

**THE EFFECT OF SOCIOECONOMIC STATUS ON
HEALTH SERVICE UTILIZATION IN KENYA: ECONOMETRIC ANALYSIS**

MAUREEN WAVINYA MBALUKA

X53/87091/2016

A RESEARCH PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENT FOR THE AWARD OF THE DEGREE OF MASTER OF SCIENCE IN
HEALTH ECONOMICS AND POLICY OF THE UNIVERSITY OF NAIROBI

NOVEMBER 2020

DECLARATION

This research project report is my original work and has not been presented for any examination in any other institution.

Signed: _____ Date: _____

Maureen Wavinya Mbaluka

APPROVAL

This project has been undertaken and submitted for examination with my approval as University Supervisor.

Signed _____ Date _____

Professor Germano Mwabu

DEDICATION

I dedicate this work to my parents, Mr David Mbaluka, and Mrs Loise D. Mbaluka for their unfailing love, and to my siblings, Caroline, Lilian, and Mercy, for their encouragement and moral support during the entire period of study. I wish them blessings of the Almighty God.

ACKNOWLEDGEMENTS

I am indebted to God's mercies and His guidance throughout the development of this project report. My sincere and special gratitude goes to my supervisor Prof. Germano Mwabu, for his dedicated supervision, guidance, patience, encouragement and whose insights and knowledge into the subject matter steered me through the development of this research project. I am grateful to my lecturers and the entire staff in the University of Nairobi, for the knowledge they imparted in me as it has been instrumental in the development of my research. My appreciation goes to my parents for support and guidance throughout my academic journey. I wish to thank my sisters and everyone else who helped me for their support.

ABSTRACT

Kenya has been in the forefront in designing policies to reduce disparities in access to health care and in health outcomes. The Government has been working to improve the health sector performance in with Vision 2030. In particular, the Government of Kenya launched the universal health coverage (UHC) agenda in 2018 as one of the big four development initiatives. The initiatives outside the health sector, which have also been piloted, include revitalizing the manufacturing sector and improving housing conditions. This was in a bid to achieve effective service coverage and financial protection in events of sickness. Despite the progress made towards achieving equity in health and in access to health care, socioeconomic disparities in health care utilization continue to persist. On this backdrop this study uses econometric methods to examine the effects of socioeconomic factors on health care utilization in Kenya shortly before the launch of the UHC agenda. This study employed LPM, logit and probit regressions on data samples from the Kenya Integrated Household Budget Survey (KIHBS), collected by the Kenya National Bureau of Statistics in 2015/16. This survey data has comprehensive information on all the variables needed for the analysis of disparities in access to health care in Kenya. The 2015/16 KIHBS consisted of 5,360 clusters, split into four equal sub-samples. The sampling frame is stratified into urban and rural areas within each of 47 counties resulting in 92 sampling strata with Nairobi city and Mombasa counties being wholly urban. The sample size was determined independently for each county, resulting in a national sample of 24,000 households. The econometric analysis shows that an individual's age, income per capita, household size, per capita income, gender, education level, employment status, area of residence and insurance are all significant determinants of health service utilization. The descriptive statistics reveal that insurance coverage is low (16.8%) and that women with formal sector employment comprise only 11.0% of total wage employment. As expected, per capita income, a major determinant of health service utilization in this data set, is highly skewed, and at the survey time (2015/16), it averaged around Ksh 4,600 per month. Separately, economic variables (income, employment, health insurance) and the social background variables (gender, marital status, and age) have large impacts on health service utilization but the impacts of their interactions are relatively minor. In particular, only the interactions of gender with wage employment and with insurance have utilization effects that are different from zero. Insurance coverage alone is associated with a large increase in health service utilization. Utilization of health services in Kenya is still low; hence the government needs to come up with measures to reduce health service utilization disparities related to gender, literacy levels, different income groups, enrolment into health insurance, health awareness levels, and also should design other kinds of assistance for the most affected groups, like men, and persons associated with disadvantageous marriage arrangements, such as the widowed, divorced, and the separated.

TABLE OF CONTENTS

DECLARATION	II
DEDICATION	III
ACKNOWLEDGEMENTS	IV
ABSTRACT	V
LIST OF TABLES	VIII
ABBREVIATIONS AND ACRONYMS	IX
CHAPTER ONE	1
INTRODUCTION	1
1.1 Background Information	1
1.2 Statement of the Problem	3
1.3 Research Questions	3
1.4 Research Aims.....	4
1.5 Justification of the Study.....	4
CHAPTER TWO	5
LITERATURE REVIEW	5
2.1 Theoretical Literature	5
2.1.2 Mechanic’s general theory	5
2.3 Overview of the Literature	8
CHAPTER THREE	9
RESEARCH METHODOLOGY	9
3.1 Introduction	9
3.2 Theoretical model.....	9
3.3 Economic models and specification issues	10
3.4 Data source.....	14
CHAPTER FOUR	15
DATA ANALYSIS, INTERPRETATION AND DISCUSSIONS OF RESULTS	15
4.1 Introduction	15
4.2 Descriptive Statistics	16
4.3 Diagnostic pre-estimation and post-estimation tests.....	17
4.3.1 Heteroskedasticity	17
4.3.2 Multicollinearity test.....	17
4.3.3 Normality tests: Shapiro-Wilk W test for normal data.....	18
4.3.4 Correlation Analysis	18
4.4 Econometric Analysis	19
CHAPTER FIVE	26
SUMMARY, CONCLUSIONS AND RECOMMENDATIONS	26
5.1 Introduction	26
5.2 Summary of the findings	26
5.3 Conclusion.....	27
5.4 Policy Recommendation	27

5.5 Recommendations for further research	28
REFERENCES	29
APPENDICES	31
Appendix 1: Heteroskedasticity Test	31
Appendix 2: Multicollinearity Test	32
Appendix 3: Normality Tests: Shapiro-Wilk W Test for Normal Data	33
Appendix 4: Logit and Probit Model Estimates for Effects of Social Economic Status on Health Care Utilization.....	34
Appendix 5: Correlation Matrix	35

LIST OF TABLES

Table 3.1: Description Of Covariates.....	13
Table 4.2: Summary Statistics Of The Study Variables.....	16
Table 4.3: LPM, Logit And Probit Model Estimates For Health Service utilization.....	19
Table 4.4: A LPM on effects of social-economic status on health service utilization.....	22

ABBREVIATIONS AND ACRONYMNS

MOH	Ministry of Health
KHHEUS	Kenya Household Health Expenditure and Utilization Survey
KNBS	Kenya National Bureau of Statistics
NHIF	National Health Insurance Fund
GOK	Government of Kenya
KHPF	Kenya Health Policy Frameworks
TB	Tuberculosis WHO World Health Organization
UHC	Universal Health Coverage
NIH	National Institute of Health
PHC	Primary Health Care
SDG	Sustainable Development Goal
KIBHS	Kenya Integrated Household Budget Survey
LPM	Linear Probability Model
SHI	Social Health Insurance

CHAPTER ONE

INTRODUCTION

1.1 Background Information

Utilization of health care is the rate at which health services are used by the population over a given time period. Being ill is unavoidable and this is part of life. There is a need to examine the social and economic factors that affect health service utilization in order to ensure that there are guidelines for promoting efficiency in health care access which will lead to reduction in health inequalities. The issue of health service utilization has been of a concern for policy makers as well as researchers for a long time.

Access to health services not only promotes the general quality but can achieve the best possible results for status in a population. Studies on socio-economic inequalities in health service utilization have increased remarkably in developed and developing countries. The findings of the research have shown that there is significant evidence on inequalities in health service utilization. According to Kutzin (2013), for countries aiming to achieve universal health coverage, there is a need to ensure equitable accessibility of health service utilization and ensure that individuals seeking health services are protected from out-of-pocket expenses.

There has been some considerable evidence showing effects of socioeconomic variables on health care utilization both in developed and developing countries. A study by Celik and Hotchkiss (2000) documented major contributions of socio-economic factors to utilization of maternal healthcare services in Turkey. In Turkey at the time, the main drivers of maternal health service utilization included medical insurance cover, education level, household wealth, residence region and ethnicity. The reduction in infant and child mortality rates were also associated with these factors.

In one of the early studies on this issue in Kenya, Muriithi (2013) demonstrated that proximity to health facilities was one of the determinants of outpatients' utilization of health facilities and that proximity had an additional effect of reducing self-treatments. The key determinants of service usage in that study included gender (with female dummy increasing the probability of service uptake), and with user fees decreasing usage.

Additionally, the main predictors of health service usage included age, gender, income, education, medical insurance, type of care, availability of health personnel, sickness symptoms, illness type and size of the household. Other Kenyan studies report similar findings, e.g., Ellis and Mwabu (2004).

Mills et al (2012) show that patterns of health service demand in South Africa, Ghana and Tanzania are influenced by socioeconomic factors, and that the highest burden of illness is associated with the lowest socioeconomic groups. This can be attributed to the fact that the higher socioeconomic groups have a benefit of using preventive health services like, having regular check-ups, taking screening tests for potential health disorders, enrolling into insurance plans and have the ability to pay for health services compared to lower social economic groups. In order to achieve Universal Health Coverage, the health systems have come up with measures for promoting access to preventive, curative, promotive and rehabilitative health services at affordable costs (see, Chuma& Okungu (2011).

Remarkable improvement in Kenya regarding healthcare has been noted in the recent years. Health sector has been priority for the government of Kenya as it works towards Vision 2030 health goals, with progress in this direction having accelerated after 1994. The progress made so far was facilitated by the national health sector plans since the 1990s, particularly; the health annual plans coupled with preparation of health policy frameworks, as long term health development initiatives. For effective service coverage and financial protection, the government of Kenyan in 2018 launched Universal Health Coverage (UHC)and piloted it in four counties as one of the big four agenda of the Kenya's plan for socio-economic development.

While the Ministry of Health operates on limited resources on a fixed budget there have been increased inefficiencies in utilization of these resources and this is a big challenge to the health sector and hence slowing down the embracing of technology to ensure there is improvement in the health outcomes and a reduction in mortality rates (MoH, 2014). The current policy, KHPF (Kenya Health Policy Framework) 2014-2030, pays attention to promotion of equity in health, collaborating with various other key stakeholder groups, to ensure there is active participation and efficiency in usage of health services, social

accountability and there is a focus on the general public health facilities to ensure that basic services are accessible by the disadvantaged groups.

Even though there is progress made so far towards achieving equity in quality health access to health services, the uptake of health services is still low.

1.2 Statement of the Problem

The world has been committed towards achieving universal health coverage (UHC) and Sustainable Development Goal (SDG) 3. This translates to equal access to health services irrespective of social and economic status. As the Kenyan Government commits itself towards improving healthcare, access to health services has been a policy concern. The cost of health services rendered in the public hospitals ranging from dispensaries, health centres to hospitals in Kenya are highly subsidized with patients paying very little user fees or nothing.

These initiatives notwithstanding, healthcare utilization in Kenya remains a challenge, with a large proportion of Kenyans (13 per cent) not seeking health care services from qualified health practitioner in case of ill health (MOH, 2014). Productivity of an individual is affected by the state of illness. The effect of social economic status causes existence of inequalities in health access and in the health status. For better health outcomes people need to have access and utilize health care so as to avoid poor health persisting into the future. Evidence on effects of social and economic factors on usage of health can be used to identify sustainable policies to reduce inequalities in the utilization of health care and therefore reduce disparities in health outcomes in Kenya.

Evidence on how interactions between economic and social factors affect health service utilization in Kenya is currently lacking. This knowledge gap needs to be filled to facilitate implementation of UHC and improvements in people's health.

1.3 Research Questions

- i. How important are social and economic variables alone in influencing health service utilization?
- ii. How does interaction between social and economic variables affect health service utilization?

1.4 Research Aims

1.4.1 The broad Aim

To investigate role of socio economic status in health care utilization.

1.4.2 Particular Aims

The particular purposes of the study are the following:

- i. To assess roles of social and economic variables as determinants of access to and utilization of health services.
- ii. To examine effects of interactions between economic and social variables on the usage of health services in Kenya.
- iii. To suggest on policy recommendations in accordance to the conclusions made on the research results.

1.5 Justification of the Study

Vision 2030 provides high standards of highest affordable and equitable health services as also envisioned by the Government towards its citizens. From earlier demand studies in Kenya, the main determinants of health care usage have emerged to be the following: quality of care, user fees, travel time, distance to facilities and demographic factors. My study looks at these issues in detail using complementary methods of econometric analysis. Understanding the effects of socio-economic status on healthcare utilization may provide useful information to policy makers in the preparation of actionable plan. In order to improve healthcare policies and implementations, providing this information is of paramount importance to key stakeholders in the health sector.

The study further will provide information guiding the policy makers and Ministry of health on how to effectively allocate resources to different regions. This will ensure there is an equitable preventive, curative and rehabilitative services in line with UHC principles. The community will be informed of the effect on socioeconomic status in usage of medical and related services.

CHAPTER TWO

LITERATURE REVIEW

2.1 Theoretical Literature

In this section, three theories of health care utilization are outlined, including Parsons' sick role theory, Mechanic's general theory of help seeking, and Suchman's stages of illness and medical care seeking.

2.1.1 Parsons' sick role theory

Talcott Parsons (1951), described how an individual once diagnosed by a qualified health personnel to have a medical condition, she or he is unable to do activities of daily living compared to a healthy person or persons without such diagnosis. There is therefore some degree of deviation from behaviour that a society can allow from a person who is well and healthy.

Being sick according to Parsons is a form of deviance from normal social roles or states of functioning. Though in an event of illness, an individual is unable to carry out his or her activities of daily living, such as going to work, during such situations, labour productivity is low. This form of deviance from normal social states is approved and sanctioned by the authority figures like health practitioners or community leaders.

Parsons further describes how an ill person has the rights not to be blamed for her or his illness and has a right to be given some freedom in regards to normal obligations by other community members. On the other hand a patient has the responsibility to ensure that she or he prioritizes health needs and seeks medical attention for his or her conditions promptly.

2.1.2 Mechanic's general theory

People need to monitor how their bodies perceive and interpret sickness symptoms (Mechanic, 1995) and thereafter make informed decisions to seek health services.

According to (Mechanic, 1978), the illness behaviour as explained by this theory includes such aspects as bodily symptoms, individual nature, how a person perceives the illness, interprets the illness and the residual category of illness which includes the financial and psychological costs that have to be borne by the person or his or her family.

If the symptoms are persistent and less tolerant it affects the illness behaviour which enables an individual to interpret this symptoms based on the individual needs assessment.

2.1.3 Suchman's (1965) theory of disease and health care demand behaviour

According to this theory, when an illness occurs, a person must decide whether to seek medical treatment or not. The factors at issue in such a consideration: symptoms experienced by an individual such as physical changes, the acceptance or denial of a symptom or illness, the progression of the bodily symptoms as time goes on; the extent to which a person takes the sick role seriously, the severity of illness, and treatment choices available. The final decision as to where treatment is sought depends on extent of acceptance of an illness and the ability to bear treatment expenses. Upon seeking care, recovery and rehabilitation typically would involve interventions outside the health system (Rebhan, 2015).

2.1.4 Consumers' demand for healthcare

Grossman (1972) explained how health is treated as a stock that consumers desire because the health stock has productivity and utility benefits. He noted that health is stock which degrades over time if there is absence of maintenance made to health. Further the model describes that health is a consumption good yielding a direct utility and an investment good yielding higher labour productivity. From the model the predictions that can be made in a bid to explain demand for health services are that: (a) as people grow older their stock of health declines and their consumption of healthcare care increases to offset the decline; (b) an increase in individual's income tends to place increased value on healthy days; hence increasing their consumption of health care; and (c) with education attainment, the demand for health care declines as educated people are more efficient in producing health.

2.1.5 Behavioural model of health Services Utilization

The models of Andersen & Newman's (1973) explain how individual's usage of health services is dependent on factors such as: factors that predispose service usage, such as demographic and social characteristics; enabling factors, the most important of which are incomes, insurance cover, and service availability. Another class of factors includes individual's perception on nature and severity of illness and the kinds of illness diagnoses that are available to the patient.

2.2 Empirical Literature

Numerous studies have been undertaken to examine the effect of socio-economic status on healthcare utilization, one of which is Celik and Hotchkiss (2000). The study looks at impacts of socio-economic factors on maternal healthcare utilization in Turkey. The findings of the study revealed that age, household wealth, ethnicity, educational level, health insurance cover and geographic region are key factors in health service utilisation decisions. The same factors were found to be associated rates of infant and child mortality. In yet another study, Veugelers and Yip (2002) examined whether lower socioeconomic groups use more health services, given their poorer health status. The findings showed that people at lower socioeconomic status utilized health services less whereas the use of specialist services was much higher in the highest income groups. Onwujekwe and Uzochukwu (2005) studied how the social, economic and geographical factors in South Nigeria affected health services utilization. The study concluded that, lower socio-economic groups and rural dwellers experience large inequalities in health service utilization relative to their urban counterparts.

Kimani et al. (2016) carried out an econometric study on health care utilization in Kenya. The findings showed that out-of-pocket expenses, distance to facilities, waiting time at clinics, income, household size, chronic illness, the working status of the household head and area of residence are all major factors affecting usage of health services in Kenya.

Muriithi (2013) studies showed a strong negative association between distance to a nearest health facility and the probability of service uptake. The roles of gender, user fees, quality of service were highlighted as key determinants of demand for health services in Kenya.

Ellis & Mwabu (2004) in their study in Western Kenya found that the standard of health facility, payment of services rendered, the household's resource base, costs to be paid on transportation, the time expected to be spent at the facility to obtain treatment. All these factors determine in various ways, the choice made by a person as to where to seek health services during a period of the so called deviance from the social norm.

A study done by Oladipo (2014) regarding usage of health services in both urban and rural areas found that the most important determinants of health services utilization include, age, gender, income, education level, health insurance cover, quality of care, illness stage, household size and social beliefs.

2.3 Overview of the Literature

There is strong evidence in the literature on the association between socioeconomic factors and health service utilization. Mills et al (2012) confirm this association in their care utilization studies in Tanzania, Ghana and South Africa. The lower socioeconomic groups were found to have the highest burden of illness and to be disadvantaged compared to the higher socioeconomic groups in terms of service utilization. This can be attributed to the fact that the higher socioeconomic groups have a benefit of using preventive health services, like having regular check-ups, taking screening tests for potential health disorders, enrolling into insurance covers, and also the ability to pay for health care compared to lower social economic households and individuals. In order to achieve universal health coverage (UHC), the health systems have come up with measures for promoting access to preventive and curative services, promotive and rehabilitative health services at an affordable cost (see e.g., Chuma & Okungu, 2011).

The studies done in Kenya show that quality of care, user fees, travel time, distance to facilities and demographic factors affect utilization of healthcare but the studies do not go into details about effects of socioeconomic variables, and their interactions on utilization of health care; my study focuses on this issue. The study will enlighten policy making on how socioeconomic status and how independently, the interactions between social and economic variables in Kenya influence health services usage in the country for proper planning and decision making on how to improve health conditions of the population.

CHAPTER THREE RESEARCH METHODOLOGY

3.1 Introduction

In this chapter, I will explain my methodology and the data I used to estimate various health service utilization models. The theoretical and econometric models are outlined, the study variables are defined.

3.2 Theoretical model

This theory is based on utility maximization theory whereby in existence of competing alternatives, a consumer's decision will be guided by the choice which aims at maximizing the utility. In this study the sick person is faced with having to choose utilization of health care services (or not), subject to a fixed budget. The benefit that an individual expects from using a health service can be given by the following relationship:

$$(i) \quad U = g(X1, X2).$$

Where:

U is the expected net benefit that the individual anticipates to get from after using health services.

$X1$ comprises goods and services that confer direct benefit to an individual but have no direct link to health;

$X2$ is a good that directly affects a person's health, e.g., medical treatment.

If the sick individual is faced with a budget constraint and has well defined health production, s(he) can be considered as trying to maximize benefits from health service usage.

The budget constraint for such a person is as follows:

$$(ii) \quad Y = X1P_{g1} + X2P_{g2}$$

Where;

Y is a person's resource base, X1 is a general good, X2 is a good that when used restores health, and P_{g1} and P_{g2} represent the respective prices of the goods.

The solution to utility maximization problem is a health care demand function of the form:

$$(iii) \quad I_i = g(P_{g1}, P_{g2}, Y, S)$$

Where;

I_i is the demand for health care facing person i.

Where; P_{g1} and P_{g2} are prices paid for the two goods and S comprises social-economic variables and the enabling factors that shift service utilization down or up.

Health care service demand by an individual is affected prices and by his or her socioeconomic status.

3.3 Economic models and specification issues

This study employed LPM, logit and probit regressions to analyze determinants of health service utilization. Since the dependent variable is dichotomous the study majored on the decision of whether or not individual utilizing the healthcare services hence taking two values, 0 if no and 1 if yes. The main purpose of the study was to analyze the social and economic factors affecting the likelihood of utilizing healthcare services in Kenya.

The assumption was made in a linear association between the probability of using health services, Y, and covariates (x_i) that determine the level of Y; as shown below:

$$Y = x_i\beta + \varepsilon$$

Where;

Y is the dependent variable of a probability of a sick individual visiting a health professional. The estimation procedures take the average of all the variables (Y and X) for all individuals. The x_i s are covariates determine the decision to use or not to use health services, such as age, education, among others.

Hence the expected benefit (utility) of a sick person from utilizing health care services as influenced by independent variables may be expressed as follows:

$$U_h = \beta_0 + \beta_1 X_1 + \dots + \beta_n X_n + \varepsilon$$

The probability of utilization of health care services will be presented as a function of the same explanatory variables which are: education level, age, health insurance, marital status, income, family size, age, gender, employment.

Thus the logit model can be shown as:

$P = \exp(U) / [1 + \exp(U)]$ (Gareth et al. 2014). This can also simply be rewritten as

$$P = 1 / [1 + \exp(-U)]$$

Where;

P = Probability of visiting a health facility

U = perceived benefit (utility) from health service utilization, which in linear form can be expressed:

$$U = \beta_0 + \beta_1 \text{education level} + \beta_2 \text{age} + \beta_3 \text{gender} + \beta_4 \text{employment status} + \beta_5 \text{distance} + \beta_6 \text{cost of health} + \beta_7 \text{disease} + \beta_8 \text{quality of care} + \beta_9 \text{income} + \beta_{10} \text{marital status} + \beta_{11} \text{health insurance} + \beta_{12} \text{family size} + \text{error term } (\varepsilon)$$

Since utility, U , from service utilization, upon visiting a health facility is affected by economic factors (E) and social factors (S) and the interaction between (E and S), the U -equation can be shortened and re-written as:

$$U = \beta_0 + \beta_1 E + \beta_2 S + \beta_3 (E * S) + \varepsilon$$

The parameters (β_s) of U (logit/probit index) was estimated using the MLE (Maximum Likelihood Estimator) using the stata software.

Effects of **E** and/or **S** (i.e., w) on P will be obtained using the following marginal effects equation:

$$dP/dw_i = dP/dU * dU/dw_i,$$

Where, $dU/dw_i = \beta_i$.

However, in the LPM equation to be estimated is:

$$P = \beta_0 + \beta_1 \text{education level} + \beta_2 \text{age} + \beta_3 \text{gender} + \beta_4 \text{employment status} + \beta_5 \text{distance} + \beta_6 \text{cost of health} + \beta_7 \text{disease} + \beta_8 \text{quality of care} + \beta_9 \text{income} + \beta_{10} \text{marital status} + \beta_{11} \text{health insurance} + \beta_{12} \text{family size} + \text{error term } (\varepsilon),$$

Where, P is the probability of service utilization and β_s are the marginal effects.

This model is shown to fit the data. In particular, the logit/probit results do not deviate substantially from the LPM estimates. Furthermore, it is also shown that the heteroskedasticity and the problem of the predicted probability being outside the 0-1 range (the unit interval problem) are easily addressed.

Table 3.1 shows the definitions of the key variables, and the expected signs of the estimated coefficients based on the theoretical model presented in section 3.2 and previous studies

Table 3.1 Description of covariates

Variables	Measurement	Hypothesized sign
Dependent variable		
Utilization of health services (<i>Prob</i>)	1 if yes. 0 if no	
Independent variables		
Economic variables		
Employment status	1 if employed 0 if not employed	Negative
Income	Per capita positive	Positive
Health Insurance	Enrolled=1 Not enrolled=0	Positive
Social Variable		
Age	Years	Negative
Gender	Male=1 Female=0	Negative
Education level	Years of school	Positive
Marital status	married=0 never married = 1 monogamous= 1 0 polygamous = 1 0 living together = 1 0 Divorced = 1 0 Separated = 1 0 Widowed= 1 0	Positive
Household size	Number	Positive
Interaction terms		
Employment Status * Gender	1 if male and employed 0 if otherwise Positive	Negative
Employment status * marital status	1 If married and employed 0 if otherwise	Negative
Insurance * Gender	1 if insured and male 0 if otherwise	positive
Control variables		
Rural	If Rural = 1 urban = 0	Positive

3.4 Data source

The study used primary data from a nationally representative household survey collected by the Kenya National Bureau of Statistics (KNBS) in 2015/2016 – the KIHBS 2015/2016 data set. The sample was drawn from the 47 counties, providing estimates for various indicators representative at the national level.

The 2015/16 KIHBS sample had 5,360 enumeration clusters, divided evenly into four sub-samples. The sampling frame was stratified into urban and rural areas, whereby within each of 47 counties, 92 sampling strata were created with Nairobi and Mombasa counties being completely urban. The sample size was calculated independently for each county, resulting in a national sample of 24,000 households. The study was a cross-sectional research conducted within a 24 months stretching from September 2015 to August 2016. The data collection tool used was a comprehensive questionnaire on health module that gathered information about previous visits to health facilities four weeks prior to survey.

CHAPTER FOUR
DATA ANALYSIS, INTERPRETATION AND DISCUSSIONS OF RESULTS
4.1 Introduction

This chapter covers data analysis, presentation and interpretation of findings on the effect of socio-economic status on healthcare utilization in Kenya.

The descriptive statistics will be presented first followed by the inferential statistics that provide the effect of socio economic status on healthcare utilization in Kenya. STATA was used to compute the estimation of the Logit model, Probit Model, Logit Model and the average marginal effects.

4.2 Descriptive Statistics

Table 4.2: Summary statistics of the study variables

Variable	Observations	Mean	Standard Deviation	Minimum	Maximum
Dependent variable					
Utilization	18,950	0.8188	0.3852	0	1
Independent Variables					
Logage	18,232	2.7842	1.1881	0	6.9058
Logpce_p	18,950	8.4386	0.6556	4.0475	11.4524
Loghsize	18,950	1.4935	0.5949	0	3.3322
Gender	18,950	0.4365	0.4960	0	1
Employed	15,270	0.1557	0.3626	0	1
Insurance	18,943	0.1677	0.3736	0	1
Rural	18,950	0.6586	0.4742	0	1
Education	18,950	0.7354	0.4411	0	1
Polygamous	12,639	0.0700	0.2552	0	1
Living together	12,639	0.0022	0.0470	0	1
Separated	12,639	0.0324	0.1770	0	1
Divorced	12,639	0.0139	0.1172	0	1
Widow or Widower	12,639	0.1164	0.3207	0	1
Never married	12,639	0.3492	0.4767	0	1
Interaction terms					
Employed_gender	15,270	0.0861	0.2805	0	1
Education_gender	18,950	0.3297	0.4701	0	1
Married_employed	15,270	0.1229	0.3284	0	1
Insurance gender	18,943	0.0760	0.2649	0	1

Table 4.1 shows that out of 18,950 respondents, 81.9% sought health care services; 43.7% being male while the rest being female. The mean age was 16 years. From the findings the household sizes mean was 4 members with household per capita being KSh 4,600. The respondents who reported to have a formal education were 73.5% and 15.6% being in formal wage employment. Broken down by gender, 22% of men had formal sector wage jobs, compared to 11% for women. Only 16.8% reported to have an insurance cover. Those residing in rural areas were found to be 65.9% which is much higher compared to those living in urban areas.

On marital status of the respondents, the results show that, 35 % were never married, 11.6% were widows or widowers, 3.2% were separated, 1.4% were divorced and 0.2% were living together. On interaction of the variables, the findings show that males who were employed were 0.09% and males who had attained a formal education were 33%. Out of the respondents, males who had insurance cover were 7.6%. Results further show that out of the respondents, males residing in rural areas were 28.1% only.

4.3 Diagnostic pre-estimation and post-estimation tests

This research project was aimed at obtaining reliable and valid estimates of effects of social and economic variables on health utilization. The following tests were carried out to in order to find out the validity of the data used.

4.3.1 Heteroskedasticity

The results from Appendix 1 on heteroskedasticity indicate that, chi square is 344.75 with p value of 0.0000 at a significance level of 0.005. This shows that the data did not have constant variance. This was corrected by using robust option in stata when estimating the coefficients of the variables in the model, as the option computes the correct standard errors.

4.3.2 Multicollinearity test

To test for multicollinearity, the Variance Inflation Factor (VIF) was used. From the results in appendix 2, the mean VIF is 1.62 which means that the variance of the model coefficient is inflated by a factor of 1.62; hence there is presence of moderate multicollinearity, which is far from the cut-off point, which is around 4.

4.3.3 Normality tests: Shapiro-Wilk W test for normal data

Normality test was carried out to determine whether the sample data was drawn from normally distributed data values. The results in appendix 3 indicate that at the margin of 0.05, we can conclude that gender, marital status, rural, educations are normally distributed. While age, income per capita, household size, employment status are not. We can therefore conclude that the probit model that is associated with following a standard normal distribution may not be appropriate for this data set.

4.3.4 Correlation Analysis

Correlation analysis was carried out in this study to evaluate the strength of relationships between the study variables. The correlation coefficient value varies between -1.0 for a perfect negative correlation and 1.0 for a perfect positive correlation. A coefficient value of 0.0 implies no linear relationship between the movements of two study variables. In appendix 5, the results show low correlations between study variables for example a correlation coefficient between utilization of health services and income is positive and statistically significant. Therefore a 1% increase in income is associated with a 0.005% increase in utilization of health services. Similarly a 1% increase in having an insurance cover is associated with a 0.1177% decrease in household size. Since correlation is not causation, the vice versa of the above statement will be 1% increment in household size is associated with 0.1177% decrease in having an insurance cover. From the examples above and the results from Appendix 5, the correlations are low; therefore there is no perfect correlation between study variables.

4.4 Econometric Analysis

Table 4.3: Lpm, Logit And Probit Model Estimates For Health Service utilization

Variables	Linear probability		Logit model			Probit model		
	Coefficient (t statistic)	Robust Std.Err	Coefficient (t statistic)	Robust Std.Err	(dy/dx) (t statistic)	Coefficient (t statistic)	Robust Std.Err	dy/dx (t statistic)
Log age	-0.0019 (-0.41)	0.0045	-0.0119 (-0.39)	0.0305	-0.0018 (-0.39)	-0.0074 (-0.43)	0.0171	-0.0019 (-0.43)
Logpce_p	0.0549 (9.02)	0.0061	0.3482 (9.00)	0.0387	0.0523 (9.06)	0.1972 (9.03)	0.0218	0.0527 (9.05)
Loghsize	0.0365 (5.84)	0.0062	0.2253 (5.84)	0.0384	0.0338 (5.85)	0.1265 (5.76)	0.0218	0.0338 (5.77)
Gender	-0.0194 (-0.87)	0.0259	-0.0938 (-0.83)	0.1441	-0.0142 (-0.83)	-0.0552 (-0.82)	0.0838	-0.0148 (-0.82)
Employment	-0.0046 (-0.21)	0.0128	-0.0369 (-0.26)	0.0883	-0.0056 (-0.26)	-0.0231 (-0.29)	0.0496	-0.0062 (-0.28)
Insurance	0.0365 (3.59)	0.0102	0.3183 (3.63)	0.0877	0.0447 (3.91)	0.1709 (3.63)	0.0471	0.0433 (3.84)
Rural	0.0257 (3.63)	0.0071	0.1654 (3.55)	0.0466	0.0253 (3.49)	0.0952 (3.59)	0.0265	0.0258 (3.55)
Education	0.0729 (5.75)	0.0128	0.4210 (5.82)	0.0732	0.0693 (5.36)	0.2462 (5.90)	0.0422	0.0708 (5.52)
Marital status	-0.0034 (-0.44)	0.0096	-0.0268 (-0.50)	0.0703	-0.0040 (-0.50)	-0.0151 (-0.50)	0.0390	-0.0040 (-0.50)
Employment * Gender	-0.0430 (-2.37)	0.0183	-0.2547 (-2.19)	0.1168	-0.0408 (-2.06)	-0.1478 (-2.23)	0.0665	-0.0416 (-2.12)
Education * Gender	-0.0043 (-0.19)	0.0238	-0.0613 (-0.50)	0.1251	-0.0093 (-0.50)	-0.0339 (-0.47)	0.0739	-0.0091 (-0.47)
Marital status * Employment	-0.0038 (-0.17)	0.0142	-0.0271 (-0.20)	0.0975	-0.0041 (-0.20)	-0.0127 (-0.16)	0.0547	-0.0034 (-0.16)
Insurance * Gender	0.0282 (1.84)	0.0153	0.1575 (1.23)	0.1281	0.0227 (1.28)	0.0941 (1.36)	0.0693	0.0242 (1.41)
Constant	0.2285 (3.88)	0.0592	-2.16391 (-5.88)	0.3704		-1.1747 (-5.61)	0.2105	
Sample size	15,265		15,265			15,265		
F test /chi- square (p- value)	323.47 (40.0000)		305.44 (0.0000)			304.01 (0.0000)		
Pseudo (R- Squared)	0.0209		0.0212			0.0211		

Interpretation of coefficients of social economic status on Health Care Utilization

Comparing estimates from the Linear probability Model (LPM), Probit Model and Logit Model, the results in relation to utilization health services at a level of significance of $t > 2$ and $p < 0.05$ show that; a 1% increment in age lowers the chance seeking treatment by 0.19% and a 1% increase in income per capita is expected to increase the expected utility from medical care by 0.35 utils, holding other variables constant. The effect of household size is also significant in that a one percentage increase in household size increases utilization of health services by 3.7% on a 1 – 100% scale. Being male decreases the chances of utilization of health services by ≈ 0.019 .

On employment status has a negative impact in that being employed reduces the likelihood of utilizing health services decreases by 0.046 holding other variables constant. A 1% increase in having an insurance increases expected utility from utilization of health services by 0.037 utils. On the other hand residing in rural raises the chance of getting treatment by 0.025.

A one year increase in years of schooling increases expected utility from utilization of health services by 0.42 utils where as being married has a negative effect in that utilization of health services decreases by 0.0034 if one is married holding other variables constant. On interaction terms being married and employed reduces the likelihood of health utilization by 0.0038. Similarly being a female and employed decreases the probability of health utilization by 0.0038.

The F-test shows that the coefficients of the independent variables across the three models jointly are statistically significant in explaining utilization of health services at a significance level of $p < 0.05$.

Interpretation of marginal effects of social economic status on Health Care Utilization

Based on the analysis of the three models the Marginal effects reveal that; That for a 1% rise in per capita income raises the probability that a person will visit a health facility by 5.2%, holding all other variables constant. On the household size a 1% increase in

household size will increase likelihood of utilization of health services by 3.7% and when age increases by 1% the chances of utilizing health services decreases by 0.2%. Gender has a negative effect in that being a male lowers the chance of visiting a clinic by 1.5% compared relative to a female.

Being employed reduces the chances of utilizing health services by 0.5% holding other variables constant. In addition having insurance increases the likelihood of utilization of health services by 4.3%. An individual residing in rural area has a 2.6% higher probability of seeking medical help compared to an individual residing in urban areas. A one year increase in years of schooling increases probability of utilization of health services by 7.1%. On the other hand being married reduces the likelihood of health utilization by 0.4% holding other variables constant. On the interaction between social economic factors; being a male and employed increases the probability of health utilization by 4.3% while being a male and having insurance increases the probability of health utilization by 0.03%.

From the above analysis, the end results from the LPM, Probit Model and Logit Model are essentially making the same predictions. The coefficients of Logit and Probit differ usually by about 1.6. Looking at marginal effects the Logit and Probit will also make the same predictions.

Table 4.4: A LPM on effects of social-economic status on health service utilization

Variables	Coefficient	Robust Std. Err.	t statistic	p>[t]	[95% Conf. Interval]	
Logage	-0.0220	0.0100	-2.22	0.026	-0.0415	-0.0026
Logpce_p	0.0573	0.0067	8.59	0.000	0.0442	0.0704
Loghsize	0.0373	0.0067	5.54	0.000	0.0241	0.0505
Gender	-0.0064	0.0278	-0.23	0.816	-0.0610	0.0481
Employment	-0.0060	0.0130	-0.46	0.644	-0.0314	0.0195
Insurance	0.0316	0.0111	2.84	0.004	0.0098	0.0534
Rural	0.0326	0.0079	4.14	0.000	0.0172	0.0480
Education	0.0637	0.0137	4.65	0.000	0.0369	0.0906
Polygamous married	0.0014	0.0149	0.09	0.926	-0.0277	0.0305
Living together	0.0632	0.0563	1.12	0.261	-0.0471	0.1736
Separated	-0.0112	0.0214	-0.52	0.601	-0.0533	0.0308
Divorced	-0.0760	0.0345	-2.21	0.027	-0.1436	-0.0085
Widow or widower	0.0043	0.0138	0.31	0.758	-0.0228	0.0313
Never married	-0.0190	0.0130	-1.47	0.143	-0.0445	0.0064
Employed_gender	-0.0399	0.0192	-2.08	0.037	-0.0776	-0.0023
Education_gender	-0.0185	0.0261	-0.71	0.478	-0.0696	0.0326
Married_employed	-0.0055	0.0158	-0.35	0.729	-0.0365	0.0255
Insurance_gender	0.0307	0.0172	1.78	0.075	-0.0031	0.0644
Constant	0.2875	0.0732	3.93	0.000	0.1439	0.4311
Sample size	12,632					
F - test (p- value)	15.17 (0.0000)					
R- squared	0.0230					

Discussion of the results

The results of estimates in Table 4.4 from the Linear probability Model (LPM) in relation to utilization of health services at a level of significance of $t > 2$ and $p < 0.05$ and holding other factors constant show that; on onset of illness the probability of utilizing health services is affected by; income per capita, education level, household size, having an insurance, living in rural areas. Being a male decreases probability of service usage and being insured increases the probability of utilizing health services during an illness period.

From the results, increase in income significantly increases the probability of utilizing health services. This is attributable to the fact that an individual has ability to pay for the health services and hence has a purchasing power. This corresponds with (Oladipo 2014) who found that an individual with high income has a higher likelihood of utilizing health services.

The coefficient on age is statistically significant. The findings show that age is negatively associated with demand in that a one percentage increase in age lowers the likelihood of getting treated when sick. This finding contradicts (Muriithi2013) and Ellis and Mwabu (2004) whose results showed that as an individual ages the demand for health care increases.

The effect of household size is also statistically significant in that a 1% increase in household size increases utilization of health services. This finding corresponds with (Oladipo 2014) and (Muriithi 2014). The results also show that being male the coefficient is statistically significant in that it decreases the probability of utilizing health services. The findings correspond with Ellis and Mwabu (2004) and Muriithi (2014) studies that validate this findings.

Being employed is statistically significant and the likelihood of utilizing health services decreases as one is employed holding other variables constant. This could be because the employed individuals are engaged in the informal sector. Hence this can be translated to low earnings when they are away visiting health practitioner. Kimani et al. (2016) studies concur with these findings.

Having an insurance cover is significant and increases usage of services. Regarding this issue, Oladipo (2014) agrees that insurance is one of the major predictors of utilization.

Moreover, an interaction between insurance and gender shows a positive association between service usage and the male subsample (Chi-square = 37.02; $p = 0.0000$). A similar finding holds for the interaction between employment status and gender (Chi-square = 20.35; $p = 0.0000$).

The results further show that residing in rural areas as well one year increase in years of schooling are statistically significant and increases utilization of health services holding other variables constant. Kimani et al. (2016) findings showed that having a formal education and residing in rural areas increased the likelihood of visiting a health practitioner.

From the estimations done on the 3 models we can conclude that the LPM yields estimates on effects of socioeconomic status on health care utilization that are just as accurate as those estimated by Logit and Probit regressions. The F-test shows that the coefficients of the independent variables in the linear probability model jointly are statistically significant in explaining utilization of health services at a significance level of $p < 0.05$.

For this reason, the LPM (the OLS model) is preferred because it is easy to estimate and is equally straightforward to interpret its results. As already noted, the heteroskedasticity in the LPM model is addressed by using the robust estimation option in stata. However, in a few cases in the data set used, the predicted probabilities of service utilization exceed unity but are never below zero. Wooldridge (2002) has shown that this is not a problem because predicted probabilities above or equal to a cut-off upper point, say, 0.75, can be set to one, while the probabilities below that point are set to zero. It is easy to show using the *sum* command in stata that the *redefined* probability of service utilization always lies between zero and one, as dictated by the definition of probability.

In policy analysis, the interest of decision makers is to know whether, after implementation of a policy measure, such as the UHC or a cash transfer to a target group, the probability of service utilization by the whole population is above or below a particular point, say, 0.50. Interest is never on the probability of a particular person using health care. Even in that case,

Wooldridge (2013) shows that the problem is not as severe as it seems because it is entirely due to the linearity property of the LPM framework. It is easy to show that due to linearity, a few outlier variables (too large or too low values) can drive the predicted probability outside of the unit interval, the (0/1) range. Scaling (e.g., multiplying or dividing such variables by a certain number, a common practice in the literature), would restore the predicted probability to a desired interval. The outliers can also be removed from the data, since by the very definition of outliers they are not representative of the sample. Truncation of extreme probability predictions into the required interval is another alternative but that is an arbitrary correction.

CHAPTER FIVE

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

5.1 Introduction

This chapter presents a summary of the findings concerning effects of socioeconomic status on health care utilization in Kenya. In this chapter, policy conclusions based on the study findings, and recommendations to the decision makers in health sector are briefly outlined.

5.2 Summary of the findings

The main aim of the study was to examine the role played by socioeconomic variables on probabilities of visiting the health system in the event of sickness. The study used primary data collected by the KNBS, using the approach described in the methods section of this paper.

Theoretical and empirical literature was reviewed to assess what was already known as to the effects of socioeconomic status on health and health care use in Kenya. LPM, Logit and Probit regressions were used to look into the roles of socioeconomic factors in health service uptake in Kenya. The comparison of the estimates from the three models showed that the linear probability model, despite its simplicity performed quite well relative to the findings from the non-linear models. The key dependent variable in the study was a service utilization dummy, and the independent variables included age, gender, household size, household income per capita, education level, employment status, marital status and area of residence.

The findings from this study show that social factors like being a male, and aging into adulthood, decreases the likelihood of utilizing health services. The study further shows that being in certain marriage categories decreases chances of utilizing health services, highlighting the need to disaggregate the marriage dummy into different marriage groupings in gender and policy analysis. On the other hand, a large, household size, having attained a formal education and living in rural areas increases the probability of utilizing health services. Economic variables, like having a higher income and being a member of an insurance scheme raises one's chance of using services when ill. In addition, the findings show that wage sector jobs lower chances of service use. As to the relationship between interactions of social and economic variables and service use, the results show a negative link for interactions between gender and employment status and

between insurance coverage and gender. The results show that being male and having a wage job lowers the probability of using health services, whereas being insured and being a male, raises service demand. Employed women are more likely to use health services relative to women without formal sector wage employment. All the other social and economic interaction terms are not statistically significant in my sample data.

5.3 Conclusion

To conclude, the findings of the study show that socio economic factors have a role as a determinant in utilization of health services. Findings have shown that large households, higher education levels, higher income, residing in rural areas and having insurance cover increases service utilization. On the other hand utilization of health services decreases among the employed, among some marriage arrangements, and among males.

Since overall utilization of health services is still low and the disease burden is high, there is need for government to come up with measures to reduce health care utilization disparities related to gender, literacy levels, income, enrolment to health insurance, access to wage employment. Government could also come up with ways of increasing health service utilization among people in certain marriage groups, such as the widowed, divorced, and as separated.

5.4 Policy Recommendation

From the results of the study there are a number of policy implications that can be made.

First there is a need for the government in conjunction with the Ministry of Health to invest more in Health system and expand the health structures especially to places where health utilization is high for example rural areas.

Second, sensitization on health matters and health seeking behaviour among the male group is important as well empowering them with knowledge on importance of routine check-up which can be lifesaving. As the utilization in this group increases, best possible health outcome is achieved due to timely receipt of care.

Third, government of Kenya is vigorously pursuing the goal Universal Health Coverage, one of the flagships of the government's development agenda. With the recent biometric registration for universal health care coverage scheme, social health insurance is likely to be embraced. However active participation is recommended especially from low social

economic groups and people in the informal sector. This can be improved by ensuring the public access to information on the Social Health Insurance (SHI). The sustainability of the achievement of Universal Health insurance Coverage will be achieved through enrolment of an insurance cover; hence cushioning against catastrophic health expenditures. This translates to increased health improvement as people will be visiting health facilities for service uptake.

Four, from the analysis high literacy levels are associated with increased utilization of health services. This therefore calls for the government of Kenya to continue ensuring a quality education which accessible to all, especially the marginalized and lower economic social groups.

Lastly, Kenyan employers need to create a healthy environment and so that on onset of illness an employee is granted a sick leave in order to seek health services. This will increase health service utilization among the employed groups, as employees will be more productive.

5.5 Recommendations for further research

The study looked into how socioeconomic factors affected health care usage in Kenya shortly before the UHC was rolled out to all counties, first on a pilot basis. More studies in examining the effects of socioeconomic status in utilization of health services after the roll out of the UHC are needed to establish progress made so far towards attaining UHC, including its impacts on health.

REFERENCES

- Andersen, R. M., & Newman, J. F. (1973). Societal and individual determinants of Medical care utilization in the United States. *Milbank Mem Fund Q Health Soc.*, 51(1):95–124.
- Celik Y. & Hotchkiss D.R. (2000). The socio-economic determinants of maternal health care utilisation in Turkey. *Social Science & Medicine*, 50(12).
- Chuma J, Okungu V. (2011) Viewing the Kenyan health system through an equity lens: implications for universal coverage, *International Journal for Equity in Health*, 10 (22): 1-14.
- Ellis, R.P. & Mwabu, G.M. (2004). *The Demand for Outpatient Medical Care in Rural Kenya*. Nairobi: University of Nairobi.
- Gareth et al (2014). An Introduction to Statistical Learning: *With Applications in R*, 4:131-134.
- Grossman, Michael (1972), On the Concept of Health Capital and the Demand for Health, *Journal of Political Economy*, 80:223-255.
- Kimani D., Mugo M., Kioko U., (2016). An econometric Analysis of health care utilization in Kenya. *European Scientific Journal*, 12 (15).
- Kutzin J. (2013). Health financing for universal coverage and health system performance: concepts and implications for policy. *Bull World Health Organ*, 91: 602-611.
- Matehesi Z., Meulemans, H., & Timmerman, C., (2005). Healthcare-seeking behaviour among clinic-based tuberculosis patients. *Acta Academia Supplementum*, (1): 154-172.
- Mechanic D. (1978). *Medical sociology, A comprehensive text*, (2nd ed). New York: The free Press.
- Mechanic D. (1995). *Medical sociology*, (4th ed). New York: The free Press.

- Mills A, Ally M, Goudge J, et al (2012). Progress towards universal coverage: the health systems of Ghana, South Africa and Tanzania. *Health Policy Plan*, 27(1):i4–12.
- Ministry of Health Government of Kenya (MOH). (2014). Kenya health policy 2014–2030: *towards attaining the highest standard of health*. Nairobi, Kenya.
- Muriithi, M.K. (2013). The Determinants of Health-Seeking Behaviour in a Nairobi Slum, Kenya. *European Scientific Journal* 9(8): 1857-7881.
- O'Donnell O, van Doorslaer, et al (2007). Analysing Health Equity Using Household Survey Data: *A Guide to Techniques and Their Implementation*. Washington, DC: The World Bank.
- Oladipo, J.A (2014). Utilization of Health Care Services in Rural and Urban Areas. A Determinant Health Care Delivery System.
- Onwujekwe O & Uzochukwu B (2005). Socio-economic and geographic differentials in costs and payment strategies for primary health care services in Southeast Nigeria. *Health Policy*, 71 (3): 383-397.
- Parsons, T. (1951). *The Social System*. Glencoe, IL: Free Press.
- Rebhan, D. (2015). Health Care Utilization: *Understanding and applying theories and models of health care seeking behaviour*: Case Western Reserve University.
- Rebhan, D. P. (2011). *Health Care Utilization: Understanding and applying theories and models of health care seeking behaviour*, Case Western Reserve University, 1-19.
- Suchman, E. (1965). Social Patterns of Illness and Medical Care. *Journal of Health and Human Behaviour*, 6(1): 2-15.
- Wooldridge (2013). *Introductory Econometrics: A Modern Approach*, South-Western.
- Yip A, Kephart G, Veugelers PJ (2002). Individual and neighbourhood determinants of health care utilization. *Canadian Journal of Public Health*, 93 (4): 303-307.

APPENDICES
Appendix 1: Heteroskedasticity Test

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
Ho: Constant variance	
chi2(1) = 344.75	Prob> chi2 = 0.0000

Appendix 2: Multicollinearity Test

Table of Variance Inflation Factor

Variable	VIF	1/VIF
Employment * Gender	2.50	0.3996
Employment	2.26	0.4432
Insurance * Gender	2.02	0.494909
Insurance	1.91	0.5227
Logpce_p	1.45	0.6901
Logage	1.36	0.735162
Loghsize	1.34	0.744556
Education	1.24	0.804655
Rural	1.13	0.881751
Marital status	1.12	0.889052
Mean VIF	1.62	

Appendix 3: Normality Tests: Shapiro-Wilk W Test for Normal Data

Variable	Obs	W	V	z	Prob>z
Logage	18,232	0.94484	456.898	16.644	0.0000
Logpce_p	18,950	0.99488	43.806	10.281	0.0000
Loghsize	18,950	0.93012	597.701	17.390	0.0000
Gender	18,950	0.99998	0.196	-4.437	1.00000
Employment	15,270	0.99935	4.630	4.147	0.00002
Insurance	18,943	0.99953	3.988	3.762	0.00008
Rural	18,950	0.99995	0.400	-2.490	0.99360
Education	18,950	0.99987	1.087	0.228	0.40982
Marital status	18,950	0.99982	1.542	1.177	0.11956
Employment * Gender	15,270	0.99860	10.031	6.240	0.00000
Insurance * Gender	18,943	0.99869	11.197	6.571	0.00000

Appendix 4: Logit and Probit Model Estimates for Effects of Social Economic Status on Health Care Utilization

Variable	Logit Model				Probit Model			
	Coefficient	dy/dx	Robust std. Err	t statistic	Coefficient	dy/dx	Robust std.Err	t-statistic
Logage	-0.0119	-0.0018	0.0045	-0.40	-0.0073	-0.0019	0.0171	-0.44
Logpce_p	0.3483	0.0523	0.0061	9.33	0.1972	0.0527	0.0218	9.32
Loghsize	0.2263	0.0340	0.0062	6.03	0.1270	0.0338	0.0218	5.93
Gender	-0.0382	-0.0058	0.0259	-0.29	-0.0241	-0.0148	0.0838	-0.31
Employment	-0.0314	-0.0047	0.0128	-0.22	-0.0197	-0.0062	0.0496	-0.24
Rural	0.1972	0.0302	0.0071	3.22	0.1126	0.0258	0.0265	3.26
Insurance	0.3227	0.04529	0.0102	3.67	0.1733	0.0433	0.0471	3.67
Education	0.4232	0.0697	0.0128	5.93	0.2473	0.0708	0.0422	5.98
Marital status	-0.0264	-0.0039	0.0096	-0.49	-0.0148	-0.0040	0.0390	-0.49
Employment * Gender	-0.2649	-0.0426	0.0183	-2.26	-0.1538	-0.0416	0.0665	-2.31
Education * Gender	-0.0674	-0.0102	0.0238	-0.55	-0.0373	-0.0090	0.0739	-0.51
Married * Employment	-0.0290	-0.0044	0.0142	-0.21	-0.0140	-0.0034	0.0547	-0.18
Insurance * Gender	0.1469	0.0212	0.0153	1.14	0.0883	0.0242	0.0693	1.27
_cons	-2.1904			-6.07	-1.1897			-5.48
Sample size	15,265				15,265			
Chi- square (p value)	308.09 (0.0000)				306.63 (0.0000)			
Pseudo R ²	0.0207				0.0206			

APPENDIX 5: CORRELATION MATRIX

	Utilization	Logage	Logpce_p	Loghsize	Gender	Employment	Insurance	Rural	Education	Marital status	Employment* Gender	Insurance * Gender
Utilization	1.0000											
Logage	-0.0518 (0.0000)	1.0000										
Logpce_p	0.0852 (0.0000)	0.0832 (0.000)	1.0000									
Loghsize	0.0360 (0.0000)	-0.2265 (0.000)	-0.4017 (0.0000)	1.0000								
Gender	-0.0119 (0.1006)	-0.0239 (0.000)	0.0019 (0.5616)	-0.0181 (0.0000)	1.0000							
Employment	-0.0231 (0.0044)	0.2946 (0.000)	0.1772 (0.0000)	-0.2453 (0.0000)	0.1300 (0.000)	1.0000 (0.0000)						
Insurance	0.0722 (0.0000)	0.0614 (0.000)	0.3207 (0.0000)	-0.1177 (0.0000)	0.0045 (0.168)	0.1520 (0.0000)	1.0000					
Rural	-0.0019 (0.7975)	-0.0427 (0.000)	-0.3387 (0.0000)	0.1606 (0.0000)	-0.0088 (0.008)	-0.1380 (0.0000)	-0.1404 (0.0000)	1.0000				
Education	0.0349 (0.0000)	0.2508 (0.000)	0.1859 (0.0000)	-0.0533 (0.0000)	0.0519 (0.000)	0.0935 (0.0000)	0.1324 (0.0000)	-0.0919 (-0.09)	1.0000			
Married	-0.0051 (0.4787)	-0.0161 (0.000)	0.0502 (0.0000)	-0.1151 (0.000)	-0.0772 (0.000)	0.1264 (0.0000)	0.0398 (0.0000)	0.0009 (0.793)	-0.2929 (0.0000)	1.0000		
Employment* Gender	-0.0375 (0.0000)	0.2321 (0.000)	0.1367 (0.0000)	-0.2109 (0.000)	0.3430 (0.000)	0.7748 (0.0000)	0.1154 (0.0000)	-0.1055 (0.000)	0.0801 (0.0000)	0.0929 (0.000)	1.0000	
Insurance * Gender	0.0423 (0.0000)	0.0434 (0.000)	0.2241 (0.0000)	-0.1034 (0.0000)	0.2916 (0.000)	0.1660 (0.0000)	0.6768 (0.0000)	-0.1027 (0.000)	0.0937 (0.0000)	0.0224 (0.000)	0.2558 (0.0000)	1.0000