

**The Determinants of Health
Existing Theory and Application of Theory to Data from Newfoundland**

Lynn M. Gambin^{*}

**Department of Economics
Memorial University of Newfoundland**

July 29, 2002

^{*} This paper was submitted to meet the requirements for the Master of Arts program in Economics at Memorial University of Newfoundland. Many thanks are extended to Dr. Douglas May and Dr. Noel Roy of Memorial University of Newfoundland. Also thanks to Gwenda Drover and others at the Newfoundland Statistics Agency and Economic Research & Analysis Division of the Newfoundland and Labrador Department of Finance. The author accepts all responsibility for this paper. Comments or questions may be submitted directly to the author by email at s79lmg@mun.ca.

Table of Contents

1.	Introduction	1
2.	Literature Review	2
3.	Description of Data	8
3.1.	Newfoundland Adult Health Survey 2001	8
3.2.	Health Indicator	9
3.3.	Demographic Factors	10
3.4.	Financial/Work-Related Factors	11
3.5.	Lifestyle Factors	14
3.6.	Social Factors	15
3.7.	Location	16
4.	Econometric Analysis	18
4.1.	General Procedure – Ordered Logistic Regression	18
4.1.1.	Pairwise Regresssion	20
4.1.2.	Multiple Independent Variables	21
4.2.	Ordered Logistic Regression Results	23
4.2.1.	Pairwise Regressions	23
4.2.2.	Multiple Independent Variables	28
5.	Summary and Conclusions	35
	Bibliography	39
	Appendix A – Relevant Survey Questions	40
	Appendix B – Regression Output	44

1. Introduction

The Government of Newfoundland and Labrador spends millions of dollars on health care each year. One of the reasons for such considerable spending is that the government wishes to improve the health of citizens. These health care expenditures consist of salaries of doctors, nurses and other health professionals, hospital maintenance, drugs and other items. Since there is such spending on health care in the province, and worldwide, it is little wonder that health has become an increasingly important area for economic studies.

If governments and individuals are going to spend billions of dollars on health care and health related goods and services each year then it would seem desirable to know whether this money is being spent in the best possible manner. Is too much being spent on treatment and too little on preventative measures? Or is it that spending on health care does little to actually improve health? Should money instead be spent on other programs and services aimed at improving all aspects of a person's life and thereby improving health as well? To determine the answers to these and other important questions concerning health it is not only helpful, but also necessary, to examine the fundamental question of what actually determines a person's health. Through an understanding of the various determinants of health status it will become easier to determine the alternative actions to take to improve the health of individuals and populations.

In this analysis, the determinants of health are investigated using a sample from the Province of Newfoundland and Labrador. The analysis attempts to build upon existing theory concerned with the determinants of health status through use of a new data set and through rare variables concerned with financial circumstances and stress. The methods used agree with existing theory in health economics and the results of the analysis could provide much

useful information and direction for policymakers concerned with promoting health status and diminishing health inequalities within a population.

This paper is divided into six major sections. This introduction constitutes the first section. In Section 2, a brief review of some of the existing literature on the determinants of health is provided. The review focuses on identifying what factors are presently considered to determine health and the findings of major studies in the United Kingdom. The direction of current studies and the obstacles encountered in such studies are also considered. Section 3 overviews the data set used for this analysis. Summary statistics as well as some discussion of the survey used to compile the data are given. Section 4 outlines the econometric analysis of the data and provides the results of estimation of various models. Finally, Section 5 concludes the paper. The results of the data and regression analysis are interpreted, summarized and compared to expected or hypothesized relationships. It also offers some recommendations concerned with changing health status through the determinants of health examined in Section 4 and provides some suggestion for future investigation.

2. *Literature Review*

Studies of the determinants of health status have been undertaken by many researchers for quite some time. Experts from a vast number of disciplines have undertaken research to examine and determine the biological, social and economic factors that affect or determine the health of individuals or populations. Differences in health status among individuals cannot be attributed to the simple fact that individuals are different. People are unique. There is no argument against that fact. However, to say that significant differences in the health status among people are due only to genetic differences is naïve. Studies into the determinants of health have evolved from exploring the more simplistic notion of health

status being affected only by biological and genetic influences to focusing on the contribution that many other factors have on health status.

In order to perform a reliable analysis of the determinants of individual health status in Newfoundland, it is first necessary to examine previous studies concerned with the determinants of health through time. Studies of health determinants have change considerably in the past few decades. Not only are there many more statistical and econometric tools¹ now used to examine the relationship between health and various factors, but the actual factors that are considered to affect health have changed over time. In addition, the explanations of how various social and economic factors impact a person's health have also progressed much since the earliest studies. Studies into the determinants of health today are not narrow in their focus. Instead, such studies often employ expertise and viewpoints from many different fields, including epidemiology, biochemistry, economics, and biology.

One of the earlier, and most influential, publications concerning health determinants was the Report of the Working Group on Inequalities in Health. The Black Report, as it is more commonly know, was submitted to the British Secretary of State in 1980. The Report addressed the presence of health inequalities in Britain. The Black Report was an attempt, authorized by the Government, "to explain trends in inequalities in health and to relate these to the policies intended to promote health."²

The Black Report indicated that inequalities in health are affected not only by the quality of medical services and technologies available but also by such factors as income, environment, education, employment, transport, and housing. Such factors are often more favourable to the better-off individuals in society. Various observations on the relationship

¹ Jones and O'Donnell (2002) provide numerous examples of the many econometric techniques applicable to health data.

² Townsend and Davidson, p.14.

between health and such factors as those above were made through analysis of the General Household Survey (GHS), the Hospital Inpatient Enquiry (HIPE), and other morbidity and mortality data in Britain. The Working Group's analysis allowed them to not only make observations on trends in health inequalities but also to offer some explanation of how or why health is affected by various lifestyle, social, and biological factors.

The Working Group observed higher mortality rates in men than in women at all ages, higher mortality rates in occupational class V than in higher classes, as well as differences in health, as indicated by mortality rates, between regions and racial and ethnic groups. Higher mortality rates were also evident in people who rented their homes from local authorities versus people who rented from private landlords or who owned their homes. Such observations support the idea that there are numerous factors that determine health.

The Black Report offered various explanations for the links between health and the above mentioned factors. The report explains four categories of theoretical explanations of the relationship between health and inequality. These categories are: artefact explanations; theories of natural or social selection; materialist or structuralist explanations; and cultural/behavioural explanations. The authors of the report consider many of the Report's observations to be sufficiently explained by the "materialist" approach. This category of explanation argues that health is influenced by various factors such as income and employment because of deprivation. Having low income means that a person cannot afford nutritious food or high quality housing. This type of explanation has become less accepted in more recent studies.³ However, the Black Report also indicates that the materialist approach may not fully or sufficiently explain the pathway through which other factors affect health. Other explanations may better suit other factors such as race or occupational class. Other

³ Milyo & Mellor, Lynch et al, Wilkinson.

explanations may also be needed for the effect of the same variable on health at different stages in the lifecycle.

The conclusions arising from the Black Report were important in establishing direction for further research into health inequalities. The report also ignited vigorous debate over the subject and the issue of how to best employ resources in order to reduce health inequalities. This matter of best resource allocation is still unresolved as there is no absolute understanding of the determinants of health.

The knowledge base concerning the determinants of health was also expanded with the development of another study in the United Kingdom. The Whitehall Studies were conducted almost twenty years apart. In the original study⁴ an inverse relationship between employment grade and poor health was found. The first Whitehall study (1967) made clarified the importance of a person's position in the social hierarchy in the determination of health status. The individuals studied were all employed and so none were considered poor or deprived in absolute terms. The relative position of individuals then is what is important.

The second Whitehall study took place in 1985-1988. The gradient in health status across positions in the civil service hierarchy is also evident in the second study. The Whitehall II study, like its predecessor, hypothesizes that psychosocial factors, such as social support networks and work-related stress, have major influence on the social gradient in health. The Whitehall II study has generated much interest in the relationship between health status and social and biological factors.

The Black Report and the Whitehall studies have brought attention to the fact that health is not just a product of genetics and medical services. Both studies indicated that other socioeconomic factors have great influence on the health of individuals. Most obvious from

⁴ Smith.

both studies is that poorer health is associated with lower occupational classes. How such a factor influences health has been the focus of further research.

A person's position in the social or occupational hierarchy may impact his health because of what that position implies about the control that person has over his job and life. Employees further down the corporate ladder would typically have less autonomy and be required to follow the directions of higher ranked employees. Such lower ranked employees would not feel in control of their professional lives and could also possibly suffer from lower self-esteem due to lack of control.

The impact of social status is also examined in *Social Determinants of Health* by Marmot and Wilkinson. In addition to social standing, this book also covers the impacts of social cohesion, relationships, age, and unemployment on health status. Various health indicators, including mortality, morbidity, mental health and sick days, are examined by the contributors to this book. One of the more interesting findings in this publication is that being married has been found to be more beneficial to men's health than women's. Women have been found to benefit from relationships with friends and relatives just as much as they do from marriage.⁵

Income, age, gender, marital status, employment status, and lifestyle factors such as smoking are commonly cited determinants of health. Another important factor that has been examined more in recent health studies is stress. Evans (paper), Brunner and Marmot (in social marmot) as well as many others have examined the impact of stress and stress-related variables on health status. Stress is considered to not only have a direct impact on health but it is also viewed as a pathway through which other factors affect health. Income for instance,

⁵ Stansfeld, 1999.

enables one to afford the material necessities for good health. However, stress caused by lack of income or changes in income has definite impact on the health of individuals.

The presence of stressful situations does not definitely translate into poor health status. How a person deals with stress, the duration of stressful situations, and the support a person has in dealing with stress all impact how strong is the influence of stress on health. Indicators of stress provide important details about the interaction of genetics, environment and biology in the determination of health. Many theories exist on the influence of stress and these interactions.⁶

The determinants of health that are prevalent in studies of the social and economic determinants of health vary but those factors discussed here are common to most studies. One of the major obstacles to studies of this nature is the lack of consistent methodology and reliable collection of data suited to investigations into the determinants of health. The measurement of health status varies across studies and often the data does not exist for the most desirable measures. Many policy makers as well as social scientists have attempted to design structured frameworks with sound data collection and statistical and econometric methods for studying the determinants of health.⁷ This paper aims to contribute to the goal of reliable analysis of data concerned with health determinants and follows methods that are soundly rooted in prevailing theory.

⁶ Evans, 2002.

⁷ Ryan-Nicholls et al, 2000.

3. *Description of Data*

3.1. *Newfoundland Adult Health Survey 2001*

The research undertaken in this paper utilizes data gathered in 2001 through the Newfoundland Adult Health Survey. The survey was administered by Omnifacts Research on behalf of the Government of Newfoundland and Labrador. There were 7949 respondents who were 18 years of age and older. However, when those observations for which the health indicator response was invalid were dropped the effective sample size became 7936. The respondents were drawn from the populations of four health regions in the province. These health regions include the St. John's Community Health Region, the Eastern Community Health Region, the Central Community Health Region, and the Western Community Health Region.

The Newfoundland Adult Health Survey garnered responses pertaining to a person's health over the 12 months preceding the survey. Questions regarding the person's mobility, physical activities, health services usage, and physical condition were posed. Demographic and economic information was also gathered. The 2001 survey was based largely on the Newfoundland Panel on Health and Medical Care Adult Health Survey in 1995⁸. The 1995 survey was also a telephone survey. There were about 12,000 respondents to the 1995 survey.

The 1995 survey, like the 2001 survey aimed to collect information on health status, health practices, demographic characteristics and medical care utilization. The 2001 survey, in addition to being more broadly administered, also contained more information on each respondent than did the 1995 survey. Information on stress and social relationships, for

⁸ Segovia et al, 1997.

instance, are variables that add much to the 2001 survey. Such variables are particularly important for studies of the determinants of health.

Most, if not all, of the information contained in the 2001 Newfoundland Adult Health Survey could be used in analyzing the determinants of health status in the four health regions surveyed. However, the survey consists of 87 questions, some with multiple parts. While the information is useful, it is necessary to limit the number of variables used in the analysis so that the scope of the analysis is suitably focused. After reviewing the survey the number of variables considered relevant to this study was reduced to fourteen.⁹ These variables were selected because they were similar to those used in other studies of the determinants of health, as discussed in Section 2, and also because of the intuitive appeal of including them in an analysis of the determinants of health status. Some differences between this study and others are present, however, the variables used here are in line with those used in similar studies and the types of variables expand on previous studies into the determinants of health status.

3.2. *Health Indicator*

For the analysis that follows in section 4, the item used to indicate health is survey question 48. This question asks, “Would you say your health is...” and offers five responses, “excellent”, “very good”, “good”, “fair”, and “poor”. This indicator is referred to as self-assessed health status (SAH) from this point on. The distribution of responses to this question is shown in Table 3.1. As shown in Table 3.1, inequalities in health exist in the four health regions surveyed. Over forty-five percent of respondents indicated that they considered their health to be very good. Over fifteen percent of the respondents considered themselves to be in poor health. The aim of the econometric analysis in section 4 is to identify some of the

⁹ The relevant questions from the survey are shown in Appendix A.

underlying determinants of these differences in health across respondents to the 2001 survey. The SAH summarized here is used as the dependent variable in the analysis. The use of this indicator agrees with the methods used much of the existing on the determinants of health. While the SAH indicator is somewhat subjective it is still considered as a reasonable indicator of health. The administrators of the 2001 survey have attempted to ensure that the responses are consistent and realistic. By matching responses to the 2001 survey with responses on the 1995 survey, as well as comparing people's MCP records to their responses, the consistency and reliability of responses can be ascertained.

SAH	Frequency	Percentage
Excellent	962	12.12%
Very good	3588	45.21%
Good	2167	27.31%
Fair	970	12.22%
Poor	249	3.14%
Total	7936	100.00%

Table 3.1: Distribution of responses to question 48 (Self-Assessed Health Status (SAH)).

As earlier stated, the 2001 survey has resulted in a detailed data set. For the purposes of this paper there is no need to analyze all information included in the survey. For instance, question 3 asks about particular leisure activities a person might take part in. Such detailed information would not add anything to the econometric analysis undertaken in this paper. The variables that are relevant for this paper include financial/work related factors, demographic factors, lifestyle choices, and social factors. These variables are described in further detail here.

3.3. Demographic Factors

Question 20 of the survey asks the respondent's gender. 3894 (49.1%) of survey respondents were male. 4042 (50.9%) of respondents were female. As described in section 2,

previous studies have found differences in numerous health status indicators between men and women. In North America, women have longer life expectancies than men. Including gender in the analysis of health determinants is not a new idea however, it is important to include the variable so that the influence of other factors on health is isolated from the impact of gender.

Age is also an important factor to include in any study of health determinants. The survey asked for respondents' ages in question 86. Not all respondents were willing to reveal their complete date of birth. However, all respondents could be put into various ranges of age. The distribution of respondents according to age is shown in Table 3.2.

Age	Frequency	Percentage
18 – 24 years	619	7.80%
25 – 34 years	1052	13.26%
35 – 44 years	1888	23.79%
45 – 54 years	1853	23.35%
55 – 64 years	1356	17.09%
65 years and older	1168	14.72%
Total	7936	100.00%

Table 3.2: Distribution of survey respondents according to age.

Like gender, age is another factor whose affects on health inequalities are commonly known. People's health typically worsens as they get older, at least after a particular age. This is the natural progression of the lifecycle.

3.4. Financial/Work-Related Factors

Question 82 of the survey asked respondents, "What is your current employment status?" This is another important factor in this analysis as a link between poorer health and unemployment has been found in other studies. The distribution of responses to question 82 is shown in Table 3.3. The greatest proportion of respondents indicated that they were employed by a company or organization at the time of the survey. A significant proportion of

respondents also indicated that they were not employed for pay. There seems to be a great variety in employment or labour force status among respondents.

Employment Status	Frequency	Percentage
Self-employed	467	5.88%
Employed by company/organization	2922	36.82%
Not employed for pay	2351	29.62%
Retired	1869	23.55%
Student	314	3.96%
Refused	13	0.16%
Total	7936	100.00%

Table 3.3: Distribution of respondents according to employment status.

The Newfoundland Adult Health Survey also includes a very interesting variable associated with financial factors. This variable, which corresponds to question 83, will be referred to as a person's financial security throughout the remainder of the discussion. Question 83 asked, "Given your annual income and considering your monthly expenditures on food, car, housing, etc., would you consider your financial circumstances to be..." The possible responses, as well as their respective frequencies, are shown in Table 3.4. There is wide variation of responses to this question and significant proportions of respondents selected each of the available responses.

The financial security variable is not only interesting but also somewhat unusual. In reviewing the literature on the determinants of health the impact that stress due to

Financial Circumstances	Frequency	Percentage
Very good	1472	18.55%
Good	1947	24.53%
Satisfactory	2346	29.56%
Just getting by	2002	25.22%
Can't cope	135	1.70%
Refused	18	0.23%
Don't know/not sure	16	0.20%
Total	7936	100.00%

Table 3.4: Distribution of respondents according to financial security.

financial concerns has on health has certainly been addressed. However, in previous studies such a direct indicator of this stress has been absent.

An indicator of income is also certainly essential to include in any study concerning the socioeconomic determinants of health. The survey asked respondents, “What is the best estimate of the total income of all household members from all sources during the past year?” in question 87. The responses were placed into various income ranges and the distribution is shown in Table 3.5. The level of income, like health status,

Income Range	Frequency	Percentage
\$15,000 or less	1146	14.44%
\$15,001 - \$20,000	712	8.97%
\$20,001 - \$25,000	621	7.83%
\$25,001 - \$30,000	690	8.69%
\$30,001 - \$35,000	526	6.63%
\$35,001 - \$45,000	778	9.80%
\$45,001 - \$55,000	637	8.03%
\$55,001 - \$75,000	753	9.49%
\$75,001 - \$100,000	445	5.61%
More than \$100,000	243	3.06%
Refused	789	9.99%
Don't know/not sure	596	7.51%
Total	7936	100.00%

Table 3.5: Distribution of respondent's according to annual household income.

is not equal for all respondents. The greatest percentage of respondent's belonging in any one category is in the range of income less than or equal to \$15,000 annually. Most studies into the determinants of health have shown that poorer health is associated with lower income levels. This study aims to determine whether such a relationship holds for the portion of Newfoundland from which the data has been acquired.

The final financial/work-related variable considered is education. This variable is categorized as work-related since education can be viewed as an investment in human capital. The decision whether to pursue higher levels of education is typically by work-

related motives. The highest level of education completed by respondents varies considerably, as shown in Table 3.6. About half of the respondents had completed high school or less. A small proportion of respondents had completed postgraduate studies. Lower education has been associated with poorer health in numerous studies. This may be due to natural selection or income-related consequences of education. The econometric analysis undertaken here includes education as a causal variable to determine whether there is an actual relationship between education and health and also to allow to control for the influence of education.

Highest Level of Education	Frequency	Percentage
Less than or some high school	2530	31.88%
Graduated high school	1938	24.42%
Some university/community college	862	10.86%
Graduated community/technical college	1418	17.87%
Graduated university	858	10.81%
Postgraduate	318	4.01%
Refused	12	0.15%
Total	7936	100.00%

Table 3.6: Distribution of respondents according to highest level of education completed.

3.5. *Lifestyle Factors*

The first of the lifestyle factors included in this analysis is changes a person has made in an attempt to improve their health. Question 5 asked, “In the past 12 months, did you do anything to improve your health? (For example, lost weight, quit smoking, increased exercise” More than half, 55.5%, of respondents that they did not do anything to improve their health. A considerable proportion, 44.4%, indicated that they did do something to improve their health. Responding “yes” to this question may be associated with better health status if the action the person took was perceived to have improved his health. However, a

positive response may be associated with poorer health since attempting to improve one's health may be an action taken only by those with poor health.

Drinking and smoking are also commonly known determinants of health status. Questions 24a and 33 are indicators of smoking and drinking, respectively. More than half of respondents indicated that they currently were, or were at one time, regular smokers. Over 60% indicated that they drank alcoholic beverages.

The final variable that is classified as a lifestyle factor here is the level of stress people feel in their daily lives. Question 56 asked, "Thinking about the amount of stress in your life, would you say that most days are...". The distribution of responses is shown in Table 3.7. Most people indicated that they do not experience high levels of stress most days. Only about 3% felt that most days were extremely stressful.

Daily Stress	Frequency	Percentage
Not at all stressful	1290	16.26%
Not very stressful	1917	24.16%
A bit stressful	3517	44.32%
Quite a bit stressful	966	12.17%
Extremely stressful	234	2.95%
Refused	2	0.03%
Don't know/not sure	10	0.13%
Total	7936	100.00%

Table 3.7: Distribution of respondents according to amount of stress felt most days.

3.6. Social Factors

In the category of social factors, there are two variables of interest. The first of these corresponds to question 22 of the survey which asks, "What is your current marital status?" The distribution of responses to this question is given in Table 3.8. The majority of respondents were married at the time of the survey. About 15% were single and had never been married.

Marital Status	Frequency	Percentage
Married	5362	67.57%
Common-law/living with partner	383	4.83%
Single (never married)	1225	15.44%
Widowed	571	7.20%
Separated/divorced	389	4.90%
Refused	6	0.08%
Total	7936	100.00%

Table 3.8: Distribution of respondents according to marital status.

The second social factor considered in this analysis corresponds to question 57. This question asked, “Do you have someone, other than your spouse or partner that you feel at ease with, can talk to about private matter and can call on for help?” Fifteen percent (1224) answered “no” to this question. While the majority, 84% (6688), indicated that they do have the support of and place trust in an individual other than their spouse or partner.

These social factors indicate the types of relationships and what kind of a support system a person has. Stronger social and emotional ties have been linked to better health. This relationship may work through the effect that supportive relationships have on how a person copes with stressful situations.¹⁰ These indicators may then be important determinants of health status.

3.7. Location

The final variables included in the analysis here are concerned with the respondent’s geographic location. There are two variables related to location that are of interest – economic zone and distance to doctor. The respondents involved in the adult health survey were residents of fourteen of the twenty economic zones in Newfoundland and Labrador. The distribution of respondents is shown in Table 3.9. Of course, the highest numbers of

¹⁰ Evans.

respondents were residents of economic zone 19, the Capital Coast. This is the most populated zone in the province.

Economic Zone	Frequency	Percentage
Zone 7	180	2.27%
Zone 8	994	12.53%
Zone 9	551	6.94%
Zone 10	262	3.30%
Zone 11	291	3.67%
Zone 12	581	7.32%
Zone 13	151	1.90%
Zone 14	965	12.16%
Zone 15	541	6.82%
Zone 16	415	5.23%
Zone 17	722	9.10%
Zone 18	146	1.84%
Zone 19	2028	25.55%
Zone 20	109	1.37%
Total	7936	100.00%

Table 3.9: Distribution of respondents according to economic zone of residence.

The final variable considered in this appraisal of the determinants of health corresponds to survey question 43. The question asked respondents, “Approximately how many miles or km do you normally have to travel to see the doctor or nurse?” The data used here are given in kilometers. The average distance is about seventy kilometres.

4. *Econometric Analysis*

Various studies of the determinants of health status have already been explored in Section 2. Focus now moves to applying some of the techniques used in previous studies to the data compiled through the Newfoundland Adult Health Survey (described in Section 3). Through the use of econometric and statistical analysis we will see if there is a relationship between the social and economic variables discussed in Section 3 and the health status of individuals in the health regions surveyed. If there are statistically significant relationships between the factors and health then we will also explore the strength of such relationships.

4.1. *General Procedure – Ordered Logistic Regression*

The method used in this paper to determine if a relationship exists between the various factors previously mentioned and health is ordered logistic regression, or ordered logit. This procedure is well-suited to the analysis undertaken here. In ordered logistic regression, the dependent variable has multiple outcomes that are coded in such a way as to reflect a particular ranking. In this analysis, the dependent variable is self-assessed health status (SAH). There are five outcomes for SAH: 1 poor, 2 fair, 3 good, 4 very good, and 5 excellent. The ordered logistic regression was used here rather than multinomial logit or probit as the latter techniques ignore the additional information contained in the order of the dependent variable outcomes. Ordinary least squares would also be an inappropriate method as the outcomes reflect a ranking and the difference between 1 and 2 cannot be considered the same as the difference between 2 and 3. All that the order indicates is that 2 is better than 1 and 3 is better than 2. Ordered probit could be used instead of ordered logit, however, in practice, ordered logit and ordered probit yield very similar results. The difference between

the two formulations is that ordered probit assumes that the error terms are normally distributed whereas ordered logit assumes that the errors have a logistic distribution.

The ordered logistic regression is a maximum likelihood estimation procedure in which the likelihood function of the data is maximized. The likelihood ratio test (a chi-square test) can then be used to determine the relative benefit of one model over another nested model. This procedure of evaluating alternative models is used in Section 4.2.2.

In ordered logit, the likelihood of a particular outcome given some set of characteristics (independent variables) is estimated. “An underlying score is estimated as a linear function of the independent variables and a set of cut points.”¹¹ The probability of observing a particular outcome corresponds to the probability that the estimated linear function, plus random error, is within the range of cut points estimated for that outcome. If we consider a particular outcome i , for example, ordered logit would estimate the probability of this outcome as:

$$\Pr(\text{outcome}_j = i) = \Pr(\kappa_{i-1} < \beta_1 x_{1j} + \beta_2 x_{2j} + \dots + \beta_k x_{kj} + u_j \leq \kappa_i) \quad (4.1.1)$$

Where $\beta_1, \beta_2, \dots, \beta_k$ are estimated coefficients, $\kappa_1, \kappa_2, \dots, \kappa_{I-1}$ are estimated cut points with I equal to the number of possible outcomes and u_j represents unobservable error with an assumed logistic distribution. κ_0 is taken as $-\infty$ and κ_I as $+\infty$. A constant term is not present in the estimation however, the cut points serve a purpose similar to the constant term in binomial logistic regression and other regression procedures.

The results of ordered logit can allow us to predict the probability of a person having a particular SAH given their responses to the survey questions that are used as variables in the ordered logistic regression procedure. The regression procedures used for the analysis of

¹¹ STATA Reference Manual, p.453.

the determinants of health care in the four health regions are discussed in the following two sections.

4.1.1. *Pairwise Regression*

To begin the econometric analysis, ordered logit was performed on the dependent variable, SAH, and each of the independent variables individually. This procedure allows us to examine the effects of each independent variable on health status without controlling for other factors.

As an illustration of what was done for pairwise regression, we examine how the procedure was performed in the case of one of the independent variables, sex. The estimation procedure regressed SAH against sex and estimated a coefficient corresponding to female. Male was considered the reference category and was thus omitted. Four cut points were also estimated. These are denoted as cut1, cut2, cut3, and cut4.

The probabilities of each health outcome or SAH category is calculated using the estimated coefficient and cut points as follows:

$$\begin{aligned} \Pr(\text{SAH} = \text{poor}) &= \Pr(\beta_F F + u < \text{cut1}) \\ &= \Pr(u < \text{cut1} - \beta_F F) \\ &= \frac{1}{1 + e^{-(\text{cut1} - \beta_F F)}} \end{aligned} \quad (4.1.1)$$

$$\begin{aligned} \Pr(\text{SAH} = \text{fair}) &= \Pr(\text{cut1} < \beta_F F + u < \text{cut2}) \\ &= \Pr(u < \text{cut2} - \beta_F F) - \Pr(u < \text{cut1} - \beta_F F) \\ &= \frac{1}{1 + e^{-(\text{cut2} - \beta_F F)}} - \frac{1}{1 + e^{-(\text{cut1} - \beta_F F)}} \end{aligned} \quad (4.1.2)$$

$$\begin{aligned} \Pr(\text{SAH} = \text{good}) &= \Pr(\text{cut2} < \beta_F F + u < \text{cut3}) \\ &= \Pr(u < \text{cut3} - \beta_F F) - \Pr(u < \text{cut2} - \beta_F F) \\ &= \frac{1}{1 + e^{-(\text{cut3} - \beta_F F)}} - \frac{1}{1 + e^{-(\text{cut2} - \beta_F F)}} \end{aligned} \quad (4.1.3)$$

$$\begin{aligned}
\Pr(\text{SAH} = \text{very good}) &= \Pr(\text{cut3} < \beta_F F + u < \text{cut4}) \\
&= \Pr(u < \text{cut4} - \beta_F F) - \Pr(u < \text{cut3} - \beta_F F) \\
&= \frac{1}{1 + e^{-(\text{cut4} - \beta_F F)}} - \frac{1}{1 + e^{-(\text{cut3} - \beta_F F)}}
\end{aligned} \tag{4.1.4}$$

$$\begin{aligned}
\Pr(\text{SAH} = \text{excellent}) &= \Pr(\text{cut4} < \beta_F F + u) \\
&= 1 - \Pr(u < \text{cut4} - \beta_F F) \\
&= 1 - \frac{1}{1 + e^{-(\text{cut4} - \beta_F F)}}
\end{aligned} \tag{4.1.5}$$

where β_F is the estimated coefficient associated with female and F is equal to 1 when the person is female and 0 when the person is male. These calculated probabilities allow us to compare the probability that a person indicates a particular SAH given that person's gender. The same calculation of probabilities was performed for all pairwise regressions. For independent variables with more than two categories, dummy variables for each category (except the reference category) were created within STATA's ologit procedure. For an independent categorical variable with n categories, $n-1$ coefficients were estimated in the ordered logistic regression. The pairwise ordered logistic regression results as well as associated probabilities are given in section 4.2.1.

4.1.2. Multiple Independent Variables

The procedure used in the pairwise regressions (Section 4.1.1.) is nearly identical to the procedure used to estimate the probabilities with more than one independent variable at a time. Including more than one independent variable in ordered logistic regression allows us to control for the influence of more than one characteristic on the probability of SAH outcomes. To illustrate the difference between the pairwise regressions and the treatment in this section, we will look at an ordered logistic regression involving sex (as in Section 4.1.1.) and whether a person ever smoked regularly (Question 24a) as

independent variables. In this example, the predicted probabilities of the SAH outcomes would be:

$$\begin{aligned}
 \Pr(\text{SAH} = \text{poor}) &= \Pr(\beta_F F + \beta_S S + u < \text{cut1}) \\
 &= \Pr(u < \text{cut1} - \beta_F F - \beta_S S) \\
 &= \frac{1}{1 + e^{-(\text{cut1} - \beta_F F - \beta_S S)}}
 \end{aligned} \tag{4.1.7}$$

$$\begin{aligned}
 \Pr(\text{SAH} = \text{fair}) &= \Pr(\text{cut1} < \beta_F F + \beta_S S + u < \text{cut2}) \\
 &= \Pr(u < \text{cut2} - \beta_F F - \beta_S S) - \Pr(u < \text{cut1} - \beta_F F - \beta_S S) \\
 &= \frac{1}{1 + e^{-(\text{cut2} - \beta_F F - \beta_S S)}} - \frac{1}{1 + e^{-(\text{cut1} - \beta_F F - \beta_S S)}}
 \end{aligned} \tag{4.1.8}$$

$$\begin{aligned}
 \Pr(\text{SAH} = \text{good}) &= \Pr(\text{cut2} < \beta_F F + \beta_S S + u < \text{cut3}) \\
 &= \Pr(u < \text{cut3} - \beta_F F - \beta_S S) - \Pr(u < \text{cut2} - \beta_F F - \beta_S S) \\
 &= \frac{1}{1 + e^{-(\text{cut3} - \beta_F F - \beta_S S)}} - \frac{1}{1 + e^{-(\text{cut2} - \beta_F F - \beta_S S)}}
 \end{aligned} \tag{4.1.9}$$

$$\begin{aligned}
 \Pr(\text{SAH} = \text{very good}) &= \Pr(\text{cut3} < \beta_F F + \beta_S S + u < \text{cut4}) \\
 &= \Pr(u < \text{cut4} - \beta_F F - \beta_S S) - \Pr(u < \text{cut3} - \beta_F F - \beta_S S) \\
 &= \frac{1}{1 + e^{-(\text{cut4} - \beta_F F - \beta_S S)}} - \frac{1}{1 + e^{-(\text{cut3} - \beta_F F - \beta_S S)}}
 \end{aligned} \tag{4.1.10}$$

$$\begin{aligned}
 \Pr(\text{SAH} = \text{excellent}) &= \Pr(\text{cut4} < \beta_F F + \beta_S S + u) \\
 &= 1 - \Pr(u < \text{cut4} - \beta_F F - \beta_S S) \\
 &= 1 - \frac{1}{1 + e^{-(\text{cut4} - \beta_F F - \beta_S S)}}
 \end{aligned} \tag{4.1.11}$$

where β_S is the estimated coefficient on the independent variable for smoking and S is equal to 1 when the respondent being considered smoked regularly and 0 when the person is, and was always, a non-smoker.

Equations (4.1.7), (4.1.8), (4.1.9), (4.1.10) and (4.1.11) are analogous to equations (4.1.1), (4.1.2), (4.1.3), (4.1.4) and (4.1.5), respectively. These formulas for predicting probabilities can be expanded when more independent variables are included in the regression. The actual results of the regression, as well as the predicted probabilities are given and discussed in Section 4.2.

4.2. Ordered Logistic Regression Results

4.2.1. Pairwise Regression

In performing the pairwise ordered logistic regressions, observations with responses such as “don’t know/not sure,” and “refused” were included in the analysis. The dummy variables corresponding to these responses were statistically insignificant in the case of all variables. For the pairwise regressions including such responses in the analysis was considered appropriate as the objective of the pairwise regressions was simply to get a feel for possible relationships between SAH and the independent variables and look at whether the effects of particular factors on health status are negative or positive.

The STATA output corresponding to each of the pairwise regressions is given in Appendix B. The predicted probabilities of each SAH outcome are given in Tables 4.1 through 4.4.

	PROBABILITY OF SAH OUTCOME				
	POOR	FAIR	GOOD	VERY GOOD	EXCELLENT
Characteristic	Sex				
Female	0.031097	0.121316	0.272025	0.453372	0.12219
Male	0.031665	0.123179	0.274168	0.450789	0.120198
Characteristic	Age				
18 - 24 years	0.021689	0.090049	0.231948	0.494803	0.161511
25 - 34 years	0.028469	0.112863	0.262224	0.465045	0.131399
35 - 44 years	0.031114	0.1217	0.272964	0.452926	0.121295
45 - 54 years	0.04073	0.151836	0.302479	0.410422	0.094534
55 - 64 years	0.046597	0.168794	0.314789	0.386661	0.083159
65+ years	0.050002	0.178179	0.320412	0.373727	0.0776709

Table 4.1: Probability of SAH outcomes dependent upon demographic (sex and age) characteristics.

	PROBABILITY OF SAH OUTCOME				
	POOR	FAIR	GOOD	VERY GOOD	EXCELLENT
Characteristic	Question 5: Do anything to improve health status?				
No	0.029897	0.117423	0.267512	0.458755	0.126414
Yes	0.032606	0.126318	0.27779	0.446432	0.116854
Characteristic	Ever smoke regularly?				
No	0.024173	0.098334	0.244148	0.485221	0.148124
Yes	0.036653	0.139915	0.294099	0.427637	0.101696
Characteristic	Ever drink alcoholic beverages?				
No	0.039363	0.148647	0.301241	0.4148	0.09595
Yes	0.027239	0.109373	0.259013	0.469944	0.134431
Characteristic	Amount of stress in daily life				
Not at all stressful	0.019756	0.083571	0.223349	0.503757	0.169567
Not very stressful	0.023661	0.098042	0.246749	0.486389	0.145158
A bit stressful	0.033265	0.131133	0.288659	0.440122	0.106822
Quite a bit stressful	0.046201	0.170688	0.321447	0.383359	0.078305
Extremely stressful	0.073843	0.239285	0.344333	0.293458	0.049081

Table 4.2: Probability of SAH outcomes dependent upon lifestyle characteristics.

	PROBABILITY OF SAH OUTCOME				
	POOR	FAIR	GOOD	VERY GOOD	EXCELLENT
Characteristic	Marital Status				
Single (never married)	0.026328	0.105528	0.252512	0.474784	0.140847
Married	0.031648	0.123454	0.274976	0.450486	0.119436
Common-law/Living with partner	0.029575	0.116586	0.266873	0.45998	0.126985
Widowed	0.034992	0.134217	0.286497	0.43536	0.108934
Divorced or Separated	0.040634	0.151548	0.302244	0.410829	0.094746
Characteristic	Have someone other than spouse/partner for support?				
No	0.040052	0.149808	0.301077	0.413291	0.095771
Yes	0.029857	0.117528	0.268279	0.458773	0.125563

Table 4.3: Probability of SAH outcomes dependent upon social characteristics.

	PROBABILITY OF SAH OUTCOME				
	POOR	FAIR	GOOD	VERY GOOD	EXCELLENT
Characteristic	Employment Status				
<i>Not employed for pay</i>	0.042962	0.161608	0.316497	0.396164	0.082769
<i>Self-employed</i>	0.01266	0.054523	0.161248	0.514665	0.256904
<i>Employed by corporation or organization</i>	0.013048	0.056075	0.164737	0.515025	0.251115
<i>Retired</i>	0.023045	0.093948	0.23561	0.489199	0.158198
<i>Student</i>	0.010284	0.044861	0.138348	0.507461	0.299047
Characteristic	Financial Security				
<i>Very good</i>	0.043671	0.162825	0.315119	0.395107	0.083277
<i>Good</i>	0.055171	0.194507	0.332656	0.351336	0.06633
<i>Satisfactory</i>	0.072758	0.236234	0.343011	0.297785	0.050213
<i>Just getting by</i>	0.11555	0.311225	0.330477	0.211973	0.030775
<i>Can't cope</i>	0.22974	0.39985	0.247283	0.109409	0.013718
Characteristic	Total Annual Household Income				
<i>\$15,000 or less</i>	0.053609	0.191712	0.33416	0.354617	0.065902
<i>\$15,001 - \$20,000</i>	0.043888	0.164605	0.319062	0.392353	0.080092
<i>\$20,001 - \$25,000</i>	0.032934	0.130546	0.28961	0.441882	0.105027
<i>\$25,001 - \$30,000</i>	0.028603	0.115945	0.272805	0.463141	0.119506
<i>\$30,001 - \$35,000</i>	0.031119	0.12451	0.282995	0.450716	0.11066
<i>\$35,001 - \$45,000</i>	0.023668	0.098456	0.248836	0.487513	0.141527
<i>\$45,001 - \$55,000</i>	0.022699	0.09491	0.243419	0.492164	0.146808
<i>\$55,001 - \$75,000</i>	0.020847	0.088029	0.232333	0.500747	0.158044
<i>\$75,001 - \$100,000</i>	0.015808	0.068586	0.196562	0.519803	0.199242
<i>More than \$100,000</i>	0.012584	0.055566	0.168508	0.524616	0.238726
Characteristic	Education				
<i>Less than high school</i>	0.043671	0.162825	0.315119	0.395107	0.083277
<i>High school diploma</i>	0.029767	0.119051	0.274002	0.458073	0.119107
<i>Some university</i>	0.025528	0.104367	0.254911	0.478488	0.136706
<i>Graduated technical/community college</i>	0.026325	0.10718	0.258802	0.47467	0.133023
<i>University degree</i>	0.018923	0.080108	0.216292	0.507672	0.177006
<i>Postgraduate</i>	0.016423	0.070461	0.198156	0.515955	0.199006

Table 4.4: Probability of SAH outcomes dependent upon financial/work-related factors.

From the estimated coefficients in these regressions we see that there is a positive relationship between better health and education and income. Higher levels of education and higher income levels (for the most part) results in higher probabilities of better SAH outcomes. A “yes” response to question 5 (change to improve health) is also associated with a higher probability of better health. The pairwise regression including sex as the independent variable indicates that females have a higher probability of a higher SAH than do males.

The results also imply that smoking has a negative effect on health status. However, drinking alcoholic beverages is shown as having a positive effect on health. Since the question related to drinking did not distinguish between levels of consumption, the results of the pairwise regression should be interpreted with caution. Occasional to moderate consumption of alcoholic beverages has been shown to be beneficial, or at least not detrimental to health. However, excessive drinking has been proven to negatively affect health.

Employment status also affects health status. The ordered logit results involving employment status indicate that not being unemployed increases the probability of better health. The highest probability of “excellent” SAH is associated with being a student. However, since students are typically relatively young (18 – 24 years of age), this result is likely showing the influence that age has on health status as well as the impact of employment status. Other than being unemployed, being retired resulted in the lowest probability of excellent SAH. This too is likely partially due to the influence of age. Income may also play a role here as most retired people have fixed incomes and probably feel stress about financial security.

The effects of marital status on health status resulting from the pairwise regression indicated that widowed and separated or divorced persons have a lower probability of better health than single people. The estimated coefficient for “common-law/living with partner” was not statistically significant. The results also indicated that married people are more probable to be in worse health than single people. This result is not as expected however, the result may be altered when more than one independent variable is used in estimating the relationship.

When question 43, distance to doctor or nurse, was included as the independent variable, the coefficient was statistically insignificant. Using zone as the independent variable resulted in the same. Zone 19 was the only zone with a Wald statistic that indicated that the coefficient was significant at the five percent level. Zone 19 contains the capital city of Newfoundland and is the most highly populated zone in the province. The sign on the coefficient indicates that living in Zone 19 increases the probability of better health. A pairwise regression using a dummy variable equal to 1 if the respondent is from Zone 19 and 0 otherwise also indicated a statistically significant coefficient for being a resident of Zone 19. Again, this coefficient was positive.

How stressful a person considers most days to be was also shown to affect SAH. People who find most days “extremely stressful” have the highest probability of “poor” SAH. How secure or comfortable a person feels with their given financial circumstances also influences the probability of SAH outcomes. People who “can’t cope” given their financial circumstances have the lowest probability of “excellent” SAH and the highest probability of “poor” SAH.

4.2.2. *Multiple Independent Variables*

When more than one independent variable was used in estimation of the relationship between health and the various factors all observations where there was “no response” or “don’t know/not sure” were dropped from these regressions. The estimation package, STATA, would simply omit such observations from the regression if such responses had not been coded as they were in the data set used.¹² Dropping these responses resulted in a final sample size of 6,497. This is still a reasonable size sample.

The analysis using multiple independent variables in ordered logit estimation began with a very general model including all independent variables discussed above. This first model shall be referred to as the “unrestricted model”. There were fourteen independent variables included in this model. This resulted in the estimation of fifty-four coefficients due to the categorical nature of many of the independent variables. The categorical variables included: zone, income, employment status (emp), age (q86d), financial security (q83), stress in daily life (q56), education (q23a), and marital status (q22). The distance to doctor or nurse was a continuous variable. All other variables were dichotomous.

The log of the likelihood function for the estimated unrestricted model was -8208.8428 and the pseudo R^2 was 0.0536. The pseudo R^2 seems low however in logistic regression a low value is not uncommon. The individual Wald statistics (z) and corresponding p-values for the independent variables indicated that a number of the variables were not statistically significant at the five percent level. This was the criterion used to determine which variables to drop from the model until a satisfactory final model was estimated. There selection criterion was not absolute however, as intuition and prevailing

¹² In the case of most variables, “no response” was coded as 8 and “don’t know/not sure” as 9.

theory dictate that some variables are necessary despite what the Wald statistics may indicate.

After estimating the unrestricted model, the first variables dropped were zone and distance to doctor or nurse (q43). This second estimation was not satisfactory and so other variables were omitted from the model and a number of subsequent regressions performed. Table 4.5 outlines the steps of model selection¹³. After particular variables were omitted from the unrestricted model a likelihood ratio (LR) test was performed to determine whether dropping the variables was an appropriate action. The null hypothesis for the LR test was that the coefficients of the dropped variables were equal to zero. Where the p-value of the chi-square test statistic was greater than 0.05, the null hypothesis was accepted and the dropped variables were not put back into the model.

The final model dropped the variables zone, q57, q43 and q5 from the unrestricted model. In the estimated final model not all categories of all variables were statistically significant at the 5% level. The second income category (\$15,001 - \$20,000), the fourth employment status corresponding to “retired”, the 25 to 34, 35 to 44, and 45 to 54 years of age categories, the education levels lower than graduation from university, and married, common law, and separated/divorces categories of marital status had individual Wald statistics indicating that they were not statistically significant at the 5% level. However, completely omitting the associated variables from the model was not appropriate as the likelihood ratios in Table 4.5 indicate. Intuition and prevailing theory on health determinants also dictate that these particular variables be included in the model. The estimated coefficients and p-values for the final model are shown in Table 4.6.

¹³ Regression output is given in Appendix B.

Model	Variables Included	Variables Dropped	Log likelihood	Pseudo R ²	LR test: χ^2	LR test: Prob > χ^2
Unrestricted	zone, smoke, income, emp, q86d, q83, q57, q56, q43, q33, q23a, q22, q20, q5	---	-8208.8425	0.0536	---	---
Model 1	smoke, income, emp, q86d, q83, q57, q56, q33, q23a, q22, q20, q5	zone, q43	-8214.5988	0.0529	11.51	0.6454
Model 2	smoke, income, emp, q86d, q83, q57, q56, q33, q23a, q22, q20	zone, q43, q5	-8216.4839	0.0527	15.28	0.4312
Model 3	smoke, income, emp, q86d, q83, q57, q56, q33, q23a, q20	zone, q43, q5, q22	-8224.8805	0.0517	32.08	0.0306
Model 4	smoke, income, emp, q83, q57, q56, q33, q23a, q22, q20	zone, q43, q5, q86d	-8234.3287	0.0506	50.97	0.0002
Model 5	Smoke, income, emp, q86d, q83, q56, q33, q23a, q22, q20	Zone, q43, q5, q57	-8218.4157	0.0525	19.15	0.2611
Final Model	smoke, income, emp, q86d, q83, q56, q33, q23a, q22, q20	zone, q57, q43, q5	-8218.4157	0.0525	19.15	0.2611

Table 4.5: Outline of models estimated and likelihood ratio tests.

STATA label	Variable	Estimated Coefficient	p-value
income reference	\$15,000 or less	---	---
_lincome_2	\$15,001 to \$20,000	0.1016279	0.274
_lincome_3	\$20,001 to \$25,000	0.3734737	0.000
_lincome_4	\$25,001 to \$30,000	0.3682789	0.000
_lincome_5	\$30,001 to \$35,000	0.22211	0.045
_lincome_6	\$35,001 to \$45,000	0.4484088	0.000
_lincome_7	\$45,001 to \$55,000	0.467121	0.000
_lincome_8	\$55,001 to \$75,000	0.4506697	0.000
_lincome_9	\$75,001 to \$100,000	0.6568796	0.000
_lincome_10	More than \$100,000	0.7853085	0.000
emp reference	Not employed for pay	---	---
_lemp_1	Self-employed	0.4641974	0.000
_lemp_2	Employed by company/organization	0.4008154	0.000
_lemp_4	Retired	0.0629036	0.511
_lemp_5	Student	0.6266253	0.000
Age reference	18 – 24 years	---	---
_lq86d_2	25 – 34 years	0.1763513	0.177
_lq86d_3	35 – 44 years	0.0501964	0.701
_lq86d_4	45 – 54 years	-0.2053835	0.125
_lq86d_5	55 – 64 years	-0.3036433	0.038
_lq86d_6	65+ years	-0.3805574	0.025
Financial security reference	Very good	---	---
_lq83_2	Good	-0.2826052	0.000
_lq83_3	Satisfactory	-0.4549863	0.000
_lq83_4	Just getting by	-0.6181056	0.000
_lq83_5	Can't cope	-1.140598	0.000
Daily stress reference	Not at all stressful	---	---
_lq56_2	Not very stressful	-0.3348904	0.000
_lq56_3	A bit stressful	-0.7441713	0.000
_lq56_4	Quite a bit stressful	-1.162446	0.000
_lq56_5	Extremely stressful	-1.654763	0.000
Education reference	Less than or some high school	---	---
_lq23a_2	Graduated high school	0.0809798	0.227
_lq23a_3	Some university/community college	0.1655659	0.069
_lq23a_4	Graduated community/technical college	0.0928094	0.229
_lq23a_5	Graduated university	0.2993923	0.002
_lq23a_6	Postgraduate	0.4607953	0.000
Marital status reference	Single, never married	---	---
_lq22_1	Married	0.0081148	0.925
_lq22_2	Common-law	0.415587	0.737
_lq22_5	Widowed	0.4335875	0.001
_lq22_6	Separated or divorced	0.2167619	0.096
q33	Drinks alcoholic beverages	0.1814365	0.001
q20	Female	0.2038037	0.000
Smoke	Smoked/smokes regularly	-0.3296562	0.000
_cut1		-3.868348	
_cut2		-2.023474	
_cut3		-0.4753291	
_cut4		1.978932	

Table 4.6: Final model estimation.

In Section 2, the effects of various factors on health status were discussed. The estimated final model largely supports the ideas about how various characteristics determine

health. The estimated coefficient for smoking indicates, as expected, that smoking decreases the probability of better SAH. Employed persons are more likely to have better health than those not employed for pay. Older people have higher probability of lower SAH. More stress in a person's daily life increases the probability of lower SAH as do higher levels of financial insecurity. Higher education is associated with higher probabilities of better SAH outcomes. Females also have higher probabilities of better SAH outcomes. Finally, the estimated coefficients for marital status are somewhat unsatisfactory. The coefficients for married, common-law, and divorced or separated were not statistically significant. The coefficients for widowed and separated or divorced persons were both positive indicating that people in this category would have a higher probability of better SAH than single people with the same characteristics. One may expect the reverse to be true as death of a spouse or separation or divorce would likely increase the amount of stress in a person's life. However, having been married at one time may indicate that a person has had serious relationships and thus have developed coping skills. Also, the presence of children could be another factor.

The differences in probabilities are presented in Table 4.7¹⁴. The predicted probabilities shown in the table were calculated using the reference categories for all independent variables other than the characteristic for which the probability is being calculated. A hypothetical person with characteristics falling in all reference categories would be a single male with total annual household income of \$15,000 or less, not employed for pay, 18 to 24 years of age, who feels very good about his financial circumstances, feel not at all stressed most days, have less than a high school diploma, and would not drink or smoke. If the same person were to have an income of \$30,001 to \$35,000 annually than the probability of him having an excellent SAH would increase by 2.5756 percentage points

¹⁴ See page 34.

(from 12.1433% to 14.7189%). If the reference person were to smoke then the probability of a poor SAH outcome would increase from 2.0465% to 2.8231% - an increase of 0.7766 percentage points. Similar comparisons can be made for all variables. There are numerous combinations of the independent variables. However, the values of all possible predicted probabilities are not included in this paper because of space limitations and to avoid cluttering the analysis.

	Variable	Probability of SAH Outcomes				
		Pr(poor)	Pr(fair)	Pr(good)	Pr(very good)	Pr(excellent)
Total Annual Household Income	\$15,000 or less	0.020465	0.096295	0.266595	0.495212	0.121433
	\$15,001 to \$20,000	0.018524	0.088157	0.252952	0.507668	0.132699
	\$20,001 to \$25,000	0.014177	0.069228	0.216278	0.533096	0.167220
	\$25,001 to \$30,000	0.014250	0.069553	0.216972	0.532727	0.166498
	\$30,001 to \$35,000	0.016456	0.079275	0.236649	0.520431	0.147189
	\$35,001 to \$45,000	0.013167	0.064686	0.206344	0.537886	0.177917
	\$45,001 to \$55,000	0.012926	0.063594	0.203886	0.538924	0.180671
	\$55,001 to \$75,000	0.013138	0.064553	0.206046	0.538015	0.178248
	\$75,001 to \$100,000	0.010716	0.053426	0.179611	0.545769	0.210477
More than \$100,000	0.009437	0.047415	0.164013	0.546524	0.232612	
Employment Status	Not employed for pay	0.020465	0.096295	0.266595	0.495212	0.121433
	Self-employed	0.012964	0.063763	0.204269	0.538766	0.180238
	Employed by company	0.013800	0.067539	0.212638	0.534960	0.171062
	Retired	0.019242	0.091186	0.258171	0.503096	0.128305
	Student	0.011042	0.054941	0.183391	0.545133	0.205494
Age	18 – 24 years	0.020465	0.096295	0.266595	0.495212	0.121433
	25 – 34 years	0.017213	0.082553	0.242845	0.515852	0.141537
	35 – 44 years	0.019483	0.092200	0.259879	0.501548	0.126891
	45 – 54 years	0.025015	0.114649	0.293257	0.465912	0.101168
	55 – 64 years	0.027526	0.124367	0.305291	0.450240	0.092576
	65+ years	0.029662	0.132407	0.314256	0.437361	0.086314
Financial Security	Very good	0.020465	0.096295	0.266595	0.495212	0.121433
	Good	0.026969	0.122234	0.302765	0.453673	0.094359
	Satisfactory	0.031880	0.140552	0.322482	0.424463	0.080623
	Just getting by	0.037318	0.159647	0.338668	0.395037	0.069329
	Can't cope	0.061356	0.231227	0.367861	0.297248	0.042309
Stress in Daily Life	Not at all stressful	0.020465	0.096295	0.266595	0.495212	0.121433
	Not very stressful	0.028375	0.127587	0.308986	0.445067	0.089985
	A bit stressful	0.042121	0.175548	0.349140	0.371568	0.061624
	Quite a bit stressful	0.062626	0.234499	0.368201	0.293242	0.041432
	Extremely stressful	0.098537	0.310315	0.355993	0.209416	0.025738
Education	Less than or some high school	0.020465	0.096295	0.266595	0.495212	0.121433
	High School diploma	0.018903	0.089761	0.255737	0.505258	0.130340
	Some university/college	0.017397	0.083342	0.244305	0.514724	0.140232
	Community/tech college diploma	0.018685	0.088839	0.254142	0.506646	0.131687
	University degree	0.015251	0.073996	0.226212	0.527385	0.157156
	Postgraduate	0.013007	0.063961	0.204716	0.538580	0.179736
Marital Status	Single-never married	0.020465	0.096295	0.266595	0.495212	0.121433
	Married	0.020303	0.095623	0.265513	0.496260	0.122301
	Common-law	0.013601	0.066641	0.210679	0.535912	0.173167
	Widowed	0.013361	0.065562	0.208298	0.537019	0.175760
	Separated/divorced	0.016543	0.079652	0.237373	0.519913	0.146519
Drinks?	Drinks alcoholic	0.017128	0.082183	0.242156	0.516377	0.142156
Sex	Female	0.016755	0.080572	0.239127	0.518639	0.144906
Smokes?	Smoked regularly	0.028231	0.127043	0.308372	0.445940	0.090414

Table 4.7: Predicted probabilities of SAH outcomes.

5. *Summary & Conclusions*

It is important to keep in mind that an alternative relationship between self-assessed health status and the independent variables included in this paper may exist. It may be that SAH is a factor affecting labour force variables such as unemployment and low income. SAH may also be a partial cause of stress. Using SAH as an independent variable instead of the dependent variable would be a valid approach to analysis. However, exploring this relationship would require a more lengthy analysis than is intended through this paper.

The pairwise regressions and the final model estimated in the previous section are helpful in identifying factors that affect self-assessed health status in the regions surveyed by the 2001 Newfoundland Adult Health Survey. The estimated coefficients and predicted probabilities of the multivariate logistic regression allow us to see the effects of particular variables on health while controlling for other factors at the same time. The final model estimated in Section 4 indicates that the health indicator, SAH is affected by ten different factors – education, income, employment status, smoking, drinking, marital status, stress, financial security, age and sex. This result is consistent with prevailing theory on the subject of health determinants.

The idea set out in the Whitehall Studies and the Black Report about a person's independence and autonomy in the workplace is evident in the estimated final model here. In the model, self-employed people have higher predicted probabilities of better health outcomes than people who are employed by an organization or company. Self-employed people would most likely have much more control over their work than would people who work for others. The independence seems to be beneficial to health status.

In the regions surveyed, the effects of age and sex also agree with previous studies. Females have higher probabilities of better health outcomes than men. The probability of an excellent SAH also decreases with age, for the most part. The effects of income in the region are also expected. With increases in income, holding all else constant, the probability of better health outcomes increases. The same holds for educational attainment. University education is associated with better SAH outcomes in this investigation.

The effects of marital status on health outcomes are less straightforward. The final model indicates that married people have a lower probability of an excellent SAH than do people who are widowed, living common law, or divorced or separated. This result is not really what we would expect. However, only widowed category has a Wald statistic that indicates statistical significance. The increase in the probability of better SAH when widowed could be due to the influence of adult children or due to coping mechanisms developed while married. An improvement in health would be expected with marriage or close relationships as such indicate strong social support systems. Unfortunately, the results of this analysis do not allow for any strong conclusions about the influence of marital status on health.

Perhaps the two most interesting determinants of health considered in this study are the amount of stress people have in their daily lives and how financially secure people consider themselves. The marginal impact of these variables on the probability of SAH outcomes is greater than all of the other variables considered. The probability of an “excellent” SAH outcome when the respondent felt like they couldn’t cope given their financial circumstances or felt extremely stressed most days was the lowest. The estimated coefficients associated with these variables were relatively large. For the most extreme

categories in each variable, “extremely stressful” and “can’t cope”, the estimated coefficients were -1.65 and -1.14, respectively. Such relatively large coefficients indicate that the impact of these variables is fairly strong. The presence of such direct measurements of stress and security is not common in a lot of the previous studies on health determinants. Having such variables available can give us increased insight into the direct influence that stress has on health status.

The financial security variable is very telling and leads to further questions about the influence of income and finance-related factors on health status. Even when a person’s income is high the level of comfort that a person feels may not be great. A person may feel financially insecure even if he has a seemingly large annual income. Even if he can afford all the necessities of life the feeling of financial insecurity may have a detrimental effect on his health. The results of this study support this. It is not simply absolute income that affects a person’s health. The person’s perception of their financial circumstances has a great effect on health as well.

The results of the econometric analysis here are satisfactory. The data obtained from the health regions surveyed in the 2001 Newfoundland Adult Health Survey show similar patterns in health status and the determinants of health that exist other regions. Many of the same factors that have been shown to influence health status in numerous other regions in the world have been revealed to be influential in the region investigated here as well.

In studying the determinants of health one thing is clear, reliable and complete data is essential. The data concerning various health indicators and social and economic variables need to be collected from large samples of the population one wishes to study. The Government of Newfoundland and Labrador has come a long way in attempting to collect

good health data. If such data collection continues then valuable longitudinal data will be available in time. Such data will greatly expand the type of analysis that can be performed. The collection of variables concerning stress, financial security, and income are also extremely important.

With increased analysis of health data policy makers can direct resources to the proper activities to decrease health inequalities and to increase overall health status. This study has attempted to direct action somewhat however it has some important limitations. The most important of these is that the sample used is not representative of the entire Province and so conclusions about the relative effects of social and economic variables on health status can only be made for the regions surveyed. It is likely that relationships very similar to those found here exist in the remainder of the Province. The only real concern would be extending the conclusions to Labrador. Some areas of Labrador may follow the same pattern as that found here but in very isolated regions and regions with high proportions of natives may exhibit very different patterns. No conclusions should be drawn for all of Newfoundland and Labrador.

Continued collection and analysis of health data are important to ensure proper resource allocation. Once a relatively thorough understanding of the determinants of health is had further analysis into such things as the effects of preventative medicine and health care expenditures can provide added detail to the overall analysis of health in a region.

Bibliography

- Evans, Robert. *Interpreting and Addressing Inequalities in Health: From Black to Acheson to Blair to ... ?* 7th Annual Lecture, Office of Health Economics, London, 2002.
- Jones, A. and O. O'Donnell. *Econometric Analysis of Health Data*. John Wiley & Son, Ltd., West Sussex, UK, 2002.
- Lynch, J., G. Davey Smith, M.Hillemeier, M. Shaw, T. Raghunathan, and G. Kaplan, 2001, Income inequality, the psychosocial environment, and health: comparison of wealthy nations, *Lancet*, 358, July, 194-200.
- Marmot, M. and R.G. Wilkinson. *Social Determinants of Health*. Oxford University Press, Oxford, 1999.
- Mellor, J. and J. Milyo, 2001, Is Inequality Bad for Your Health?, Working Paper Series 00.16, The Harris School, Chicago.
- Ryan-Nicholls, K., F. Racher, B. Gfellner, and R. Annis, 2000. Toward a Framework to Explore the Health of Rural Communities, Discussion Document #1, Brandon University.
- Segovia, J., A.C. Edwards, and R.F. Bartlett, 1997. Newfoundland Panel on Health and Medical Care Adult Health Survey – Methodology and Descriptive Results. Division of Community Medicine, Faculty of Medicine, Memorial University of Newfoundland, St.John's, NF.
- Smith, J.P., 1999, Health Bodies and Thick Wallets: The Dual Relation Between Health and Economic Status, *Journal of Economic Perspectives*, 13(2), 145-166.
- Stansfeld, S.A., 1999, Social support and social cohesion, in *Social Determinants of Health*, M.Marmot and R.G. Wilkinson (eds), Oxford University Press, Oxford.
- Stata Reference Manual, Release 7, 2001, Texas.
- StataCorp. 2001. Stata Statistical Software: Release 7.0. College Station, TX: Stata Corporation.
- Townsend, Peter, and Nick Davidson. *Inequalities in Health: The Black Report*. Penguin Books, England, 1982.

Appendix A

Relevant Survey Questions

Note: These questions are taken from the Newfoundland Adult Health Survey 2001 Survey Questionnaire, revised May 2, 2001.

5. In the past 12 months, did you do anything to improve your health? (For example, lost weight, quit smoking, increased exercise)
- Yes
 - No
 - Don't know/Not sure
20. You are...
- Male
 - Female
22. What is your current marital status? Are you...
- Married
 - Common-law/Living with partner
 - Single (never married)
 - Widowed
 - Separated/Divorced
 - Refused
- 23A. What is the highest level of education you have completed?
- Less than or some high school
 - Graduated high school
 - Some university/community college
 - Graduated community/technical college
 - Graduated university
 - Postgraduate
 - Refused
- 24A. Did you ever smoke cigarettes regularly?
- Yes
 - No
33. Do you drink any alcoholic beverages, that is, beer, wine or liquor?
- Yes
 - No
 - Don't know/refused
43. Approximately how many miles/km do you normally have to travel to see the doctor/nurse?
- _____ miles or _____ km

48. Would you say that your health is...
- Excellent
 - Very Good
 - Good
 - Fair
 - Poor
 - Refused
 - Don't know/Not sure
56. Thinking about the amount of stress in your life, would you say that most days are:
- Not at all stressful
 - Not very stressful
 - A bit stressful
 - Quite a bit stressful
 - Extremely stressful
57. Do you have someone, other than your spouse or partner that you feel at ease with, can talk to about private matters and can call on for help?
- Yes
 - No
 - Refused
 - Don't know/not sure
82. What is your current employment status? Are you...
- Self-employed
 - Employed by company/organization
 - Not employed for pay
 - Retired
 - Student
 - Refused
83. Given your annual household income and considering your monthly expenditures on food, car, housing, etc, would you consider your financial circumstances to be:
- Very good
 - Good
 - Satisfactory
 - Just getting by
 - Can't cope
 - Refused
 - Don't know/not sure
86. What is your date of birth?
- [Placed into categories]

87. What is the best estimate of the total income of all household members from all sources during the past year? Was the total household income...

- \$15,000 or less
- \$15,001 - \$20,000
- \$20,001 - \$25,000
- \$25,001 - \$30,000
- \$30,001 - \$35,000
- \$35,001 - \$45,000
- \$45,001 - \$55,000
- \$55,001 - \$75,000
- \$75,001 - \$100,000
- More than \$100,000
- Refused
- Don't know/not sure

Appendix B

**Regression Results
(Generated by STATA)**

PAIRWISE REGRESSION RESULTS

```
. xi: ologit health i.q5
i.q5          _Iq5_0-9          (naturally coded; _Iq5_0 omitted)
```

```
Iteration 0:  log likelihood = -10591.858
Iteration 1:  log likelihood = -10589.124
Iteration 2:  log likelihood = -10589.124
```

```
Ordered logit estimates          Number of obs   =       7936
                                LR chi2(2)        =         5.47
                                Prob > chi2         =         0.0650
Log likelihood = -10589.124      Pseudo R2       =         0.0003
```

```
-----+-----
health |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
_Iq5_1 |   .089518    .041632     2.15   0.032    .0079207   .1711153
_Iq5_9 |   .5204602   .5261312    0.99   0.323   - .5107379  1.551658
-----+-----
_cut1 |  -3.390121   .0668727                    (Ancillary parameters)
_cut2 |  -1.666261   .0361368
_cut3 |   -.25451    .029388
_cut4 |   2.022563   .0393597
-----+-----
```

```
. xi: ologit health i.q20
i.q20          _Iq20_1-2          (naturally coded; _Iq20_1 omitted)
```

```
Iteration 0:  log likelihood = -10591.858
Iteration 1:  log likelihood = -10591.756
```

```
Ordered logit estimates          Number of obs   =       7936
                                LR chi2(1)        =         0.20
                                Prob > chi2         =         0.6513
Log likelihood = -10591.756      Pseudo R2       =         0.0000
```

```
-----+-----
health |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
_Iq20_2 |   .0187017   .0413822    0.45   0.651   - .0624059  .0998092
-----+-----
_cut1 |  -3.420359   .0677052                    (Ancillary parameters)
_cut2 |  -1.697099   .0375624
_cut3 |  -.2858803   .0310746
_cut4 |   1.990555   .0404703
-----+-----
```

```
. xi: ologit health i.q22
i.q22          _Iq22_0-8          (naturally coded; _Iq22_0 omitted)
```

```
Iteration 0:  log likelihood = -10591.858
Iteration 1:  log likelihood = -10580.333
Iteration 2:  log likelihood = -10580.329
```

```
Ordered logit estimates          Number of obs   =       7936
                                LR chi2(5)        =        23.06
                                Prob > chi2         =         0.0003
Log likelihood = -10580.329      Pseudo R2       =         0.0011
```

```
-----+-----
health |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
```

_Iq22_1	-.1895166	.0580654	-3.26	0.001	-.3033227	-.0757104
_Iq22_2	-.1196108	.1058363	-1.13	0.258	-.3270462	.0878246
_Iq22_5	-.293406	.0935113	-3.14	0.002	-.4766847	-.1101272
_Iq22_6	-.4487475	.1079896	-4.16	0.000	-.6604032	-.2370919
_Iq22_8	.4195536	.8956149	0.47	0.639	-1.335819	2.174926

_cut1	-3.610425	.0809442	(Ancillary parameters)			
_cut2	-1.884647	.0576972				
_cut3	-.471047	.0531925				
_cut4	1.808269	.0585017				

```
. xi: ologit health i.q23a
i.q23a      _Iq23a_1-8      (naturally coded; _Iq23a_1 omitted)
```

```
Iteration 0:  log likelihood = -10591.858
Iteration 1:  log likelihood = -10483.347
Iteration 2:  log likelihood = -10483.055
Iteration 3:  log likelihood = -10483.055
```

```
Ordered logit estimates      Number of obs   =      7936
                             LR chi2(6)             =      217.60
                             Prob > chi2            =      0.0000
Log likelihood = -10483.055  Pseudo R2       =      0.0103
```

health	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
_Iq23a_2	.3977246	.0558432	7.12	0.000	.2882739 .5071753
_Iq23a_3	.5557151	.0732117	7.59	0.000	.4122228 .6992073
_Iq23a_4	.5241419	.061215	8.56	0.000	.4041628 .644121
_Iq23a_5	.8618676	.0743085	11.60	0.000	.7162255 1.00751
_Iq23a_6	1.006116	.1110069	9.06	0.000	.788546 1.223685
_Iq23a_8	-.5077291	.5178311	-0.98	0.327	-1.522659 .5072012

_cut1	-3.086415	.0698934	(Ancillary parameters)		
_cut2	-1.346175	.042255			
_cut3	.0865167	.0379537			
_cut4	2.398635	.0480196			

```
. xi: ologit health i.q24a
i.q24a      _Iq24a_1-8      (naturally coded; _Iq24a_1 omitted)
```

```
Iteration 0:  log likelihood = -10591.858
Iteration 1:  log likelihood = -10539.758
Iteration 2:  log likelihood = -10539.691
Iteration 3:  log likelihood = -10539.691
```

```
Ordered logit estimates      Number of obs   =      7936
                             LR chi2(2)             =      104.33
                             Prob > chi2            =      0.0000
Log likelihood = -10539.691  Pseudo R2       =      0.0049
```

health	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
_Iq24a_2	.429127	.0422564	10.16	0.000	.346306 .5119479
_Iq24a_8	-.8286221	1.504836	-0.55	0.582	-3.778046 2.120801

_cut1	-3.268931	.0661552	(Ancillary parameters)		
_cut2	-1.539777	.0349624			
_cut3	-.1174667	.0285933			

```

      _cut4 |      2.178522      .039958
-----+-----
. xi: ologit health i.q33
i.q33      _Iq33_0-8      (naturally coded; _Iq33_0 omitted)

Iteration 0:   log likelihood = -10591.858
Iteration 1:   log likelihood = -10554.031
Iteration 2:   log likelihood = -10553.995
Iteration 3:   log likelihood = -10553.995

Ordered logit estimates      Number of obs   =      7936
                             LR chi2(2)             =      75.72
                             Prob > chi2            =      0.0000
Log likelihood = -10553.995   Pseudo R2       =      0.0036

```

```

-----+-----
health |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
_Iq33_1 |      .3807224   .043906     8.67   0.000     .2946681   .4667766
_Iq33_8 |      -.753     1.504712    -0.50   0.617    -3.702181   2.196181
-----+-----
_cut1 |     -3.194778   .0695881                    (Ancillary parameters)
_cut2 |     -1.462996   .0415934
_cut3 |     -.0430045   .0368201
_cut4 |      2.243059   .0462124
-----+-----

```

```

. xi: ologit health q43

Iteration 0:   log likelihood = -10591.858
Iteration 1:   log likelihood = -10590.779
Iteration 2:   log likelihood = -10590.778

Ordered logit estimates      Number of obs   =      7936
                             LR chi2(1)             =      2.16
                             Prob > chi2            =      0.1418
Log likelihood = -10590.778   Pseudo R2       =      0.0001

```

```

-----+-----
health |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
q43 |     -.0001331   .0000905    -1.47   0.141    -.0003105   .0000443
-----+-----
_cut1 |     -3.439499   .0647389                    (Ancillary parameters)
_cut2 |     -1.716183   .0318273
_cut3 |     -.3048419   .0235836
_cut4 |      1.97198    .034912
-----+-----

```

```

. xi: ologit health i.q56
i.q56      _Iq56_1-9      (naturally coded; _Iq56_1 omitted)

Iteration 0:   log likelihood = -10591.858
Iteration 1:   log likelihood = -10480.14
Iteration 2:   log likelihood = -10479.781
Iteration 3:   log likelihood = -10479.781

Ordered logit estimates      Number of obs   =      7936
                             LR chi2(6)             =      224.15
                             Prob > chi2            =      0.0000
Log likelihood = -10479.781   Pseudo R2       =      0.0106

```

health	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
_Iq56_2	-.1843988	.0669427	-2.75	0.006	-.315604	-.0531936
_Iq56_3	-.53493	.0606195	-8.82	0.000	-.6537421	-.416118
_Iq56_4	-.876902	.0800485	-10.95	0.000	-1.033794	-.7200099
_Iq56_5	-1.375258	.1355554	-10.15	0.000	-1.640942	-1.109575
_Iq56_8	-1.383627	1.342722	-1.03	0.303	-4.015313	1.248059
_Iq56_9	-.4440366	.5801076	-0.77	0.444	-1.581027	.6929534
(Ancillary parameters)						
_cut1	-3.904359	.0816979				
_cut2	-2.160793	.0584615				
_cut3	-.7232577	.0532403				
_cut4	1.588696	.0569944				

```
. xi: ologit health i.q57
i.q57          _Iq57_0-9          (naturally coded; _Iq57_0 omitted)
```

```
Iteration 0:  log likelihood = -10591.858
Iteration 1:  log likelihood = -10576.376
Iteration 2:  log likelihood = -10576.37
```

```
Ordered logit estimates          Number of obs   =       7936
                                LR chi2(3)         =       30.98
                                Prob > chi2         =       0.0000
Log likelihood = -10576.37      Pseudo R2       =       0.0015
```

health	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
_Iq57_1	.3043456	.0573614	5.31	0.000	.1919193	.4167719
_Iq57_8	1.15142	.8096442	1.42	0.155	-.4354537	2.738293
_Iq57_9	.7993375	.419592	1.91	0.057	-.0230477	1.621723
(Ancillary parameters)						
_cut1	-3.176689	.0796936				
_cut2	-1.450915	.0568457				
_cut3	-.0362523	.0534755				
_cut4	2.245117	.0604201				

```
. xi: ologit health i.q82
i.q82          _Iq82_0-8          (naturally coded; _Iq82_0 omitted)
```

```
Iteration 0:  log likelihood = -10591.858
Iteration 1:  log likelihood = -10454.284
Iteration 2:  log likelihood = -10453.839
Iteration 3:  log likelihood = -10453.839
```

```
Ordered logit estimates          Number of obs   =       7936
                                LR chi2(5)         =      276.04
                                Prob > chi2         =       0.0000
Log likelihood = -10453.839      Pseudo R2       =       0.0130
```

health	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
_Iq82_1	.7461462	.0941475	7.93	0.000	.5616205	.930672
_Iq82_2	.7155962	.0515994	13.87	0.000	.6144633	.8167292
_Iq82_4	.136569	.0576186	2.37	0.018	.0236386	.2494993
_Iq82_5	.9564287	.1129791	8.47	0.000	.7349938	1.177864
_Iq82_8	-.417195	.4713674	-0.89	0.376	-1.341058	.5066682

```

      _cut1 |    -3.10353    .0707791          (Ancillary parameters)
      _cut2 |   -1.357971    .0435614
      _cut3 |    .084318    .0394123
      _cut4 |    2.405304    .049301
-----

```

```

. xi: ologit health i.q83
i.q83          _Iq83_1-9          (naturally coded; _Iq83_1 omitted)

```

```

Iteration 0:  log likelihood = -10591.858
Iteration 1:  log likelihood = -10402.843
Iteration 2:  log likelihood = -10401.81
Iteration 3:  log likelihood = -10401.81

```

```

Ordered logit estimates                                Number of obs   =       7936
                                                       LR chi2(6)      =       380.09
                                                       Prob > chi2     =       0.0000
Log likelihood = -10401.81                            Pseudo R2       =       0.0179

```

```

-----
      health |          Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      _Iq83_2 |   -0.2458439   .0644558     -3.81  0.000   -0.3721749   -0.1195129
      _Iq83_3 |   -0.5413356   .0624289     -8.67  0.000   -0.663694   -0.4189773
      _Iq83_4 |  -0.1051153   .0651047    -16.15  0.000   -1.178756   -0.9235502
      _Iq83_5 |  -0.1876634   .1733469    -10.83  0.000   -2.216388   -1.53688
      _Iq83_8 |  -0.7501151   .4525225     -1.66  0.097   -1.637043    .1368128
      _Iq83_9 |  -0.6582548   .4353523     -1.51  0.131   -1.51153    .19502
-----
      _cut1 |   -4.030853   .0805444          (Ancillary parameters)
      _cut2 |   -2.273889   .0566808
      _cut3 |   -0.8152544   .0508525
      _cut4 |    1.523438   .0539876
-----

```

```

. xi: ologit health i.income
i.income      _Iincome_1-99      (naturally coded; _Iincome_1 omitted)

```

```

Iteration 0:  log likelihood = -10591.858
Iteration 1:  log likelihood = -10433.142
Iteration 2:  log likelihood = -10432.515
Iteration 3:  log likelihood = -10432.515

```

```

Ordered logit estimates                                Number of obs   =       7936
                                                       LR chi2(11)     =       318.69
                                                       Prob > chi2     =       0.0000
Log likelihood = -10432.515                            Pseudo R2       =       0.0150

```

```

-----
      health |          Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      _Iincome_2 |   .2103068   .088029     2.39  0.017   .0377732   .3828404
      _Iincome_3 |   .5088313   .0926385     5.49  0.000   .3272631   .6903995
      _Iincome_4 |   .6542887   .0888748     7.36  0.000   .4800973   .8284802
      _Iincome_5 |   .5673874   .0969046     5.86  0.000   .3774578   .757317
      _Iincome_6 |   .8487435   .0860957     9.86  0.000   .6799991   1.017488
      _Iincome_7 |   .8915476   .0923763     9.65  0.000   .7104934   1.072602
      _Iincome_8 |   .9785523   .0878826    11.13  0.000   .8063056   1.150799
      _Iincome_9 |   1.260369   .1055217    11.94  0.000   1.05355   1.467188
      _Iincome_10 |  1.49173   .1344917    11.09  0.000   1.228131   1.755329
      _Iincome_98 |  .6248381   .085963     7.27  0.000   .4563536   .7933226
      _Iincome_99 |  .5321979   .0938068     5.67  0.000   .3483398   .7160559
-----

```

```

_cut1 | -2.870934 .080037 (Ancillary parameters)
_cut2 | -1.123726 .0581468
_cut3 | .3206408 .0560362
_cut4 | 2.651407 .0640418

```

```

-----
. xi: ologit health i.q86d
i.q86d      _Iq86d_1-6      (naturally coded; _Iq86d_1 omitted)

```

```

Iteration 0: log likelihood = -10591.858
Iteration 1: log likelihood = -10529.427
Iteration 2: log likelihood = -10529.336

```

```

Ordered logit estimates      Number of obs   =      7936
                             LR chi2(5)             =      125.04
                             Prob > chi2              =      0.0000
Log likelihood = -10529.336  Pseudo R2       =      0.0059

```

```

-----
health |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
_Iq86d_2 | -.0803739   .0926605    -0.87  0.386    - .2619851   .1012374
_Iq86d_3 | -.1719505   .0849683    -2.02  0.043    - .3384852  -.0054157
_Iq86d_4 | -.4512241   .0854893    -5.28  0.000    - .61878    -.2836682
_Iq86d_5 | -.5919141   .0897009    -6.60  0.000    - .7677247  -.4161035
_Iq86d_6 | -.6660371   .0916803    -7.26  0.000    - .8457271  -.486347
-----+-----
_cut1 | -3.809003   .096756     (Ancillary parameters)
_cut2 | -2.073107   .0781612
_cut3 | -.6469113   .074526
_cut4 | 1.647027    .0777347

```

```

-----
. xi: ologit health i.emp
i.emp      _Iemp_0-8      (naturally coded; _Iemp_0 omitted)

```

```

Iteration 0: log likelihood = -10591.858
Iteration 1: log likelihood = -10454.284
Iteration 2: log likelihood = -10453.839
Iteration 3: log likelihood = -10453.839

```

```

Ordered logit estimates      Number of obs   =      7936
                             LR chi2(5)             =      276.04
                             Prob > chi2              =      0.0000
Log likelihood = -10453.839  Pseudo R2       =      0.0130

```

```

-----
health |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
_Iemp_1 | .7461462    .0941475     7.93  0.000     .5616205   .930672
_Iemp_2 | .7155962    .0515994    13.87  0.000     .6144633   .8167292
_Iemp_4 | .136569     .0576186     2.37  0.018     .0236386   .2494993
_Iemp_5 | .9564287    .1129791     8.47  0.000     .7349938   1.177864
_Iemp_8 | -.417195    .4713674    -0.89  0.376    -1.341058   .5066682
-----+-----
_cut1 | -3.10353    .0707791     (Ancillary parameters)
_cut2 | -1.357971   .0435614
_cut3 | .084318     .0394123
_cut4 | 2.405304    .049301

```

REGRESSIONS WITH MULTIPLE INDEPENDENT VARIABLES - MODEL SELECTION

```
. xi: ologit health i.zone smoke i.income i.emp i.q86d i.q83 q57 i.q56 q43 q33
i.q23a i
> .q22 q20 q5
i.zone      _Izone_7-20      (naturally coded; _Izone_7 omitted)
i.income    _Iincome_1-10   (naturally coded; _Iincome_1 omitted)
i.emp       _Iemp_0-5       (naturally coded; _Iemp_0 omitted)
i.q86d      _Iq86d_1-6     (naturally coded; _Iq86d_1 omitted)
i.q83       _Iq83_1-5     (naturally coded; _Iq83_1 omitted)
i.q56       _Iq56_1-5     (naturally coded; _Iq56_1 omitted)
i.q23a      _Iq23a_1-6    (naturally coded; _Iq23a_1 omitted)
i.q22       _Iq22_0-6     (naturally coded; _Iq22_0 omitted)
```

```
Iteration 0:  log likelihood = -8673.3641
Iteration 1:  log likelihood = -8217.2257
Iteration 2:  log likelihood = -8208.8563
Iteration 3:  log likelihood = -8208.8425
Iteration 4:  log likelihood = -8208.8425
```

```
Ordered logit estimates                Number of obs   =      6497
                                         LR chi2(54)     =      929.04
                                         Prob > chi2     =      0.0000
Log likelihood = -8208.8425            Pseudo R2      =      0.0536
```

health	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
_Izone_8	-.0226998	.172518	-0.13	0.895	-.3608288 .3154292
_Izone_9	-.1219689	.1804976	-0.68	0.499	-.4757376 .2317998
_Izone_10	.0007849	.1992586	0.00	0.997	-.3897547 .3913246
_Izone_11	.0929658	.1985059	0.47	0.640	-.2960986 .4820302
_Izone_12	.1436245	.1809269	0.79	0.427	-.2109857 .4982348
_Izone_13	.256242	.2279813	1.12	0.261	-.1905931 .7030772
_Izone_14	.0356593	.1727627	0.21	0.836	-.3029494 .374268
_Izone_15	.1817354	.1828404	0.99	0.320	-.1766251 .540096
_Izone_16	.0450105	.1875004	0.24	0.810	-.3224835 .4125045
_Izone_17	.0730392	.1774265	0.41	0.681	-.2747103 .4207888
_Izone_18	-.048915	.2359036	-0.21	0.836	-.5112776 .4134477
_Izone_19	.0060799	.1671607	0.04	0.971	-.3215491 .333709
_Izone_20	.1199844	.2511209	0.48	0.633	-.3722036 .6121724
smoke	-.3286801	.0495139	-6.64	0.000	-.4257257 -.2316346
_Iincome_2	.0908161	.0933164	0.97	0.330	-.0920808 .273713
_Iincome_3	.3610263	.1015724	3.55	0.000	.161948 .5601046
_Iincome_4	.3543604	.1013836	3.50	0.000	.1556522 .5530685
_Iincome_5	.2056556	.111338	1.85	0.065	-.0125629 .423874
_Iincome_6	.439221	.1053574	4.17	0.000	.2327242 .6457177
_Iincome_7	.4648743	.1142769	4.07	0.000	.2408957 .688853
_Iincome_8	.4512356	.1149437	3.93	0.000	.22595 .6765212
_Iincome_9	.6638274	.1357034	4.89	0.000	.3978537 .9298012
_Iincome_10	.7966959	.1640412	4.86	0.000	.4751811 1.118211
_Iemp_1	.464434	.1102243	4.21	0.000	.2483983 .6804697
_Iemp_2	.4067165	.063648	6.39	0.000	.2819688 .5314643
_Iemp_4	.0679794	.0958835	0.71	0.478	-.1199488 .2559077
_Iemp_5	.6414878	.1607341	3.99	0.000	.3264548 .9565208
_Iq86d_2	.1734855	.1307032	1.33	0.184	-.082688 .429659
_Iq86d_3	.0516814	.1310489	0.39	0.693	-.2051697 .3085324
_Iq86d_4	-.2002073	.1344319	-1.49	0.136	-.463689 .0632744
_Iq86d_5	-.3059176	.1469541	-2.08	0.037	-.5939423 -.0178929
_Iq86d_6	-.3886707	.1709177	-2.27	0.023	-.7236633 -.0536781
_Iq83_2	-.2800906	.074159	-3.78	0.000	-.4254396 -.1347416
_Iq83_3	-.4503836	.0731932	-6.15	0.000	-.5938396 -.3069275
_Iq83_4	-.6045011	.0812546	-7.44	0.000	-.7637573 -.4452449

_Iq83_5	-1.11085	.1948494	-5.70	0.000	-1.492747	-.7289516
q57	.1419476	.0667092	2.13	0.033	.0112001	.2726952
_Iq56_2	-.3308644	.0771198	-4.29	0.000	-.4820164	-.1797124
_Iq56_3	-.7359963	.0718661	-10.24	0.000	-.8768512	-.5951414
_Iq56_4	-1.144645	.0941205	-12.16	0.000	-1.329118	-.9601723
_Iq56_5	-1.635105	.1587767	-10.30	0.000	-1.946301	-1.323908
q43	.0000284	.0001159	0.25	0.806	-.0001987	.0002555
q33	.1901615	.0544583	3.49	0.000	.0834251	.2968978
_Iq23a_2	.0891487	.0675042	1.32	0.187	-.0431571	.2214544
_Iq23a_3	.1790748	.0921926	1.94	0.052	-.0016193	.3597689
_Iq23a_4	.1054668	.0777279	1.36	0.175	-.046877	.2578106
_Iq23a_5	.3091772	.0965569	3.20	0.001	.1199293	.4984252
_Iq23a_6	.4746211	.132533	3.58	0.000	.2148612	.734381
_Iq22_1	.0069451	.086828	0.08	0.936	-.1632346	.1771248
_Iq22_2	.0455376	.1241226	0.37	0.714	-.1977384	.2888135
_Iq22_5	.4179483	.133624	3.13	0.002	.1560502	.6798464
_Iq22_6	.2246334	.1302524	1.72	0.085	-.0306566	.4799235
q20	.1882326	.0512627	3.67	0.000	.0877595	.2887057
q5	-.089094	.0478653	-1.86	0.063	-.1829082	.0047202

_cut1	-3.755705	.2520019			(Ancillary parameters)	
_cut2	-1.909488	.2440919				
_cut3	-.358152	.2428685				
_cut4	2.100575	.2442065				

. lrtest, saving (0)

```
. xi: ologit health smoke i.income i.emp i.q86d i.q83 q57 i.q56 q33 i.q23a i.q22
q20 q5
i.income      _Iincome_1-10      (naturally coded; _Iincome_1 omitted)
i.emp         _Iemp_0-5        (naturally coded; _Iemp_0 omitted)
i.q86d       _Iq86d_1-6      (naturally coded; _Iq86d_1 omitted)
i.q83        _Iq83_1-5     (naturally coded; _Iq83_1 omitted)
i.q56        _Iq56_1-5     (naturally coded; _Iq56_1 omitted)
i.q23a       _Iq23a_1-6    (naturally coded; _Iq23a_1 omitted)
i.q22        _Iq22_0-6    (naturally coded; _Iq22_0 omitted)
```

```
Iteration 0:  log likelihood = -8673.3641
Iteration 1:  log likelihood = -8222.7722
Iteration 2:  log likelihood = -8214.612
Iteration 3:  log likelihood = -8214.5988
```

```
Ordered logit estimates                Number of obs   =       6497
                                        LR chi2(40)      =       917.53
                                        Prob > chi2     =       0.0000
Log likelihood = -8214.5988           Pseudo R2      =       0.0529
```

health	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
smoke	-.3317028	.0494069	-6.71	0.000	-.4285386 -.234867
_Iincome_2	.0979888	.0929481	1.05	0.292	-.0841861 .2801636
_Iincome_3	.3740152	.1012069	3.70	0.000	.1756533 .5723772
_Iincome_4	.3633512	.1010335	3.60	0.000	.1653291 .5613732
_Iincome_5	.2212836	.1107787	2.00	0.046	.0041614 .4384058
_Iincome_6	.448414	.1047429	4.28	0.000	.2431216 .6537064
_Iincome_7	.4665504	.1135217	4.11	0.000	.2440519 .6890488
_Iincome_8	.4526114	.1139925	3.97	0.000	.2291903 .6760326
_Iincome_9	.6581094	.13457	4.89	0.000	.3943572 .9218617
_Iincome_10	.7941005	.1624722	4.89	0.000	.4756609 1.11254
_Iemp_1	.4602973	.109961	4.19	0.000	.2447777 .6758169
_Iemp_2	.400214	.0632967	6.32	0.000	.2761547 .5242733

_Iemp_4	.0635077	.0957203	0.66	0.507	-.1241007	.2511161
_Iemp_5	.6309197	.1603219	3.94	0.000	.3166945	.9451449
_Iq86d_2	.17801	.1305476	1.36	0.173	-.0778587	.4338786
_Iq86d_3	.0520206	.1308117	0.40	0.691	-.2043656	.3084068
_Iq86d_4	-.2010625	.1341887	-1.50	0.134	-.4640676	.0619426
_Iq86d_5	-.3072068	.1466767	-2.09	0.036	-.594688	-.0197257
_Iq86d_6	-.3917819	.1703922	-2.30	0.021	-.7257445	-.0578192
_Iq83_2	-.2844089	.0740651	-3.84	0.000	-.4295739	-.1392439
_Iq83_3	-.4561795	.0730507	-6.24	0.000	-.5993561	-.3130028
_Iq83_4	-.6132084	.0809903	-7.57	0.000	-.7719465	-.4544702
_Iq83_5	-1.125987	.1946421	-5.78	0.000	-1.507479	-.7444959
q57	.1391856	.0666378	2.09	0.037	.008578	.2697932
_Iq56_2	-.3372676	.0769778	-4.38	0.000	-.4881414	-.1863938
_Iq56_3	-.741561	.0717214	-10.34	0.000	-.8821324	-.6009896
_Iq56_4	-1.150519	.0939741	-12.24	0.000	-1.334705	-.9663326
_Iq56_5	-1.641962	.1585567	-10.36	0.000	-1.952728	-1.331197
q33	.1817921	.0540398	3.36	0.001	.0758761	.2877081
_Iq23a_2	.0842552	.0671659	1.25	0.210	-.0473875	.2158979
_Iq23a_3	.1731573	.0912365	1.90	0.058	-.0056629	.3519776
_Iq23a_4	.0980792	.0772688	1.27	0.204	-.0533649	.2495233
_Iq23a_5	.3024413	.095657	3.16	0.002	.1149572	.4899255
_Iq23a_6	.4576426	.1316031	3.48	0.001	.1997053	.7155799
_Iq22_1	.0168233	.0863018	0.19	0.845	-.1523252	.1859717
_Iq22_2	.0446569	.1237569	0.36	0.718	-.1979021	.2872158
_Iq22_5	.4291056	.1333028	3.22	0.001	.1678369	.6903743
_Iq22_6	.2201365	.1300975	1.69	0.091	-.0348499	.4751229
q20	.1884769	.0510117	3.69	0.000	.0884958	.2884579
q5	-.092816	.0478055	-1.94	0.052	-.1865131	.0008811

_cut1	-3.807349	.194228			(Ancillary parameters)	
_cut2	-1.961897	.183887				
_cut3	-.4128028	.1821838				
_cut4	2.043427	.1838485				

```

. lrtest, using(0)
Ologit: likelihood-ratio test                chi2(14)    =    11.51
                                              Prob > chi2 =    0.6454

. lrtest, saving(1)

. xi: ologit health smoke i.income i.emp i.q86d i.q83 q57 i.q56 q33 i.q23a i.q22
q20
i.income      _Iincome_1-10      (naturally coded; _Iincome_1 omitted)
i.emp         _Iemp_0-5        (naturally coded; _Iemp_0 omitted)
i.q86d        _Iq86d_1-6      (naturally coded; _Iq86d_1 omitted)
i.q83         _Iq83_1-5      (naturally coded; _Iq83_1 omitted)
i.q56         _Iq56_1-5      (naturally coded; _Iq56_1 omitted)
i.q23a        _Iq23a_1-6     (naturally coded; _Iq23a_1 omitted)
i.q22         _Iq22_0-6     (naturally coded; _Iq22_0 omitted)

Iteration 0:  log likelihood = -8673.3641
Iteration 1:  log likelihood = -8224.6149
Iteration 2:  log likelihood = -8216.4969
Iteration 3:  log likelihood = -8216.4839

Ordered logit estimates                Number of obs    =    6497
                                         LR chi2(39)     =    913.76
                                         Prob > chi2     =    0.0000
Log likelihood = -8216.4839            Pseudo R2       =    0.0527

```

```

-----
health |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]

```

smoke	-.3320528	.0494061	-6.72	0.000	-.4288868	-.2352187
_Iincome_2	.0965714	.092939	1.04	0.299	-.0855857	.2787284
_Iincome_3	.3714562	.1011773	3.67	0.000	.1731522	.5697601
_Iincome_4	.3613994	.1009981	3.58	0.000	.1634467	.5593521
_Iincome_5	.2159602	.1107558	1.95	0.051	-.0011172	.4330377
_Iincome_6	.4425594	.1047022	4.23	0.000	.2373469	.6477719
_Iincome_7	.4627946	.1135012	4.08	0.000	.2403364	.6852528
_Iincome_8	.4459064	.1139611	3.91	0.000	.2225468	.6692661
_Iincome_9	.6506125	.134516	4.84	0.000	.386966	.9142591
_Iincome_10	.7821174	.1623367	4.82	0.000	.4639433	1.100292
_Iemp_1	.4621264	.10995	4.20	0.000	.2466283	.6776244
_Iemp_2	.4006815	.0632937	6.33	0.000	.2766281	.524735
_Iemp_4	.0615203	.0957382	0.64	0.520	-.126123	.2491637
_Iemp_5	.6244382	.1603056	3.90	0.000	.3102451	.9386313
_Iq86d_2	.1816046	.1305902	1.39	0.164	-.0743474	.4375567
_Iq86d_3	.0608086	.1307882	0.46	0.642	-.1955316	.3171488
_Iq86d_4	-.1940043	.1341952	-1.45	0.148	-.4570221	.0690134
_Iq86d_5	-.2953565	.1466028	-2.01	0.044	-.5826927	-.0080204
_Iq86d_6	-.3703586	.1700807	-2.18	0.029	-.7037107	-.0370065
_Iq83_2	-.2839419	.0740681	-3.83	0.000	-.4291127	-.1387711
_Iq83_3	-.4552719	.0730464	-6.23	0.000	-.5984402	-.3121036
_Iq83_4	-.6140877	.0809835	-7.58	0.000	-.7728124	-.455363
_Iq83_5	-1.128378	.1946116	-5.80	0.000	-1.50981	-.7469465
q57	.130714	.066462	1.97	0.049	.0004508	.2609771
_Iq56_2	-.3354892	.0769591	-4.36	0.000	-.4863263	-.1846521
_Iq56_3	-.7421706	.0717148	-10.35	0.000	-.8827291	-.6016121
_Iq56_4	-1.151904	.093974	-12.26	0.000	-1.336089	-.9677179
_Iq56_5	-1.636838	.1584991	-10.33	0.000	-1.947491	-1.326186
q33	.1809817	.0540348	3.35	0.001	.0750753	.286888
_Iq23a_2	.0784924	.0670973	1.17	0.242	-.0530158	.2100006
_Iq23a_3	.1615494	.0910269	1.77	0.076	-.01686	.3399588
_Iq23a_4	.0909893	.0771753	1.18	0.238	-.0602714	.24225
_Iq23a_5	.2920764	.095499	3.06	0.002	.1049017	.4792511
_Iq23a_6	.4503631	.1315481	3.42	0.001	.1925335	.7081928
_Iq22_1	.0153155	.0863164	0.18	0.859	-.1538615	.1844925
_Iq22_2	.0430059	.1237299	0.35	0.728	-.1995002	.285512
_Iq22_5	.4305419	.1332886	3.23	0.001	.169301	.6917828
_Iq22_6	.2160986	.1300931	1.66	0.097	-.0388791	.4710764
q20	.1824966	.0509203	3.58	0.000	.0826947	.2822985
_cut1	-3.782831	.1938038				
_cut2	-1.937472	.1834487				
_cut3	-.3888243	.1817629				
_cut4	2.066078	.1834951				

(Ancillary parameters)

```
. lrtest, using(0)
Ologit: likelihood-ratio test          chi2(15) = 15.28
                                         Prob > chi2 = 0.4312
```

```
. lrtest, saving(2)
```

```
. xi: ologit health smoke i.income i.emp i.q86d i.q83 q57 i.q56 q33 i.q23a q20
i.income      _Iincome_1-10      (naturally coded; _Iincome_1 omitted)
i.emp         _Iemp_0-5          (naturally coded; _Iemp_0 omitted)
i.q86d       _Iq86d_1-6        (naturally coded; _Iq86d_1 omitted)
i.q83        _Iq83_1-5         (naturally coded; _Iq83_1 omitted)
i.q56        _Iq56_1-5         (naturally coded; _Iq56_1 omitted)
i.q23a       _Iq23a_1-6        (naturally coded; _Iq23a_1 omitted)
```

```
Iteration 0: log likelihood = -8673.3641
Iteration 1: log likelihood = -8232.7266
```

Iteration 2: log likelihood = -8224.8927
 Iteration 3: log likelihood = -8224.8805

Ordered logit estimates

Number of obs	=	6497
LR chi2(35)	=	896.97
Prob > chi2	=	0.0000
Pseudo R2	=	0.0517

Log likelihood = -8224.8805

health	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
smoke	-.3317123	.0492905	-6.73	0.000	-.4283199 -.2351046
_Iincome_2	.0348526	.0910126	0.38	0.702	-.1435288 .2132339
_Iincome_3	.2967236	.0983512	3.02	0.003	.1039588 .4894883
_Iincome_4	.2800512	.0975787	2.87	0.004	.0888004 .4713019
_Iincome_5	.1322241	.1071015	1.23	0.217	-.0776909 .3421391
_Iincome_6	.3574441	.1008211	3.55	0.000	.1598384 .5550498
_Iincome_7	.3741336	.109059	3.43	0.001	.1603819 .5878852
_Iincome_8	.3584003	.1095073	3.27	0.001	.1437699 .5730307
_Iincome_9	.5627359	.1301028	4.33	0.000	.3077391 .8177327
_Iincome_10	.6989775	.1587126	4.40	0.000	.3879064 1.010048
_Iemp_1	.4754757	.1098392	4.33	0.000	.260195 .6907565
_Iemp_2	.421165	.0630524	6.68	0.000	.2975846 .5447454
_Iemp_4	.088411	.0954384	0.93	0.354	-.0986449 .2754669
_Iemp_5	.6311563	.1593281	3.96	0.000	.318879 .9434336
_Iq86d_2	.2016371	.1224312	1.65	0.100	-.0383237 .4415978
_Iq86d_3	.092386	.1177781	0.78	0.433	-.1384547 .3232268
_Iq86d_4	-.1544555	.1194131	-1.29	0.196	-.3885009 .0795898
_Iq86d_5	-.2458982	.1320022	-1.86	0.062	-.5046176 .0128213
_Iq86d_6	-.2368874	.1547279	-1.53	0.126	-.5401484 .0663736
_Iq83_2	-.2778399	.074002	-3.75	0.000	-.422881 -.1327987
_Iq83_3	-.4486866	.0729263	-6.15	0.000	-.5916195 -.3057537
_Iq83_4	-.6083241	.0807732	-7.53	0.000	-.7666365 -.4500116
_Iq83_5	-1.127086	.194128	-5.81	0.000	-1.50757 -.7466024
q57	.1441704	.0662982	2.17	0.030	.0142282 .2741125
_Iq56_2	-.3384064	.0769118	-4.40	0.000	-.4891508 -.187662
_Iq56_3	-.7433348	.071689	-10.37	0.000	-.8838426 -.602827
_Iq56_4	-1.149998	.0938422	-12.25	0.000	-1.333926 -.9660711
_Iq56_5	-1.627974	.1582297	-10.29	0.000	-1.938099 -1.31785
q33	.1823119	.0539951	3.38	0.001	.0764834 .2881403
_Iq23a_2	.0836718	.0669554	1.25	0.211	-.0475584 .2149021
_Iq23a_3	.162191	.0907224	1.79	0.074	-.0156217 .3400036
_Iq23a_4	.0948177	.0769615	1.23	0.218	-.0560242 .2456595
_Iq23a_5	.2919325	.0949125	3.08	0.002	.1059075 .4779575
_Iq23a_6	.4504838	.1311077	3.44	0.001	.1935174 .7074501
q20	.210597	.0503618	4.18	0.000	.1118897 .3093042
(Ancillary parameters)					
_cut1	-3.774626	.1929192			
_cut2	-1.932234	.1824985			
_cut3	-.3871731	.1808096			
_cut4	2.064961	.1825863			

. lrtest, using(0)
 Ologit: likelihood-ratio test

chi2(19)	=	32.08
Prob > chi2	=	0.0306

. lrtest, saving(3)

. xi: ologit health smoke i.income i.emp i.q83 q57 i.q56 q33 i.q23a i.q22 q20
 i.income _Iincome_1-10 (naturally coded; _Iincome_1 omitted)
 i.emp _Iemp_0-5 (naturally coded; _Iemp_0 omitted)
 i.q83 _Iq83_1-5 (naturally coded; _Iq83_1 omitted)

```

i.q56          _Iq56_1-5          (naturally coded; _Iq56_1 omitted)
i.q23a         _Iq23a_1-6        (naturally coded; _Iq23a_1 omitted)
i.q22          _Iq22_0-6        (naturally coded; _Iq22_0 omitted)

```

```

Iteration 0:  log likelihood = -8673.3641
Iteration 1:  log likelihood = -8241.8029
Iteration 2:  log likelihood = -8234.3397
Iteration 3:  log likelihood = -8234.3287

```

```

Ordered logit estimates          Number of obs   =      6497
                                LR chi2(34)         =      878.07
                                Prob > chi2         =      0.0000
Log likelihood = -8234.3287     Pseudo R2       =      0.0506

```

health	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
smoke	-.3371128	.0493208	-6.84	0.000	-.4337798 -.2404459
_Iincome_2	.0891576	.0927704	0.96	0.337	-.0926689 .2709842
_Iincome_3	.3775481	.1011163	3.73	0.000	.1793637 .5757324
_Iincome_4	.3752342	.1005839	3.73	0.000	.1780934 .572375
_Iincome_5	.2272146	.1105541	2.06	0.040	.0105326 .4438967
_Iincome_6	.4547776	.1043284	4.36	0.000	.2502976 .6592575
_Iincome_7	.4609694	.112908	4.08	0.000	.2396738 .682265
_Iincome_8	.4604107	.1134062	4.06	0.000	.2381386 .6826828
_Iincome_9	.6534942	.1338063	4.88	0.000	.3912387 .9157498
_Iincome_10	.7785692	.1614351	4.82	0.000	.4621623 1.094976
_Iemp_1	.4355485	.1097538	3.97	0.000	.2204351 .6506619
_Iemp_2	.4145723	.0631798	6.56	0.000	.2907421 .5384025
_Iemp_4	-.1565771	.0726904	-2.15	0.031	-.2990477 -.0141064
_Iemp_5	.6579527	.1516126	4.34	0.000	.3607974 .955108
_Iq83_2	-.2527908	.0737558	-3.43	0.001	-.3973496 -.108232
_Iq83_3	-.4330492	.0728252	-5.95	0.000	-.5757839 -.2903145
_Iq83_4	-.5857695	.0805162	-7.28	0.000	-.7435783 -.4279607
_Iq83_5	-1.071427	.194216	-5.52	0.000	-1.452083 -.6907707
q57	.137539	.0663268	2.07	0.038	.0075409 .2675371
_Iq56_2	-.3304623	.0768328	-4.30	0.000	-.4810519 -.1798728
_Iq56_3	-.7196058	.0714435	-10.07	0.000	-.8596324 -.5795792
_Iq56_4	-1.122632	.0937559	-11.97	0.000	-1.30639 -.9388738
_Iq56_5	-1.621492	.1582301	-10.25	0.000	-1.931618 -1.311367
q33	.2049722	.0538077	3.81	0.000	.099511 .3104334
_Iq23a_2	.1172268	.0665752	1.76	0.078	-.0132583 .2477118
_Iq23a_3	.2026796	.090404	2.24	0.025	.0254909 .3798683
_Iq23a_4	.174343	.0757844	2.30	0.021	.0258083 .3228776
_Iq23a_5	.352219	.0946797	3.72	0.000	.1666502 .5377878
_Iq23a_6	.5201911	.130765	3.98	0.000	.2638964 .7764857
_Iq22_1	-.0559378	.0763764	-0.73	0.464	-.2056328 .0937571
_Iq22_2	.07031	.1213021	0.58	0.562	-.1674377 .3080577
_Iq22_5	.3003048	.1239041	2.42	0.015	.0574573 .5431523
_Iq22_6	.1452023	.1232327	1.18	0.239	-.0963294 .386734
q20	.2038931	.0507342	4.02	0.000	.1044558 .3033303

_cut1	-3.63596	.1699368	(Ancillary parameters)		
_cut2	-1.797637	.1582468			
_cut3	-.2556818	.1565674			
_cut4	2.192512	.1587997			

```

. lrtest, using(0)
Ologit: likelihood-ratio test          chi2(20)   =      50.97
                                         Prob > chi2 =      0.0002

```

```

. lrtest, saving(4)

```

```
. xi: ologit health smoke i.income i.emp i.q86d i.q83 i.q56 q33 i.q23a i.q22 q20
i.income      _Iincome_1-10      (naturally coded; _Iincome_1 omitted)
i.emp         _Iemp_0-5         (naturally coded; _Iemp_0 omitted)
i.q86d       _Iq86d_1-6       (naturally coded; _Iq86d_1 omitted)
i.q83        _Iq83_1-5       (naturally coded; _Iq83_1 omitted)
i.q56        _Iq56_1-5       (naturally coded; _Iq56_1 omitted)
i.q23a       _Iq23a_1-6     (naturally coded; _Iq23a_1 omitted)
i.q22        _Iq22_0-6     (naturally coded; _Iq22_0 omitted)
```

```
Iteration 0:  log likelihood = -8673.3641
Iteration 1:  log likelihood = -8226.5105
Iteration 2:  log likelihood = -8218.4286
Iteration 3:  log likelihood = -8218.4157
```

```
Ordered logit estimates                                Number of obs   =       6497
                                                       LR chi2(38)    =       909.90
                                                       Prob > chi2    =       0.0000
Log likelihood = -8218.4157                            Pseudo R2      =       0.0525
```

health	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
smoke	-.3296562	.0493879	-6.67	0.000	-.4264547 -.2328576
_Iincome_2	.1016297	.0928961	1.09	0.274	-.0804434 .2837028
_Iincome_3	.3734737	.101156	3.69	0.000	.1752115 .5717359
_Iincome_4	.3682789	.1009288	3.65	0.000	.1704621 .5660957
_Iincome_5	.22211	.1107209	2.01	0.045	.005101 .4391191
_Iincome_6	.4484088	.1046706	4.28	0.000	.2432581 .6535594
_Iincome_7	.467121	.1134785	4.12	0.000	.2447072 .6895348
_Iincome_8	.4506697	.1139379	3.96	0.000	.2273556 .6739838
_Iincome_9	.6568796	.1344977	4.88	0.000	.393269 .9204902
_Iincome_10	.7853085	.1622772	4.84	0.000	.4672511 1.103366
_Iemp_1	.4641974	.1099169	4.22	0.000	.2487642 .6796307
_Iemp_2	.4008154	.0632991	6.33	0.000	.2767514 .5248794
_Iemp_4	.0629036	.0957216	0.66	0.511	-.1247073 .2505145
_Iemp_5	.6266253	.1603893	3.91	0.000	.312268 .9409826
_Iq86d_2	.1763513	.1305492	1.35	0.177	-.0795203 .432223
_Iq86d_3	.0501964	.1306473	0.38	0.701	-.2058675 .3062604
_Iq86d_4	-.2053835	.1340533	-1.53	0.125	-.4681233 .0573562
_Iq86d_5	-.3036433	.1465114	-2.07	0.038	-.5908003 -.0164864
_Iq86d_6	-.3805574	.1699661	-2.24	0.025	-.7136848 -.04743
_Iq83_2	-.2826052	.0740658	-3.82	0.000	-.4277716 -.1374389
_Iq83_3	-.4549863	.0730558	-6.23	0.000	-.598173 -.3117996
_Iq83_4	-.6181056	.0809724	-7.63	0.000	-.7768086 -.4594026
_Iq83_5	-1.140598	.1945274	-5.86	0.000	-1.521865 -.7593312
_Iq56_2	-.3348904	.0769379	-4.35	0.000	-.4856859 -.1840949
_Iq56_3	-.7441713	.0716924	-10.38	0.000	-.8846859 -.6036567
_Iq56_4	-1.162446	.0938192	-12.39	0.000	-1.346329 -.9785641
_Iq56_5	-1.654763	.1582061	-10.46	0.000	-1.964841 -1.344684
q33	.1814365	.0540337	3.36	0.001	.0755324 .2873406
_Iq23a_2	.0809798	.0670696	1.21	0.227	-.0504742 .2124337
_Iq23a_3	.1655659	.0910006	1.82	0.069	-.012792 .3439239
_Iq23a_4	.0928094	.0771737	1.20	0.229	-.0584482 .244067
_Iq23a_5	.2993923	.0954075	3.14	0.002	.1123971 .4863876
_Iq23a_6	.4607953	.1314552	3.51	0.000	.2031478 .7184427
_Iq22_1	.0081148	.0862292	0.09	0.925	-.1608914 .177121
_Iq22_2	.0415587	.1236771	0.34	0.737	-.200844 .2839613
_Iq22_5	.4335875	.1332465	3.25	0.001	.1724291 .6947459
_Iq22_6	.2167619	.1300472	1.67	0.096	-.0381258 .4716497
q20	.2038037	.0497516	4.10	0.000	.1062923 .301315
_cut1	-3.868348	.1888746			(Ancillary parameters)
_cut2	-2.023474	.1781428			

```

      _cut3 |   - .4753291   .1763105
      _cut4 |    1.978932   .1779572
-----+-----
. lrtest, using(0)
Ologit: likelihood-ratio test                chi2(16)   =    19.15
                                              Prob > chi2 =    0.2611

. lrtest, saving(5)

. xi: ologit health smoke i.income i.emp i.q83 i.q56 q33 i.q23a i.q22 q20
i.income      _Iincome_1-10      (naturally coded; _Iincome_1 omitted)
i.emp          _Iemp_0-5          (naturally coded; _Iemp_0 omitted)
i.q83          _Iq83_1-5         (naturally coded; _Iq83_1 omitted)
i.q56          _Iq56_1-5         (naturally coded; _Iq56_1 omitted)
i.q23a         _Iq23a_1-6        (naturally coded; _Iq23a_1 omitted)
i.q22          _Iq22_0-6         (naturally coded; _Iq22_0 omitted)

Iteration 0:   log likelihood = -8673.3641
Iteration 1:   log likelihood = -8243.9168
Iteration 2:   log likelihood = -8236.4871
Iteration 3:   log likelihood = -8236.4762

Ordered logit estimates                Number of obs   =    6497
                                         LR chi2(33)     =    873.78
                                         Prob > chi2     =    0.0000
Log likelihood = -8236.4762            Pseudo R2      =    0.0504

```

```

-----+-----
      health |      Coef.   Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
      smoke |   - .3347387   .0493045    -6.79   0.000   - .4313736   - .2381037
  _Iincome_2 |   .0945332   .0927249     1.02   0.308   - .0872043   .2762707
  _Iincome_3 |   .3795289   .1010928     3.75   0.000   .1813906   .5776672
  _Iincome_4 |   .3820065   .1005169     3.80   0.000   .1849971   .579016
  _Iincome_5 |   .2334268   .1105245     2.11   0.035   .0168028   .4500508
  _Iincome_6 |   .4603234   .1043034     4.41   0.000   .2558925   .6647543
  _Iincome_7 |   .4648739   .1128845     4.12   0.000   .2436244   .6861234
  _Iincome_8 |   .4649638   .1133858     4.10   0.000   .2427318   .6871958
  _Iincome_9 |   .6595087   .1337938     4.93   0.000   .3972776   .9217398
  _Iincome_10 |   .7815016   .1613725     4.84   0.000   .4652174   1.097786
  _Iemp_1     |   .4372136   .1097223     3.98   0.000   .2221619   .6522652
  _Iemp_2     |   .4147277   .063184      6.56   0.000   .2908893   .5385661
  _Iemp_4     |  - .1551925   .0726699    -2.14   0.033   - .297623   - .012762
  _Iemp_5     |   .6646288   .151671      4.38   0.000   .3673591   .9618985
  _Iq83_2     |  - .2510413   .0737523    -3.40   0.001   - .3955931   - .1064895
  _Iq83_3     |  - .4324976   .072834     -5.94   0.000   - .5752497   - .2897456
  _Iq83_4     |  - .5903014   .0805032    -7.33   0.000   - .7480848   - .4325179
  _Iq83_5     |  - 1.084255   .1941192    -5.59   0.000   -1.464722   - .7037882
  _Iq56_2     |  - .3299469   .0768096    -4.30   0.000   - .480491   - .1794028
  _Iq56_3     |  - .7218509   .0714175   -10.11   0.000   - .8618267   - .5818751
  _Iq56_4     |  - 1.133605   .0936051   -12.11   0.000   -1.317068   - .9501424
  _Iq56_5     |  - 1.640473   .1579089   -10.39   0.000   -1.949969   -1.330977
      q33     |   .2056899   .0538048     3.82   0.000   .1002345   .3111454
  _Iq23a_2     |   .1205161   .0665372     1.81   0.070   - .0098944   .2509266
  _Iq23a_3     |   .2075952   .090369      2.30   0.022   .0304753   .3847151
  _Iq23a_4     |   .1770859   .0757726      2.34   0.019   .0285743   .3255975
  _Iq23a_5     |   .3604152   .094577      3.81   0.000   .1750477   .5457827
  _Iq23a_6     |   .5316858   .1306655     4.07   0.000   .2755861   .7877856
  _Iq22_1     |  - .0671729   .0761658    -0.88   0.378   - .216455   .0821093
  _Iq22_2     |   .0673747   .1212268     0.56   0.578   - .1702254   .3049748
  _Iq22_5     |   .29963     .1238684     2.42   0.016   .0568524   .5424075
  _Iq22_6     |   .1418976   .1231638     1.15   0.249   - .099499   .3832942

```

q20	.2263729	.0495589	4.57	0.000	.1292392	.3235066

_cut1	-3.719078	.1651768	(Ancillary parameters)			
_cut2	-1.881329	.153017				
_cut3	-.3399825	.1511624				
_cut4	2.107562	.153329				

```
. lrtest, using(0)
Ologit: likelihood-ratio test          chi2(21)    =    55.27
                                         Prob > chi2 =    0.0001
```

```
. lrtest, saving(6)
```

```
. * FINAL MODEL
. xi: ologit health smoke i.income i.emp i.q86d i.q83 i.q56 q33 i.q23a i.q22 q20
i.income      _Iincome_1-10      (naturally coded; _Iincome_1 omitted)
i.emp         _Iemp_0-5          (naturally coded; _Iemp_0 omitted)
i.q86d       _Iq86d_1-6        (naturally coded; _Iq86d_1 omitted)
i.q83        _Iq83_1-5         (naturally coded; _Iq83_1 omitted)
i.q56        _Iq56_1-5         (naturally coded; _Iq56_1 omitted)
i.q23a       _Iq23a_1-6        (naturally coded; _Iq23a_1 omitted)
i.q22        _Iq22_0-6         (naturally coded; _Iq22_0 omitted)
```

```
Iteration 0:  log likelihood = -8673.3641
Iteration 1:  log likelihood = -8226.5105
Iteration 2:  log likelihood = -8218.4286
Iteration 3:  log likelihood = -8218.4157
```

```
Ordered logit estimates          Number of obs   =    6497
                                LR chi2(38)          =    909.90
                                Prob > chi2          =    0.0000
Log likelihood = -8218.4157      Pseudo R2       =    0.0525
```

health	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
smoke	-.3296562	.0493879	-6.67	0.000	-.4264547	-.2328576
_Iincome_2	.1016297	.0928961	1.09	0.274	-.0804434	.2837028
_Iincome_3	.3734737	.101156	3.69	0.000	.1752115	.5717359
_Iincome_4	.3682789	.1009288	3.65	0.000	.1704621	.5660957
_Iincome_5	.22211	.1107209	2.01	0.045	.005101	.4391191
_Iincome_6	.4484088	.1046706	4.28	0.000	.2432581	.6535594
_Iincome_7	.467121	.1134785	4.12	0.000	.2447072	.6895348
_Iincome_8	.4506697	.1139379	3.96	0.000	.2273556	.6739838
_Iincome_9	.6568796	.1344977	4.88	0.000	.393269	.9204902
_Iincome_10	.7853085	.1622772	4.84	0.000	.4672511	1.103366
_Iemp_1	.4641974	.1099169	4.22	0.000	.2487642	.6796307
_Iemp_2	.4008154	.0632991	6.33	0.000	.2767514	.5248794
_Iemp_4	.0629036	.0957216	0.66	0.511	-.1247073	.2505145
_Iemp_5	.6266253	.1603893	3.91	0.000	.312268	.9409826
_Iq86d_2	.1763513	.1305492	1.35	0.177	-.0795203	.432223
_Iq86d_3	.0501964	.1306473	0.38	0.701	-.2058675	.3062604
_Iq86d_4	-.2053835	.1340533	-1.53	0.125	-.4681233	.0573562
_Iq86d_5	-.3036433	.1465114	-2.07	0.038	-.5908003	-.0164864
_Iq86d_6	-.3805574	.1699661	-2.24	0.025	-.7136848	-.04743
_Iq83_2	-.2826052	.0740658	-3.82	0.000	-.4277716	-.1374389
_Iq83_3	-.4549863	.0730558	-6.23	0.000	-.598173	-.3117996
_Iq83_4	-.6181056	.0809724	-7.63	0.000	-.7768086	-.4594026
_Iq83_5	-1.140598	.1945274	-5.86	0.000	-1.521865	-.7593312
_Iq56_2	-.3348904	.0769379	-4.35	0.000	-.4856859	-.1840949
_Iq56_3	-.7441713	.0716924	-10.38	0.000	-.8846859	-.6036567
_Iq56_4	-1.162446	.0938192	-12.39	0.000	-1.346329	-.9785641

_Iq56_5	-1.654763	.1582061	-10.46	0.000	-1.964841	-1.344684
q33	.1814365	.0540337	3.36	0.001	.0755324	.2873406
_Iq23a_2	.0809798	.0670696	1.21	0.227	-.0504742	.2124337
_Iq23a_3	.1655659	.0910006	1.82	0.069	-.012792	.3439239
_Iq23a_4	.0928094	.0771737	1.20	0.229	-.0584482	.244067
_Iq23a_5	.2993923	.0954075	3.14	0.002	.1123971	.4863876
_Iq23a_6	.4607953	.1314552	3.51	0.000	.2031478	.7184427
_Iq22_1	.0081148	.0862292	0.09	0.925	-.1608914	.177121
_Iq22_2	.0415587	.1236771	0.34	0.737	-.200844	.2839613
_Iq22_5	.4335875	.1332465	3.25	0.001	.1724291	.6947459
_Iq22_6	.2167619	.1300472	1.67	0.096	-.0381258	.4716497
q20	.2038037	.0497516	4.10	0.000	.1062923	.301315

_cut1	-3.868348	.1888746	(Ancillary parameters)			
_cut2	-2.023474	.1781428				
_cut3	-.4753291	.1763105				
_cut4	1.978932	.1779572				

. lrtest, using(0)

Ologit: likelihood-ratio test

chi2(16) = 19.15
 Prob > chi2 = 0.2611