13:18]) + lnmp*beta[3] +
(lamda[3]*lnmp2);
    w4hat = alpha[4]*ones(n,1) + lnp*(gam[19:24]) + lnmp*beta[4] +
(lamda[4]*lnmp2);
    w5hat = alpha[5]*ones(n,1) + lnp*(gam[25:30]) + lnmp*beta[5] +
(lamda[5]*lnmp2);
    w6hat = alpha[6]*ones(n,1) + lnp*(gam[31:36]) + lnmp*beta[6] +
(lamda[6]*lnmp2);


    w1hat = cdf[.,1].*w1hat ;
    w2hat = cdf[.,2].*w2hat + phi[2]*pdf[.,2];
    w3hat = cdf[.,3].*w3hat ;
    w4hat = cdf[.,4].*w4hat + phi[4]*pdf[.,4];
    w5hat = cdf[.,5].*w5hat + phi[5]*pdf[.,5];
    w6hat = cdf[.,6].*w6hat + phi[6]*pdf[.,6];

    what = w1hat|w2hat|w3hat|w4hat|w5hat|w6hat;

    retp(what);
endp;

/* PROCEDURE TO CALCULATE INCOME ELASTICITIES  */
proc Inc_elast(bhat);
    local e;
    /* Data prepared for calculating elasticities */
    alphat = bhat[1]|bhat[11]|bhat[21]|bhat[31]|bhat[41]|bhat[51];
    gamhat = bhat[2:7]'|bhat[12:17]'|bhat[22:27]'|bhat[32:37]'|bhat[42:47]'|bhat[52:57]';
    lnap = a0*ones(n,1) + lnp*alphat + 0.5*quadform(lnp,gamhat);
    betahat = bhat[8]|bhat[18]|bhat[28]|bhat[38]|bhat[48]|bhat[58];
    lnbp = lnp*betahat;
```

```
    bp = exp(lnbp);
    lnmp = (ln(Ms) - lnap);
    lamhat = bhat[9]|bhat[19]|bhat[29]|bhat[39]|bhat[49]|bhat[59];

    /* THE CALCULATION OF INCOME ELASTICITIES */

    mu1 = cdf1.*(betahat[1]*ones(n,1)+2*(lamhat[1]./bp).*lnmp);
    mu26 = cdf26.*(betahat[2]*ones(n,1)+2*(lamhat[2]./bp).*lnmp);
    mu3 = cdf3.*(betahat[3]*ones(n,1)+2*(lamhat[3]./bp).*lnmp);
    mu4 = cdf4.*(betahat[4]*ones(n,1)+2*(lamhat[4]./bp).*lnmp);
    mu5 = cdf5.*(betahat[5]*ones(n,1)+2*(lamhat[5]./bp).*lnmp);
    mu7 = cdf7.*(betahat[6]*ones(n,1)+2*(lamhat[6]./bp).*lnmp);

    e1 = mu1./what1 + 1; e2 = mu26./what26 + 1; e3 = mu3./what3 + 1;
    e4 = mu4./what4 + 1; e5 = mu5./what5 + 1; e6 = mu7./what7 + 1;

    e1 = meanc(e1); e2 = meanc(e2); e3 = meanc(e3);
    e4 = meanc(e4); e5 = meanc(e5); e6 = meanc(e6);

    e = e1|e2|e3|e4|e5|e6;
    retp(e);
endp;

proc ownprice(bhat);

    /* Data prepared for calculating elasticities */
    alphat = bhat[1]|bhat[11]|bhat[21]|bhat[31]|bhat[41]|bhat[51];
    gamhat = bhat[2:7]'|bhat[12:17]'|bhat[22:27]'|bhat[32:37]'|bhat[42:47]'|bhat[52:57]';
    lnap = a0*ones(n,1) + lnp*alphat + 0.5*quadform(lnp,gamhat);
    betahat = bhat[8]|bhat[18]|bhat[28]|bhat[38]|bhat[48]|bhat[58];
    lnbp = lnp*betahat;
    bp = exp(lnbp);
    lnmp = (ln(Ms) - lnap);
    lamhat = bhat[9]|bhat[19]|bhat[29]|bhat[39]|bhat[49]|bhat[59];
    diaggam = diag(gamhat);

    /* THE CALCULATION OF OWN-PRICE ELATICITIES   */
    mu11 = cdf1.*(diaggam[1]*ones(n,1)-
mu1.*(alphat[1]*ones(n,1)+lnp*gamhat[1,.]')-
((lamhat[1]*betahat[1]./bp).*(lnmp)^2));
    mu22 = cdf26.*(diaggam[2]*ones(n,1)-
mu26.*(alphat[2]*ones(n,1)+lnp*gamhat[2,.]')-
((lamhat[2]*betahat[2]./bp).*(lnmp)^2));
    mu33 = cdf3.*(diaggam[3]*ones(n,1)-
mu3.*(alphat[3]*ones(n,1)+lnp*gamhat[3,.]')-
((lamhat[3]*betahat[3]./bp).*(lnmp)^2));
    mu44 = cdf4.*(diaggam[4]*ones(n,1)-
mu4.*(alphat[4]*ones(n,1)+lnp*gamhat[4,.]')-
((lamhat[4]*betahat[4]./bp).*(lnmp)^2));
```

```
    mu55 = cdf5.*(diaggam[5]*ones(n,1)-
mu5.*(alphat[5]*ones(n,1)+lnp*gamhat[5,.]')-
((lamhat[5]*betahat[5]./bp).*(lnmp)^2));
    mu66 = cdf7.*(diaggam[6]*ones(n,1)-
mu7.*(alphat[6]*ones(n,1)+lnp*gamhat[6,.]')-
((lamhat[6]*betahat[6]./bp).*(lnmp)^2));

    e11 = (mu11./what1)-1; e22 = (mu22./what26)-1; e33 = (mu33./what3)-1;
    e44 = (mu44./what4)-1; e55 = (mu55./what5)-1; e66 = (mu66./what7)-1;

    obs = seqa(1,1,n);
    dw1 = w1-what1;
    dw2 = w26-what26;
    dw3 = w3-what3;
    dw4 = w4-what4;
    dw5 = w5-what5;
    dw6 = w7-what7;

    e11 = meanc(e11); e22 = meanc(e22); e33 = meanc(e33); e44 = meanc(e44); e55 =
meanc(e55); e66 = meanc(e66);

    retp(e11|e22|e33|e44|e55|e66);
endp;

proc Crossprice(bhat);

    /* Data prepared for calculating elasticities */
    alphat = bhat[1]|bhat[11]|bhat[21]|bhat[31]|bhat[41]|bhat[51];
    gamhat = bhat[2:7]'|bhat[12:17]'|bhat[22:27]'|bhat[32:37]'|bhat[42:47]'|bhat[52:57]';
    lnap = a0*ones(n,1) + lnp*alphat + 0.5*quadform(lnp,gamhat);
    betahat = bhat[8]|bhat[18]|bhat[28]|bhat[38]|bhat[48]|bhat[58];
    lnbp = lnp*betahat;
    bp = exp(lnbp);
    lnmp = (ln(Ms) - lnap);
    lamhat = bhat[9]|bhat[19]|bhat[29]|bhat[39]|bhat[49]|bhat[59];
    diaggam = diag(gamhat);

    /* THE CALCULATION OF CROSS-PRICE ELASTICITIES */
    mu12 = cdf1.*(gamhat[1,2]*ones(n,1)-
mu1.*(alphat[2]*ones(n,1)+lnp*gamhat[2,.]')-(lamhat[1]*betahat[2]./bp.*(lnmp)^2));
    mu13 = cdf1.*(gamhat[1,3]*ones(n,1)-
mu1.*(alphat[3]*ones(n,1)+lnp*gamhat[3,.]')-(lamhat[1]*betahat[3]./bp.*(lnmp)^2));
    mu14 = cdf1.*(gamhat[1,4]*ones(n,1)-
mu1.*(alphat[4]*ones(n,1)+lnp*gamhat[4,.]')-(lamhat[1]*betahat[4]./bp.*(lnmp)^2));
    mu15 = cdf1.*(gamhat[1,5]*ones(n,1)-
mu1.*(alphat[5]*ones(n,1)+lnp*gamhat[5,.]')-(lamhat[1]*betahat[5]./bp.*(lnmp)^2));
    mu16 = cdf1.*(gamhat[1,6]*ones(n,1)-
mu1.*(alphat[6]*ones(n,1)+lnp*gamhat[6,.]')-(lamhat[1]*betahat[6]./bp.*(lnmp)^2));
```

```
   e12 = mu12./what1; e13 = mu13./what1; e14 = mu14./what1; e15 = mu15./what1;
e16 = mu16./what1;

   e12 = meanc(e12); e13 = meanc(e13); e14 = meanc(e14); e15 = meanc(e15); e16 =
meanc(e16);

   mu21 = cdf26.*(gamhat[2,1]*ones(n,1)-
mu26.*(alphat[1]*ones(n,1)+lnp*gamhat[1,.]')-
(lamhat[2]*betahat[1]./bp.*(lnmp)^2));
   mu23 = cdf26.*(gamhat[2,3]*ones(n,1)-
mu26.*(alphat[3]*ones(n,1)+lnp*gamhat[3,.]')-
(lamhat[2]*betahat[3]./bp.*(lnmp)^2));
   mu24 = cdf26.*(gamhat[2,4]*ones(n,1)-
mu26.*(alphat[4]*ones(n,1)+lnp*gamhat[4,.]')-
(lamhat[2]*betahat[4]./bp.*(lnmp)^2));
   mu25 = cdf26.*(gamhat[2,5]*ones(n,1)-
mu26.*(alphat[5]*ones(n,1)+lnp*gamhat[5,.]')-
(lamhat[2]*betahat[5]./bp.*(lnmp)^2));
   mu26 = cdf26.*(gamhat[2,6]*ones(n,1)-
mu26.*(alphat[6]*ones(n,1)+lnp*gamhat[6,.]')-
(lamhat[2]*betahat[6]./bp.*(lnmp)^2));

   e21 = mu21./what26; e23 = mu23./what26; e24 = mu24./what26; e25 =
mu25./what26; e26 = mu26./what26;

   e21 = meanc(e21); e23 = meanc(e23); e24 = meanc(e24); e25 = meanc(e25); e26 =
meanc(e26);

   mu31 = cdf3.*(gamhat[3,1]*ones(n,1)-
mu3.*(alphat[1]*ones(n,1)+lnp*gamhat[1,.]')-(lamhat[3]*betahat[1]./bp.*(lnmp)^2));
   mu32 = cdf3.*(gamhat[3,2]*ones(n,1)-
mu3.*(alphat[2]*ones(n,1)+lnp*gamhat[2,.]')-(lamhat[3]*betahat[2]./bp.*(lnmp)^2));
   mu34 = cdf3.*(gamhat[3,4]*ones(n,1)-
mu3.*(alphat[4]*ones(n,1)+lnp*gamhat[4,.]')-(lamhat[3]*betahat[4]./bp.*(lnmp)^2));
   mu35 = cdf3.*(gamhat[3,5]*ones(n,1)-
mu3.*(alphat[5]*ones(n,1)+lnp*gamhat[5,.]')-(lamhat[3]*betahat[5]./bp.*(lnmp)^2));
   mu36 = cdf3.*(gamhat[3,6]*ones(n,1)-
mu3.*(alphat[6]*ones(n,1)+lnp*gamhat[6,.]')-(lamhat[3]*betahat[6]./bp.*(lnmp)^2));

   e31 = mu31./what3; e32 = mu32./what3; e34 = mu34./what3; e35 = mu35./what3;
e36 = mu36./what3;

   e31 = meanc(e31); e32 = meanc(e32); e34 = meanc(e34); e35 = meanc(e35); e36 =
meanc(e36);

   mu41 = cdf4.*(gamhat[4,1]*ones(n,1)-
mu4.*(alphat[1]*ones(n,1)+lnp*gamhat[1,.]')-(lamhat[4]*betahat[1]./bp.*(lnmp)^2));
   mu42 = cdf4.*(gamhat[4,2]*ones(n,1)-
mu4.*(alphat[2]*ones(n,1)+lnp*gamhat[2,.]')-(lamhat[4]*betahat[2]./bp.*(lnmp)^2));
```

```
    mu43 = cdf4.*(gamhat[4,3]*ones(n,1)-
mu4.*(alphat[3]*ones(n,1)+lnp*gamhat[3,.]')-(lamhat[4]*betahat[3]./bp.*(lnmp)^2));
    mu45 = cdf4.*(gamhat[4,5]*ones(n,1)-
mu4.*(alphat[5]*ones(n,1)+lnp*gamhat[5,.]')-(lamhat[4]*betahat[5]./bp.*(lnmp)^2));
    mu46 = cdf4.*(gamhat[4,6]*ones(n,1)-
mu4.*(alphat[6]*ones(n,1)+lnp*gamhat[6,.]')-(lamhat[4]*betahat[6]./bp.*(lnmp)^2));

    e41 = mu41./what4; e42 = mu42./what4; e43 = mu43./what4; e45 = mu45./what4;
e46 = mu46./what4;

    e41 = meanc(e41); e42 = meanc(e42); e43 = meanc(e43); e45 = meanc(e45); e46 =
meanc(e46);

    mu51 = cdf5.*(gamhat[5,1]*ones(n,1)-
mu5.*(alphat[1]*ones(n,1)+lnp*gamhat[1,.]')-(lamhat[5]*betahat[1]./bp.*(lnmp)^2));
    mu52 = cdf5.*(gamhat[5,2]*ones(n,1)-
mu5.*(alphat[2]*ones(n,1)+lnp*gamhat[2,.]')-(lamhat[5]*betahat[2]./bp.*(lnmp)^2));
    mu53 = cdf5.*(gamhat[5,3]*ones(n,1)-
mu5.*(alphat[3]*ones(n,1)+lnp*gamhat[3,.]')-(lamhat[5]*betahat[3]./bp.*(lnmp)^2));
    mu54 = cdf5.*(gamhat[5,4]*ones(n,1)-
mu5.*(alphat[4]*ones(n,1)+lnp*gamhat[4,.]')-(lamhat[5]*betahat[4]./bp.*(lnmp)^2));
    mu56 = cdf5.*(gamhat[5,6]*ones(n,1)-
mu5.*(alphat[6]*ones(n,1)+lnp*gamhat[6,.]')-(lamhat[5]*betahat[6]./bp.*(lnmp)^2));

    e51 = mu51./what5; e52 = mu52./what5; e53 = mu53./what5; e54 = mu54./what5;
e56 = mu56./what5;

    e51 = meanc(e51); e52 = meanc(e52); e53 = meanc(e53); e54 = meanc(e54); e56 =
meanc(e56);

    mu61 = cdf7.*(gamhat[6,1]*ones(n,1)-
mu5.*(alphat[1]*ones(n,1)+lnp*gamhat[1,.]')-(lamhat[6]*betahat[1]./bp.*(lnmp)^2));
    mu62 = cdf7.*(gamhat[6,2]*ones(n,1)-
mu5.*(alphat[2]*ones(n,1)+lnp*gamhat[2,.]')-(lamhat[6]*betahat[2]./bp.*(lnmp)^2));
    mu63 = cdf7.*(gamhat[6,3]*ones(n,1)-
mu5.*(alphat[3]*ones(n,1)+lnp*gamhat[3,.]')-(lamhat[6]*betahat[3]./bp.*(lnmp)^2));
    mu64 = cdf7.*(gamhat[6,4]*ones(n,1)-
mu5.*(alphat[4]*ones(n,1)+lnp*gamhat[4,.]')-(lamhat[6]*betahat[4]./bp.*(lnmp)^2));
    mu65 = cdf7.*(gamhat[6,5]*ones(n,1)-
mu5.*(alphat[5]*ones(n,1)+lnp*gamhat[5,.]')-(lamhat[6]*betahat[5]./bp.*(lnmp)^2));

    e61 = mu61./what7; e62 = mu62./what7; e63 = mu63./what7; e64 = mu64./what7;
e65 = mu65./what7;

    e61 = meanc(e61); e62 = meanc(e62); e63 = meanc(e63); e64 = meanc(e64); e65 =
meanc(e65);

    CPelast = e12|e13|e14|e15|e16;
    CPelast = CPelast|(e21|e23|e24|e25|e26);
    CPelast = CPelast|(e31|e32|e34|e35|e36);
```

```
    CPelast = CPelast|(e41|e42|e43|e45|e46);
    CPelast = CPelast|(e51|e52|e53|e54|e56);
    CPelast = CPelast|(e61|e62|e63|e64|e65);

     retp(CPelast);

endp;
```