UNIVERSITY
OF WOLLONGONG
AUSTRALIA

University of Wollongong - Research Online

Thesis Collection

Title: The multiple inter-relationships among health status, education, income and lifestyle factors: evidence from Australia

Author: Jae Hyung Lee

Year: 1997

Repository DOI:

Copyright Warning
You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.
You are reminded of the following: This work is copyright. Apart from any use permitted under the Copyright Act 1968, no part of this work may be reproduced by any process, nor may any other exclusive right be exercised, without the permission of the author. Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material.
Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

Unless otherwise indicated, the views expressed in this thesis are those of the author and do not necessarily represent the views of the University of Wollongong.

Research Online is the open access repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au

University of Wollongong Thesis Collections

University of Wollongong Thesis Collection

University of Wollongong *Year* 1997

The multiple inter-relationships among health status, education, income and lifestyle factors: evidence from Australia

Jae Hyung Lee
University of Wollongong

Lee, Jae Hyung, The multiple inter-relationships among health status, education, income and lifestyle factors: evidence from Australia, Doctor of Philosophy thesis, Department of Economics, University of Wollongong, 1997. http://ro.uow.edu.au/theses/1325

This paper is posted at Research Online.

NOTE

This online version of the thesis may have different page formatting and pagination from the paper copy held in the University of Wollongong Library.

UNIVERSITY OF WOLLONGONG

COPYRIGHT WARNING

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site. You are reminded of the following:

Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material. Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

UNIVERSITY OF
WOLLONGONG
LIBRARY

THE MULTIPLE INTER-RELATIONSHIPS AMONG HEALTH STATUS, EDUCATION, INCOME, AND LIFESTYLE FACTORS: EVIDENCE FROM AUSTRALIA

By

JAE HYUNG LEE

Ph.D Thesis submitted to the Department of Economics
The University of Wollongong,
Northfields Avenue, Wollongong,
New South Wales, 2522
Australia

Telephone: 61 42 213666
Facsimile: 61 42 213725

ACKNOWLEDGEMENTS

I would like to thank Professor Don Lewis of the University of Wollongong for his encouragement and guidance throughout my study.

Insightful comments on the thesis were made by Professor Phil Lewis of Murdoch University, Associate Professor Rob Castle and Senior Lecturer Ed Wilson of the University of Wollongong.

Addressing the points raised by thesis examiners, Dr. Pat Wilson, Dr. Jim Butler of Australian National University, and Associate Professor John Mangan of Lancaster University and the University of Queensland, improved the thesis.

The Australian Bureau of Statistics was most helpful in supplying data and detailed information about the data used in this study.

Finally, I would like to thank Mr David McCoy and Mrs Rhonda McCoy for their careful typing of numerous drafts and the final version of this thesis.

TABLE OF CONTENTS

TABLE OF CONTENTS - (Continued)

LIST OF TABLES

LIST OF TABLES - (Continued)

LIST OF TABLES - (Continued)

LIST OF TABLES - (Continued)

LIST OF TABLES - (Continued)

LIST OF TABLES - (Continued)

LIST OF FIGURES

ABSTRACT

This dissertation examines the inter-relationships among health, education, income, and health-related behaviour as measured by alcohol consumption and smoking. The cross-section models are estimated using data from Australia's sixty-one statistical regions to analyse multiple causal relationships among factors taken as endogenous. Classifying health indicators as 'no recent illness', 'no chronic condition', and 'self-assessed good or excellent health', ten models are presented and econometrically evaluated. The first, second, third, fourth, and fifth models take the proportions of persons with no recent illness and no chronic conditions, the proportion of persons aged 18 years or over with self-assessed good or excellent health, the proportion of persons aged 15 years or over with post-school qualifications, and nominal gross annual median income of persons aged 15 years or over, respectively, as endogenous. The inter-relationships among the proportion of persons with no recent illness, the proportion of persons aged 18 years or over with self-assessed good or excellent health, and the proportion of persons aged 18 years or over with moderate alcohol consumption are specified in the sixth, seventh, and eighth models. The proportions of persons aged 18 years or over with cigarettes consumption and with excessive alcohol consumption are taken as endogenous in the ninth and tenth models.

Diagnostic checks are conducted to evaluate all the models. The tests include the non-nested tests, the tests of independence, the tests for endogeneity and exogeneity, the RESET tests for functional form misspecification, and the tests for heteroskedasticity. The parameter stability of the models is also tested and then (if any instability is apparent) parameter instability analysis is carried out. In addition, the direct, indirect, and total effects of variables exogenously determined in the individual models are manifested, since all the models are identified.

The empirical evidence is consistent with the hypotheses that there are multiple inter-relationships (a) among the proportion of persons with each of no recent illness and no chronic conditions, the proportion of persons aged 18 years or over with self-assessed good or excellent health, the proportion of persons aged 15 years or over with post-school

qualifications and nominal gross annual median income of persons aged 15 years or over, (b) among the proportion of persons with no recent illness, the proportion of persons aged 18 years or over with self-assessed good or excellent health and the proportion of persons aged 18 years or over with moderate alcohol consumption, and (c) between the proportions of persons aged 18 years or over with cigarettes consumption and with excessive alcohol consumption.

1.0 INTRODUCTION

Even though many countries have experienced rises in health care spending over the course of the last few years, it has become increasingly recognised that medical care is but one of the factors contributing towards good health. This has led to a growth of interest in the non-medical determinants of health; people want good health and health care is one means to that end. Studies using a production of health function view health as the outcome of a production process involving health inputs such as education, income and lifestyle factors, as well as medical care. This literature suggests that once basic levels of medical sophistication, personnel, and facilities become available, additional inputs of medical care do not have much effect on health; while the total contribution of medical care is probably substantial in modern societies, its marginal contribution in generating health improvements is small. Furthermore, the marginal products of other variables (education, in particular) are generally significantly different from zero [Fuchs (1974) and Wagstaff (1989)]. Often it is the marginals, not the totals, that are most relevant in policy formulation.

The notion that medicine played a relatively minor historical role is asserted by Fuchs (1974). Fuchs concludes that rising living standards, the spread of literacy and education, and a substantial fall in the birth rate all played a part in the sharp reduction in the infant mortality rate between 1900 and 1930. In the 1930s, sulphonamide, the first of the anti-microbial drugs, was introduced. Fuchs argues that during the period 1935-1950, a period in which the fall in infant death rates accelerated, both medical advances and rising living standards contributed to the reduction in infant deaths. Fuchs (1974) also notes that "how medical care is used may be more important than how much is used" and that "what is required is some sense of balance so that the contribution of medical care is not oversold".

The production theory of health predicts that, ceteris paribus, education is positively associated with health, since the educated understand the technology needed to stay healthy, and they have better knowledge of how to use medical and other market inputs and their own time to produce better health outcomes. It is also possible to suggest that the completion of formal schooling increases self-confidence and thus reduces the stress associated with many social and work situations. This point is based on the view that increased education contributes to better health status. The positive correlation between health and length of schooling also holds after allowing for the effects of such other variables as income, lifestyle factors, occupation, gender, etc. This relationship may also reflect a chain of causality that begins with good health and results in more schooling.

Higher income is generally associated with better health; people with higher incomes tend to consume higher quality goods and better housing, have a better diet, and use more and better medical care, all of which may favourably affect their health. This suggests a causality which operates from income to health. Assuming that healthy workers are more productive and have lower rates of absenteeism, employers will be willing to pay a higher wage to healthy persons. Therefore, better health will increase income. Hence, the relationship between these two variables might suggest the direction of causation from health to income, as well as the direction of causation from income to health.

Increasing the skills and knowledge of the labour force is viewed as a fundamental way of increasing productivity and economic growth [Lewis (1991)]. The evidence from the study of intra-family differences in education and their correlation with intra-family differences in income also supports the hypothesis that additional schooling is responsible for increases in earnings [Ashenfelter (1993)]. The crux of the issue is whether higher education causes higher income (via higher productivity).

In brief, the relationships between health status and education, between education and income, and between health status and income justify a system of multiple causations; ceteris paribus, an increased level of one is associated with increased levels of the other two.

An implication of this is that the cross effects among health, education, and income are positive. It turns out that they complement each other and therefore health and education can be viewed as normal goods. Lewis (1994, Lecture notes) uses Figure 1.1 to illustrate the multiple causal relationships among these variables and formulate several testable propositions:

Figure 1.1. Inter-relationships Among Health, Education, and Income

'Please see print copy for image'

Note: The direction of causation is indicated by the arrows.
Source: Lewis (1994, Lecture notes).

The testable propositions consistent with letters (a) through (f) in the figure are:

(a) Good health makes it easier to study and proceed through the educational system, and thus facilitates learning. Poor health limits educational progress.

(b) Higher education leads to higher income and limited education causes low income, since productivity and income increase with the skills and knowledge.

(c) Higher income is generally associated with better health as a result of better diet, and access to the better medical care; while low income is related to poor health as a result of unhealthy environment, poor diet, and lack of access to health care system.

(d) Health status affects the ability to earn income through lower absenteeism and higher productivity.

(e) High income facilitates additional years of schooling.

(f) Good education contributes to better health by fostering better knowledge about health issues, increasing receptability to health information, and increasing the ability to deal with health bureaucracies.

The relevance of this framework is that the above propositions suggest a full simultaneous equations model, implying that the choice of policy instruments should not be based upon stability analysis of a single final target variable. This implies, for example, that health status may be improved directly by increased expenditure on medical care or indirectly through increased expenditure on education and through higher income, that education may be facilitated directly by increased expenditure on additional years of schooling or indirectly through increased expenditure on health and through higher income, and that income may be raised directly by increased levels of skill or indirectly through increased expenditures on health and education.

The estimation of the correlations among health status, education, and income is considered in Section 5.1 of Chapter 5.

In Section 5.1 of Chapter 5, a secondary objective of this study is to examine the inter-relationship between differentials in health and the gender distribution. It is suggested that universities yield beneficial health effects to communities in which they are located, due to beneficial education effects. On the other hand, it is also suggested that communities with a university campus have many service industries such as finance and business services, recreation, personal and other services, and community services. Given that the service sector is female-intensive, women move from a community without a service sector to a community with it. An implication of this is that the proportion of the population which is female is greater in communities with a university than in those without. Since women generally have better health than men, the combined effects of a university on health and gender imply that health status has an impact on the gender distribution. Further, women possess more knowledge about health-related issues than men, are more likely to monitor their own health status, are less likely to engage in a number of risky behaviours, and provide medical services to other members of their families [Sindelar (1982) and Umberson (1992)]. Thus, the higher the proportion of females in a community, ceteris paribus, the more likely it is that its population will be healthy and able to enjoy a full and long life, which in turn leads to more education and higher income. On the basis of these arguments, it can also be expected that the contributions of universities to communities in which they are located is large.

The positive association between education and income is one of the most consistent empirical findings of the human capital literature. The conventional view, the productivity augmenting view, is that education enhances earnings via the production of marketable skills - a fundamental way of increasing productivity [Lewis (1991)]. It is suggested implicitly from this view that the level of skill is influenced by income, through education. Further, skilled workers are more attracted to communities where income is expected to be higher. Therefore, communities with higher income tend to have people with higher levels of skill. On the other hand, lack of self-esteem raises the perceived costs of education or training, while a lack of optimism or a short-term orientation limits the expected benefits [Ehrenberg and Smith (1985)]. This suggests that communities with better health are probably much more likely to undertake human capital investments with distant payoffs from becoming skilled. Thus, communities with better health also tend to have people with higher levels of skill. Therefore, it is necessary to observe the direction of causation from each of education and income to the level of skill, and the indirect effect of health on the level of skill via education and income. Section 5.1 of Chapter 5 presents the results of this analysis.

Different patterns of behaviour may explain differentials in health. While such lifestyle differences make an important contribution to socio-economic differences in health, there may be higher morbidity and mortality in lower socio-economic groups, even after controlling for lifestyle effects. For instance, moderate drinking yields beneficial physical and psychological effects which may have beneficial health effects [Hamilton and Hamilton (1993)].

On the other hand, levels of education and income affect moderate drinking directly and indirectly, through good health, which in turn is an important determinant of drinking behaviour. This implies that two causal relationships potentially exist; moderate drinking affecting health, and health affecting moderate drinking. Other things being equal, the well educated tend to be future-oriented, since the internal rate of return to education exceeds the alternative rate of return or the discount rate [Ehrenberg and Smith (1985)]. For example, the results for males in South Australia in 1968-69 indicate that a profitable investment in higher education is to preserve a discount rate at 5 per cent or 10 per cent and that the private internal rate of return to the bachelor's degree is 13.9 per cent [Blandy and Goldsworthy (1984)]. The social rate of return to the bachelor's degree is 13.6 per cent on the whole investment since age 15 for Australian-born males in 1976 [Miller (1984)]. This suggests that communities which have a high propensity to invest in education would engage in other forward-looking behaviour such as health habits, income, and moderate drinking. However, these beneficial health effects deteriorate as alcohol use increases. Recent Australian research [National Health Strategy, Research Paper No.1 (1992)] uses both poisson and logistic regressions for adults aged 25-64 years in 1989-90, and demonstrates that behavioural or lifestyle factors such as smoking and excessive drinking have detrimental effects on health. Another recent Australian study by Conway, Pinyopusarerk, Carter, Penm, and Stevenson (1993) which uses 1989-90 data from the Australian Institute of Health and Welfare (AIHW) argues that smoking-related disease, alcohol-related disease, and disease related to illicit and licit drug abuse have important effects on morbidity indicators (hospital bed days, hospital separations, medical consultations, scripts, allied health professional referrals, and cost of illness), respectively.

On the other hand, the poor and less educated as well as the young appear to discount the future more heavily. Addicts with higher discount rates respond more to changes in money prices of addictive goods, partly because they generally place a smaller monetary value on health. On the contrary, addicts with lower rates of discount respond more to changes in harmful future consequences of addictive goods, such as negative effects of smoking and excessive drinking on health [Becker, Grossman, and Murphy (1991)]. It follows that a higher future cost (due perhaps to greater information about health hazards) is likely to reduce the demand for cigarettes and alcohol, particularly among the rich. Thus, the issue is whether a higher price for cigarettes and alcohol is likely to have a substantial negative effect on the demand for cigarettes and alcohol which in turn, is likely to have a substantial positive effect on health. More specifically, the analysis takes into account the possibility that what holds for smoking generally tends to hold also for excessive drinking, implying that they complement each other. Therefore, it is necessary to test if excessive drinking is related to smoking and smoking is related to excessive drinking.

Given the importance of the population's health, many have argued that the community would be better off if it substantially reduced tobacco smoking by reducing the proportion of the population who smoke on a regular basis and increasing the proportion of the population who have never smoked, and if it reduced the proportion of persons who are drinking regularly at hazardous or harmful levels [Conway, Pinyopusarerk, Carter, Penm, and Stevenson (1993)]. For example, the campaign against smoking and excessive drinking seeks to improve policy formulation and health outcomes [Collins and Lapsley (1993)]. Section 5.2 of Chapter 5 discusses the empirical analysis of these issues.

Australia was selected for study because of the availability of in-depth data across sixty-one statistical regions. Even though sixty-one statistical regions do not provide sufficient observations to allow a full econometric evaluation of the causal relationships among the selected variables of endogeneity, the Australian data have the advantage of accessible health status indicators (that is, no recent illness, no chronic condition, and self-assessed good or excellent health) by statistical regions. Note that each of these three indicators of health status is likely to be an important determinant of the values of education and income [Lewis, O'Brien, and Thampapillai (1990)].

The plan of this study is as follows. A review of specific studies which have estimated the relationships between health status and education, between health status and income, between education and income, and between lifestyle factors and health status is undertaken in Section 2.1 of Chapter 2. Previous studies of the effects of other inputs in the respective production functions on health, education, income, and lifestyle factors are also reviewed. The analytical frameworks used for estimation are developed in Chapter 3. Data sources are described in Chapter 4. The estimated results of the positive inter-relationships among health status, education, and income are considered in Section 5.1 of Chapter 5. The causal links between health and gender, and the indirect effect of gender on education and income via health, are discussed in Section 5.1 of the chapter. The causal links between education, income and the level of skill, and the indirect effect of health on the level of skill via education and income, are also discussed. In Section 5.2 of Chapter 5, the estimated results of each equation for lifestyle factors in terms of moderate drinking, smoking, and excessive drinking are presented. More specifically, diagnostic double-checkings are conducted to evaluate all the models (in the sense that the more tests that are carried out, the less the chance of accepting a poor model [Beggs (1988)]) and summarised in this chapter. Among those performed are: the non-nested tests, the tests of independence, the tests for endogeneity and exogeneity, the RESET tests for functional form misspecification, the test for autoregression, and the tests for heteroscedasticity. The parameter stability of the model is also tested and, where appropriate, parameter instability analysis is carried out, since all the equations undertaken in this study are identified. The summary of principal findings and the policy implications are presented in Chapter 6.

2.0 REVIEW OF PREVIOUS STUDIES

In this chapter, we review previous empirical studies on the association between each pair of health, education and income, justifying the inter-relationships among these variables; ceteris paribus, an increased level of one is associated with increased levels of the other two. Given the positive effect of moderate drinking and the negative effects of smoking and excessive drinking on health, we also review some earlier studies which examine the proposal that differentials in health, education, and income cause different patterns of these lifestyle behaviours. For each statement, wherever necessary, the qualifications 'ceteris paribus' and 'on average' are to be understood.

2.1 THE INTER-RELATIONSHIPS AMONG HEALTH STATUS, EDUCATION, AND INCOME

This section reviews previous empirical studies within the following structure; the first subsection is on health, the second on education, and the third on income. If there are causal relationships between health and education, between health and income, and between education and income, then we could expect to find that there are multiple inter-relationships among these variables.

2.1.1 Endogenous Health

This subsection examines previous empirical studies which examine the proposal that different levels of education and income cause differentials in health. It also presents previous studies of possible determinants of health such as gender and lifestyle factors in terms of smoking, moderate drinking, and excessive drinking.

Auster, Leveson, and Sarachek (1969) examine the logarithmic form of the production model of health as measured by 1960 mortality rates of the white population in the labour force across 51 states of the United States. Using two-stage least squares they find that, other things being equal, 10 per cent increases in education and the proportion of married females out of the labor force with husband present reduce mortality by 3.1 per cent and 2.8 per cent, respectively, while 10 per cent increases in income and the proportion of workers in white-collar occupations raise mortality by 1.8 per cent and 1.7 per cent, respectively. They also use ordinary least squares, adding the composite of medical capital and the number of paramedical personnel per capita (plant assets), as exogeneous variables and find that ceteris paribus, education has an adverse effect on mortality rates, whereas higher levels of income are associated with higher mortality rates.[1]

Grossman and Benham (1974) utilise the 1963 health interview survey with sample size of 1,049 white males between the ages of 18 and 64 which was conducted by the National Opinion Research Center and the Center for Health Administration Studies of the University of Chicago. They present evidence from the reduced form estimates of the ill health parameter of schooling that a small part of the negative direct effect of schooling on ill health (8.9 percent) is offset by the positive effect of the wage rate on ill health (2.6 percent) and, therefore, that the total effect of schooling on ill health is negative (6.3 percent).

Newhouse and Friedlander (1980) estimate a linear structural equation by the ordinary least squares (OLS) method with the cross-section data from the Health Examination Survey by the United States National Center for Health Statistics which examined 6,672 individuals between the ages of 18 and 79 from October 1959 to December 1962. They find that education and income are positively related to health status as measured by the Abrahamse and Kisch's Health Status Age Index (pp.204-205); an

---

1 Hitiris and Posnett (1992) argue that crude mortality rates are a very imperfect measure of the output of health care. In their regression analysis, per capita total health expenditure has a negative effect on mortality.

additional five years of education or an additional US$12,800 of income makes an individual roughly one year healthier than is average for his or her age group. On the other hand, they find that the estimated elasticity of health status with respect to the number of federal (non-military) hospital beds per 1,000 population and its estimated t-value are 0.006 and 4.08, respectively.

Van De Ven and Van Der Gaag (1982) use individual data for male family heads among 8,000 privately insured households from a health care survey in the Netherlands in 1976 and specify the reduced-form equations of the linear structural models. They obtain the maximum likelihood (ML) estimates which assume normality of the disturbances and conclude that, besides the positive direct effect of education on health (0.017), more education leads to a higher income, which in turn leads to a better health status (0.019).

Narendranthan, Nickell, and Metcalf (1985) use information from the National Training Survey conducted on behalf of the Manpower Services Commission. This provides a unique retrospective longitudinal data set for 17,708 British males in the labour force over the period 1965-75. They find from a logit analysis that the probability of sickness spells is negatively correlated with schooling and that sickness spells fall dramatically as people rise through the social classes, with those unskilled in 1975 being ten times more prone to lengthy sickness spells as professionals in the previous ten years; the lower down the socio-economic scale one goes, the higher the incidence of sickness. Their results are consistent with the Culyer's result (1976) from data of the 1973 General Household Survey - Introductory Report London that semi and unskilled manual workers have a substantially higher incidence of sickness than any other social class in England and Wales for 1971 (males only).

Wagstaff (1986), using data from 2,243 randomly selected individuals between the ages of 20 and 70 from the 1976 Danish Welfare Survey, estimates the health structural demand equation of a pure investment model by using a maximum likelihood (ML)

3 0009 03201178 0

procedure. According to his estimation results, other things being equal, the number of years of education and log of hourly wage variables have a positive influence on the log of good health variable. The coefficient on the urbanization variable is found to be negative.

Van Vliet and Van Praag (1987), using the cross-section data from a survey on living conditions held in 1980 by the Netherlands Central Bureau of Statistics for 2,153 respondents aged 18 years and over, find that, under the ceteris paribus conditions, the levels of education and log of family income raise the unobservable health status index which is increasing if health improves. Classifying education into four levels (i.e., 1 = primary school, 4 = university level), one education-level stands for eight years difference in health ($0.049/0.0063 \approx 8$). The same study suggests that the female variable is highly significant (t-ratio = 5.7) and positive in explaining self-rated good or excellent health status [see also Sindelar (1982) and Umberson (1992)]. A significant positive effect of females on health status may imply that females generally possess more knowledge about health-related issues than men, are more likely to monitor their own health status, and are less likely to engage in a number of risky health behaviours such as excessive alcohol consumption and dangerous sports. In addition, women provide medical services to other members of their families. It might not be highly skilled care, but it is personal and in the right place at the right time. In sum, the higher the proportion of females of a region, ceteris paribus, the more likely it is that its population will be healthy and able to enjoy a full and long life. In our view, the results reported by Björklund (1985) in a cross-section analysis for 6,500 individuals between 15 and 75 years of age from the Swedish Level of Living Survey conducted in the late spring of 1968, 1974, and 1981 support this notion; in a logit analysis, the probability of an increase in mental symptoms is smaller among females than males as unemployment levels increase, ceteris paribus.[2]

[2] Note that Björklund does not make this inference. Nevertheless it is a reasonable interpretation of his empirical results.

Kemna (1987) using ordinary least squares (OLS) estimates of the health production function as an index of self-reported ill health reaches several conclusions. First, an increase in the level of schooling, ceteris paribus, is always significantly related to better health. His results also show that marginal effects of the completion of high school and college on ill health are -0.156 and -0.070, respectively. Second, both metropolitan area and doctor visits positively affect ill health. Kemna suggests that a positive coefficient estimate on the log of the doctor visit variable is due to the bias that results when differences in endowment of health at the beginning of the period are unobserved and when no correction is made, and argues that those people who are in poor health at the beginning of the period see a doctor more often (the marginal effect of doctor visits is equal to 0.0458).[3] His empirical analysis is based on data for 12,907 United States civilians in the labor force from the 1980 Health Interview Survey, combined with occupational information from the Dictionary of Occupational Titles.

Wooden (1990) utilises data for 17,891 dwellings from the 1983 Australian Health Survey by the Australian Bureau of Statistics in conjunction with the monthly population survey. He concludes from a probit analysis that persons in higher income brackets are less likely to suffer an illness as a result of workplace accidents and suggests that they pursue employment in less hazardous jobs. His estimated results also suggest that the probability of workplace accident occurrence is lowest for white-collar workers, that accident rates are lower among females than males, and that persons who reside outside the major metropolitan areas are less likely to suffer an injury as a result of workplace accidents after controlling for inter-occupation and inter-industry differences. Contrary to many previous findings, Wooden's results show that, holding constant all other variables, accident probability rises with education.

---

[3] Suppose that the following semilogarithmic equation is given:
$Y = \beta_0 + \beta_1 \ln X + \text{error}$
where $\beta_1 = (\partial Y/\partial \ln X) = (\partial Y/\partial X) * \overline{X}$.
Then, the impact at margin and elasticity of the estimated coefficient can be obtained, respectively, by:
$\beta_1 * (1/\overline{X})$ and $\beta_1 * (1/\overline{Y})$
where the barred symbols are mean values of the variable.

The Australian research reviewed by McClelland (1991a) suggests that differences in the level of economic resources as measured by income levels and the relationship of income to the Henderson poverty line are related to increased morbidity.

More recently, Australian research [National Health Strategy of Australia, Research Paper No.1 (1992)] finds that for both males and females aged 25-64 years, differences in education, income, and lifestyle factors (such as smoking and excessive drinking) significantly explain differentials in health. The research uses the Poisson regression model for serious chronic illness and the logistic model for self-assessed poor health. They find that men and women with low education have, respectively, 1.23 times and 1.15 times the number of serious chronic illnesses compared to men and women with high education. On the other hand, the odds of reporting self-assessed poor health are, respectively, 81 per cent and 64 per cent higher for men and women with low income compared to men and women with high income. The control variables in each of Poisson and logistic estimates are age, socio-economic index of area, metropolitan location, family composition and workforce status, period of residence, and language spoken at home. The data set is based on the 1985-87 mortality data, the 1988 survey of disability and ageing and the 1989-90 National Health Survey by the Australian Bureau of Statistics, and the 1989 National Heart Foundation Risk Factor Prevalence Survey.

2.1.2 Endogenous Education

In this subsection, we review previous studies on the effects of factors which influence the level of education such as health and income and find their marginal contributions to be substantial.

Blaug (1970) argues that under the conditions of perfect information and perfectly competitive capital markets the lengthening of an individual's life expectancy through improved health raises the returns on investment in her or his education and, conversely, an improvement in her or his productivity through education raises the returns on investment in her or his health. Grossman (1972) notes that poor health increases the amount of time lost from market and non-market activities. These contentions imply that good health facilitates learning.

Fuchs (1986), based on a single cross-section survey for 295 respondents aged 25-64 living in Nassau and Suffolk counties of New York City in November 1979, presents the results of ordinary least squares (OLS) regressions in which number of years of schooling is regressed on self-assessed good or excellent health status when variations in time preference measured as implicit interest rate and expected inflation are controlled for. Additional explanatory variables are age, parents' education, a dummy variable with 1 if the respondent lived with both parents until age 16 and 0 otherwise, two dummy variables with 1 if the respondent is Catholic (or Jewish) and 0 if Protestant or other, and the respondent's scholastic performance in high school.. He finds that self-assessed good or excellent health is positively related to schooling for both males and females. The marginal effects are 0.109 for males and 0.130 for females.

Williams and Carpenter (1990) use data on two national probability samples of Australian age cohorts (born in 1961 and 1965, respectively) - the class of 1978 and the class of 1982 - and provide the connections between completion of Year 12 and family wealth, a weighted composite based on indirect measures (bedrooms, bathrooms, dishwashers, and telephones present in each respondent's home), and between completion of Year 12 and the non-government school systems (Catholic and independent schools). The sample is divided into older respondents (n = 1,213) and younger respondents (n = 1,080). Their results from the logistic regression estimates are identical to those derived from the oreinary least squares (OLS) estimates; other things equal, increases in family wealth and attendance at the non-government schools increase the likelihood of completing Year 12.

On regressing a simple nonlinear transformation of literacy across 22 developing countries in 1985, Anand and Ravallion (1993) find that a 10 per cent increase in average income increases the literacy rate by 9.5 per cent. After controlling for poverty index and social spending on education per capita, a 10 per cent increase in average income increases the literacy rate by 11.2 per cent.

Cook and Moore (1993) utilise data from the National Longitudinal Survey of Youth in 1979 managed by the Human Resource Research Center at Ohio State University with sponsorship from the United States Department of Labor and other federal agencies. Their sample includes 1,904 youths between the ages of 14 and 21. They also make extensive use of the high school subsample, which includes 753 youths who were enrolled in twelfth grade at the time of the 1982 interview. The second stage of the structural equations is estimated by weighted least squares (WLS) to correct for the heteroskedasticity caused by use of the predicted value. They provide the results from the WLS estimates for the High School Senior Sample (753) and show that, other things being equal, the highest year completed is positively related to the log of family income. The estimated income elasticity of schooling equals 0.0031. Cook and Moore also estimate the reduced form of the pairs of structural equations, drinking and school enrolment, and present logit regression results for an alternative dependent variable, an indicator of whether the respondent had graduated from college by 1988. They find that, in the complete sample, family income has a very substantial effect on the likelihood of attaining a college degree; a one unit increase in the log of family income increases the probability of college graduation by 5.8 percent and its standard error is 0.145.[4]

2.1.3 Endogenous Income

The primary purpose of this subsection is to review evidences of the effects of health status, education, skill level, and moderate drinking on income. We review overseas and Australian studies on individual and social returns to education and find additional education to be an important influence upon income for both males and females. There is also

---

[4] In the college graduation logit, the effect of a change in the log of income on P (graduate) equals $ß*P*(1-P)$ where ß denotes the coefficient on the log of income and P the probability of college graduation taken as the sample mean value of P = 0.154, and ß = 0.442 (see Cook and Moore, 1993, p.428).

evidence of positive effects of skill level and moderate drinking, implying for the latter that moderate drinking yields beneficial health effects, which in turn increase income.

Grossman and Benham (1974) employ data for 1,049 white males between the ages of 18 and 64 from the 1963 health interview survey by the National Opinion Research Center and the Center for Health Administration Studies of the University of Chicago, and estimate the reduced-form wage parameter. They decompose it into a direct component and an indirect component, and report that the reduced-form wage parameter of schooling indicates a rate of return to investment in schooling of 7.1 percent. The indirect component, which arises from the fact that schooling raises health and health raises market wage rates, is approximately one-half as large as the direct component.

Wise (1975) investigates the job performance of individuals as measured by salary with a given level of education, but different levels of academic achievement, who work in a particular environment and perform similar tasks. His econometric analysis is based on the 1968 United States data for 976 college graduates among white males hired before 1965 and aged 30 years and less. The estimated results of the parameters obtained by ordinary least squares (OLS) suggest that persons with a bachelor's degree had a higher initial salary than those who started work before obtaining a degree and that the estimated rate of salary increase goes up consistently with college selectivity, college grades, and rank in graduate class. His findings also imply that the level of education is a more important influence upon salary than previous experience; the estimated coefficients of persons with the bachelor's degree and with the previous experience in the monthly salary equation are 0.0259 and 0.0165, respectively.

Brown and Medoff (1978) estimate the earnings functions for 38,065 female workers and 57,067 male workers in private sector with merged 1973-75 Current Population Survey (CPS) files and show regression results which suggest that years of schooling and union membership have significantly positive effects on the log of the hourly wage in real terms for females (t-values = 69.00 and 34.43, respectively) as well as for males (t-values = 64.00 and 52.75, respectively). They deflate the hourly wage to 1973 dollars with hourly wage data for production workers provided by the U.S. Bureau of Labor Statistics. Additional variables in the earnings functions are age, age squared, and a vector of region dummies.

Van De Ven and Van Der Gaag (1982) utilise data for male family heads among 8,000 privately insured households from a Health Care Survey (HCS) in the Netherlands in 1976 and attempt to estimate the direct and indirect effects of education and income. They use a linear structural model and present the maximum likelihood (ML) estimates, together with comparable results found by Grossman, concluding that besides the positive direct effect (0.037), more education leads to a better health status, which in turn leads to a higher income (0.002). Van De Ven and Van Der Gaag specify income (permanent income in their context) as a proxy for someone's lifestyle or quality of life. On the other hand, Grossman's results show that the direct and indirect effects are 0.052 and 0.003, respectively. In both studies the indirect effect of education (via health status) on earnings is only a small fraction (5%) of the total effect.

Podgursky (1983) estimates the wages of union and nonunion workers using a subsample of 13,413 private sector, full-time, year-round production workers from the March 1971 United States Current Population Survey, of whom 42 per cent are union workers. He suggests from ordinary least squares (OLS) estimates that employers of nonunion workers in heavily unionised industries often raise wages in tandem with the unionised segment of the industry as a defensive strategy to prevent unionisation, that the wages of both union and nonunion workers are significantly and positively related to the level of education, and that the level of education has a more important effect on wage than previous experience.

Blandy and Goldsworthy (1984) conducted a study of the private returns to education for males aged 15 to 70 years in South Australia for the financial year 1968-69. They utilised the data for 1,386 males aged 30 to 70 years of a sample of children in third, fourth, and fifth years of secondary government and private schools from the South Australian Education Department for State school teachers at 1971 and from the Australian Research Grants Committee (ARGC) survey undertaken by Blandy, Goldsworthy, and Hancock during 1970 and the data for younger males aged 15 to 30 years from the various sources (for the sources see p.203). They argue that a profitable investment in higher education would generate a discount rate between 5 per cent and 10 per cent (p.198). They estimate the private internal rates of return to secondary schooling and the bachelor's degree are, respectively, 14.0 per cent and 13.9 per cent (p.200). It can be concluded from their results that education in Australia (judging by the South Australian evidence) pays handsome private financial dividends, both at secondary and university levels.

Using the data from the 1976 Australian Census of Population and Housing, Miller (1984) calculates the rate of return to education since age 15 for Australian-born males and females in 1976. According to his findings, the internal rates of return to the diploma and bachelor's degree are 11.1 per cent and 13.5 per cent for males, and 9.9 per cent and 12.4 per cent for females (p.217). On the other hand, the social rates of return to the diploma and bachelor's degree are estimated to be 11.0 per cent and 13.6 per cent for males, and 10.1 per cent and 12.5 per cent for females (p.219). His results show that the internal rates of return to education exceed the return which can be earned on alternative investments available (5 per cent or 10 per cent); as expected, better educated persons are observed to raise their earnings. Miller's study also implies that education is a profitable investment from the viewpoint of society. In summary, additional education, whether it be acquired in the schools system or at a tertiary institution, is an important influence upon income.

Chirikos and Nestel (1985) use cross-sectional data from the National Longitudinal Surveys (NLS) of white men, black men, white women, and black women in the high health-risk age category 45-64 years of age in the United States in 1976 and 1977. They calculate the net effects of health history (measured in reference to a ten-year period (1966-1976 or 1967-1977) longitudinal profile or retrospective history of self-reported health appraisals, disability limitations, and functional impairments) on annual earnings and family income. Their findings show that the average reduction in individual income as measured by annual earnings for any history of poor health represents a loss of 12.5% to 27.8% of the earnings reported by the continuously healthy for each of four subgroups and that a history of poor health reduces family income from 6.0% to 21.7% for each subgroup. Chirikos and Nestel also find that continuously poor health reduces family income more than individual income.

Boissiere, Knight, and Sabot (1985) estimate a cross-section human capital earnings function for Kenya and Tanzania in East Africa. Their econometric analysis is based on the data for 205 employees in Kenya and 179 employees in Tanzania from comparable surveys administered by a team (including the authors) within a few months of each other in 1980. Pre-tax earnings of the employees is used, in logarithmic form, as the measure of earnings, primary or secondary school completers as the measure of education (a dichotomous variable), the literacy and numeracy test score as the measure of cognitive achievement, and the test score of reading achievement and mathematics as the measure of reasoning ability. Their estimated results of the semilogarithmic earnings function reveal that the coefficient on secondary education (0.476) is positive and large relative to the coefficient on the number of years of employment experience (0.042) in Kenya and also in Tanzania (0.280 and 0.054, respectively). Even when the achievement score and the ability score are held constant, the positive effect of secondary education on earnings exceeds that of employment experience in both countries. In short, educational qualification or years of schooling has a more important effect on earnings than employment experience. They also find from the mean characteristics of workers by earnings quintile that in both countries the contribution of years of education is large, reflecting the size of its coefficient and the tendency for the proportion with secondary education to rise with their earnings quintile (p.1025).

Wagstaff (1986) obtains the ordinary least squares (OLS) estimates of the semilogarithmic wage equation, treating chronic ill health as predetermined. Using data for 2,243 individuals between the ages of 20 and 70 from the 1976 Danish Welfare Survey, he finds that, once sex, job tenure, local unemployment, industry unemployment, age, and age squared are controlled, the effect of years of schooling on wage rates is significantly positive (t-value = 13.98), whereas the effect of chronic ill health is significantly negative (the absolute t-value = 2.46).

Berger and Leigh (1988) employ data for 924 males and 440 females living in the United States aged 18 years old or over who were working for pay for 20 or more hours per week from the 1972-73 Quality of Employment Survey and examine the hypothesis that the moderate drinkers received the largest wage premiums. In a log wage regression on dummy variables indexing frequency of drinking, they find from ordinary least squares (OLS) estimates that both males and females who have at least one drink per week earn significantly higher wages than nondrinkers and conclude that the beneficial health effects of moderate drinking carry over to the labour market, raising productivity and wages (p.1349).

Ashenfelter (1993) analyses the intra-family schooling-income relations studied with Zimmerman. Their findings are based on cross-sectional data for 332 father-son pairs in the National Longitudinal Survey of the United States, and suggest that more education for the son increases the son's income regardless of the father's education level (the rate of return to schooling is about 5%).

Acs and Danziger (1993) estimate linear probability models on 11,322 men in 1979 and 11,094 men in 1989 from the U.S. Bureau of the Census:1992 and perform a parallel decomposition using the means and coefficients from the models. They find that increasing educational attainment partially offsets the rise in low earnings rates by shifting patterns of industrial employment; that is, the growth in low earnings would have been 34 per cent greater for whites had educational attainment not increased. For blacks, low earnings rates would have been almost 75 per cent higher if education had remained constant. Hispanics at the low end of the earnings distribution experienced much less educational upgrading, and this reduced the incidence of low earnings by only 0.6 percentage points.

In order to test the hypothesis that moderate drinking yields beneficial health effects which lead to increased productivity in the labour market, Hamilton and Hamilton (1993) categorise non- and moderate drinkers as those who drank less often than once a month or not at all over the previous year and those who drank at least once a week or everyday and drink 8 or fewer drinks on a single day in the previous week, respectively, from the 1985 General Social Survey (GSS) Health and Social Support collected by Statistics Canada for 1,954 males between the ages of 20 and 59. They perform wage decomposition and find the mean wage differentials between moderate and non-drinkers and between moderate and heavy drinkers to be 0.067 and 0.088, respectively. This suggests that the mean wage of moderate drinkers is higher than the mean wages of both non-drinkers and heavy drinkers. On the other hand, Hamilton and Hamilton also measure the differences in the productivity between moderate and non-drinkers, and between moderate and heavy drinkers to be 0.963 and 0.555, respectively. This suggests that the returns to moderate drinking are greater than the returns to non-drinking and heavy drinking. An implication of this is that moderate drinking leads to productivity gain via a beneficial health effect, which increases earnings in the labour market (p.13). In short, moderate drinking appears to yield a substantial positive effect on earnings. They also provide results for the selection corrected wage regressions which show that even when holding age, marital status, region, industry, religious attendance, and other characteristics constant, higher education and the skill level (managerial and professional) have positive impacts upon earnings for non, moderate, and heavy drinkers.

Conclusions reached in this section allow the following summary remarks:

1. Evidence from the studies reviewed suggest that both education and income make a substantial contribution in generating health improvements. Evidence alÿo suggests that different patterns of lifestyle factors cause differentials in health; moderate drinking yields beneficial health effects, while smoking and excessive drinking jeopardise good health. It has also been suggested that the higher the proportion of females in a region, ceteris paribus, the more likely it is that its population will be healthy and able to enjoy a full and long life.

2. Evidence from these studies suggest that both health and income make a substantial contribution to the level of education.

3. Evidence from the aforementioned studies suggest that better health enhances income and that additional education is an important influence upon income via an increase in productivity. The evidence also imply that the beneficial health effects of moderate drinking carry over into the labour market, raising productivity and income. It has also been suggested that the level of skill has a positive effect on income.

In brief, evidence reviewed here has revealed that there are inter-relationships among health, education, and income; ceteris paribus, an increased level of one is associated with increased levels of the other two. Therefore, this study attempts to examine the multiple inter-relationships among these variables using in-depth data from sixty-one statistical regions throughout Australia provided by the Australian Bureau of Statistics (ABS).

2.2 LIFESTYLE FACTORS - SMOKING AND DRINKING BEHAVIOURS

It has been suggested that different patterns of personal behaviour contribute to the explanation of differentials in health. For instance, moderate drinking yields beneficial physical and psychological effects which have beneficial health effects [Hamilton and Hamilton (1993)]. If education and income affect moderate drinking directly and indirectly via good health, then we could expect to find that two causal relationships potentially exist; moderate drinking affecting health, and health affecting moderate drinking. However, these beneficial health effects deteriorate as alcohol use increases, together with a negative effect of smoking on health [National Health Strategy of Australia, Research Paper No.1 (1992), and Conway, Pinyopusarerk, Carter, Penm and Stevenson (1993)]. Given these effects of lifestyle factors on health, this section reviews previous empirical studies suggesting that differentials in health, education, and income cause different patterns of lifestyle behaviour.

Berger and Leigh (1988) employ data for 924 males and 440 females living in the United States aged 18 years old or over who are working for pay for 20 or more hours per week from the 1972-73 Quality of Employment Survey, and report probit estimates of the drinking decision equation for males and females. Their results show that for both males and females higher education raises the probability of drinking when holding race, marital status, union membership status, experience, occupation, city, and so forth constant.

Kenkel (1991) uses data from the Health Promotion/Disease Prevention (HPDP) supplement to the 1985 Health Interview Survey, a continuing survey of the civilian noninstitutionalised population of the United States, providing sample sizes of 14,177 males and 19,453 females. He finds from a tobit analysis that, after differences in health knowledge are controlled for, education has significant positive effects on the total number of drinks in past two weeks for both males and females. The estimated coefficients and t-values are, respectively, 0.762 and 10.873 for males, and 0.766 and 12.876 for females.

Umberson (1992) performs regression analysis using data for 1,688 males and 1,897 females aged 24 and older of the United States population from a national two-wave panel survey, Americans' Changing Lives, conducted in 1986 and 1989 under the auspices of the Institute for Social Research of The University of Michigan. She measures the number of drinks by the number of days in the past month on which the respondent drank, multiplied by the number of drinks the respondent usually drinks on days that he/she drinks. Income refers to a ten-category variable; cases within each category are assigned the midpoint value of that category. These values are then divided by 1,000. Values range from 2.5 (indicating a midpoint value of $2,500) to 110 (indicating a midpoint value of $110,000). From the results of ordinary least squares (OLS) regressions, Umberson finds that income has positive effects on the number of drinks for both males and females (significant at the 0.10 and 0.05 levels, respectively). The estimated coefficients are 0.062 for males and 0.086 for females.

Cook and Moore (1993) utilise data from the National Longitudinal Survey of Youth in 1979 managed by the Human Resource Research Center at Ohio State University with sponsorship from the United States Department of Labor and other federal agencies. Their sample includes 1,904 youths between the ages of 14 and 21. They also make extensive use of the high school subsample, which includes 753 youths who were enrolled in twelfth grade at the time of the 1982 interview. Cook and Moore use tobit analysis for the high school senior sample (753) and find that, other things being equal, the log of family income is positively and significantly related to drinking (number of drinks in the preceding week). The estimated coefficient and standard error are 2.545 and 1.141, respectively.

Hamilton and Hamilton (1993) utilise cross-section data for 1,954 males between the ages of 20 and 59 from the 1985 General Social Survey (GSS) Health and Social Support collected by Statistics Canada. They categorise non-drinkers as those who drank less often than once a month or not at all over the previous year and moderate drinkers as those who drank at least once a week or everyday and drank 8 or less drinks on a single day in the previous week. Hamilton and Hamilton provide the results of the multinomial logit regression for non- and moderate drinkers. Their findings suggest that self-assessed poor health has a positive impact on the probability of being a non-drinker and that higher education has a negative effect on the propensity for individuals to be non-drinkers as opposed to moderate drinkers (p.14). Differences in occupation, age, marital status, region, industry, religious attendance, and other characteristics are controlled for.

In sum, evidence from the aforementioned studies suggests that health is one of the important determinants of lifestyle factors in terms of moderate drinking. It has been suggested from the previous studies that moderate drinking yields beneficial health effects, while smoking and excessive drinking deteriorate health. Therefore, this study attempts to examine the proposition that two causal relationships exist; moderate drinking affecting health, and health affecting moderate drinking. On the other hand, evidence from the aforementioned studies suggests that education and income have positive effects on lifestyle factors in terms of moderate drinking, smoking, and excessive drinking. Given two causal relationships between health and moderate drinking, we can expect that education and income affect moderate drinking directly and indirectly via good health. Further, given the importance of health, smoking and excessive drinking may be one package in policy formulation. Thus, this study also attempts to examine two causal relationships: excessive drinking affecting smoking, and smoking affecting excessive drinking.

3.0 ANALYTICAL FRAMEWORK

This chapter first examines Lewis' proposition that a set of three health indicators (viz., no recent illness, no chronic condition, and self-assessed good or excellent health), education, and income are positively inter-related. If moderate drinking reduces the levels of education and income, we reject the hypothesis that differences in education, income, and three health indicators are manifestations of moderate drinking. We also examine if excessive drinking has detrimental effects on three health indicators. The first specification of the model is the strongest generalisation.

The second specification of the model examines the possibility that the positive correlations among education, income, and moderate drinking develop only after the hypothesis that moderate drinking yields beneficial physical and psychological effects which have beneficial health effects. This implies that the mechanism behind the education-income-moderate drinking relationships may also give rise to the education-income-a set of two health indicators (viz., no recent illness and self-assessed good or excellent health) relation. Assume further that more education and higher income affect moderate drinking through better health. Thus, we examine the positive correlation between a set of two health indicators and moderate drinking. The second specification of the model is influenced by the first specification of the model.

The first specification of the model examines if excessive drinking has detrimental effects on the three good health indicators. The second specification of the model examines if both smoking and excessive drinking have detrimental effects on the two good health indicators, either directly or indirectly. The detrimental total (indirect) effects of smoking on two health indicators via moderate drinking suggest that smoking has a negative effect on moderate drinking. This implies that the negative correlations between moderate drinking and excessive drinking develop only after the observation that moderate drinking is negatively associated with smoking. Thus, the mechanism behind the excessive drinking-moderate drinking relationships may also give rise to the excessive drinking-smoking relationship.

Given the importance of health, smoking and excessive drinking may be considered as a package in policy formulation and health outcomes. In order to comply with our argument, the third specification of the model attempts to examine if smoking and excessive drinking are positively inter-related. This then suggests that differences in both smoking and excessive drinking are causal to differentials in three indicators of health. The third specification of the model also examines the possibility that increases in education and income increase both smoking and excessive drinking. We then describe some of the policy implications.

Assume that excessive drinking is a causal factor for more education and higher income in the first specification. Assume further that the negative indirect effects of excessive drinking on education and income via three health indicators offset the positive direct effects of excessive drinking on education and income. We can then reject the hypothesis that differences in both education and income are causal to excessive drinking differences in favour of the existence of smoking. This suggests that the third specification of the model is influenced by both the first and second specifications.

3.1 THE INTER-RELATIONSHIPS AMONG HEALTH STATUS, EDUCATION, AND INCOME

The primary purpose of this section is to specify Lewis' proposition of the multiple inter-relationships among health status, education, and income; ceteris paribus, an increased level of one is associated with increased levels of the other two. The relevance of this framework is that the above proposition suggests a full simultaneous equations model. An implication of this is that instrument choice should be based upon stability analyses among these target variables. Distinguishing three indicators of health status (viz. no recent illness, no chronic condition, and self-assessed good or excellent health), inter-relationships among these variables in describing the characteristics of the stock of individual statistical region i can be written, in general form, as:

$$H1_i = H\,(E_i, Y_i, F_i, \text{REGRESSORS}_i) \quad \textbf{(1)}$$

$$H2_i = H\,(E_i, Y_i, F_i, \text{REGRESSORS}_i) \quad \textbf{(2)}$$

$$H3_i = H\,(E_i, Y_i, F_i, \text{REGRESSORS}_i) \quad \textbf{(3)}$$

$$E_i = E\,(H1_i, H2_i, H3_i, Y_i, \text{REGRESSORS}_i) \quad \textbf{(4)}$$

$$Y_i = Y\,(H1_i, H2_i, H3_i, E_i, \text{REGRESSORS}_i) \quad \textbf{(5)}$$

where REGRESSORS denote other exogenous variables in the respective equations. Table 3.1 provides a description of the variables used in the model as well as their means and standard deviations.

Equations (1) - (5) stand for well-behaved production functions exhibiting everywhere diminishing returns to inputs. Health, education, and income are complementary to each other and therefore health and education can be viewed as normal goods. This means that the cross effects among these variables are positive; the second cross partial derivatives are positive (e.g., $\partial^2 H_3/\partial E \cdot \partial Y$, $\partial^2 E/\partial H_3 \cdot \partial Y$, and $\partial^2 Y/\partial H_3 \cdot \partial E > 0$).

In equations (1) and (2), the positive partial effects of education and of income on health reveal that regions with a relatively higher proportion of persons holding post-school qualifications and with relatively higher income may be able to produce better health for children as well as for adults.

A schematic view (Figure 3.1) presents the logical linkages among the null hypotheses given in equations (1)-(5). It also depicts the gender distribution affecting three indicators of health directly and thereby each of education and income indirectly. The direct effect captures the impact of a change in the gender distribution on each of three indicators of health. The indirect effects measure the effect on education and income of a change in the gender distribution, due to a change in each of three indicators of health. For example, the gender distribution affects education indirectly through three indicators of health, and through income via three indicators of health.

Table 3.1. Definition of Variables

		Variable	Mean	Standard Deviation
Endogenous Variables[1]				
H_1	=	Proportion of persons reporting no recent illness in two weeks before interview to population	29.16	3.75
H_2	=	Proportion of persons reporting no long-term chronic illness to population	34.07	7.45
H_3	=	Proportion of persons reported being in good or excellent health by their perception to population aged 18 years or over	78.99	3.87
E	=	Proportion of persons with post-school qualifications to population aged 15 years or over	47.06	6.93
Y	=	Nominal gross annual median income of persons aged 15 years or over ($ '000)	16.19	3.05
R^2	=	Proportion of persons with risk factors to population aged 18 years or over subcategorised as		
		R_1 = proportion of smokers	28.57	4.19
		R_3 = proportion of moderate drinkers	51.23	4.93
		R_4 = proportion of excessive drinkers	11.26	3.23
F	=	Proportion of females to population	50.09	1.30
OCC_1^3	=	Proportion of managers and administrators and professionals to all employees aged 15 years or over	31.34	7.29
Explanatory Variables				
R_2^2	=	Proportion of non-drinkers to population aged 18 years or over; $R_2 + R_3 + R_4 = 100$	37.51	5.64
OCC_2^4	=	Proportion of unskilled and semi-skilled workers to all employees aged 15 years or over	21.78	5.69
C	=	Proportion of persons who reported doctor (general practitioners and specialists) consultations in two weeks before interview to population	19.83	3.04
B_1^5	=	Rate of acute care private hospital beds	1.26	1.28
B_2^5	=	Rate of acute care public hospital beds	4.02	3.28
B^5	=	Rate of acute care hospital beds; $B = B_1 + B_2$	5.28	4.19

Table 3.1. (Continued)

		Variable	Mean	Standard Deviation
METD	=	1 for capital cities in New South Wales, Victoria, Queensland, South Australia, Western Australia, Tasmania and Australian Capital Territory, and 0 elsewhere including Northern Territory		
G_S^1	=	Nominal state and local governments outlays on health per person aged 15 years or over by States and Territories ($ '000)	1.82	0.15
SD	=	State dummies defined by SD_1 = 1 for statistical regions in New South Wales (23), SD_2 = 1 for statistical regions in Victoria (12), SD_3 = 1 for statistical regions in Queensland (11), SD_4 = 1 for statistical regions in South Australia (6), SD_5 = 1 for statistical regions in Western Australia (6), SD_6 = 1 for Tasmania (1), SD_7 = 1 for Northern Territory (1) and SD_8 = 1 for Australian Capital Territory (1), and 0 elsewhere		
AGE	=	Proportion of persons by age groups		
		AGE_1 = proportion of persons aged 0-4	7.35	1.75
		AGE_2 = proportion of persons aged 5-14	14.60	3.04
		AGE_3 = proportion of persons aged 15-24	16.13	2.21
		AGE_4 = proportion of persons aged 25-44	31.60	3.09
		AGE_5 = proportion of persons aged 45-64	19.03	2.60
		AGE_6 = proportion of persons aged 65-74	7.15	2.19
		AGE_7 = proportion of persons aged 75 and over	4.14	1.79
		AGE_8 = proportion of persons aged 25-74	57.78	3.43
FAGE	=	Proportion of females to total females by age group		
		$FAGE_3$ = proportion of females aged 15-24	15.85	2.45
		$FAGE_4$ = proportion of females aged 25-44	31.28	2.98
		$FAGE_5$ = proportion of females aged 45-64	18.65	2.83
NES[6]	=	Proportion of persons born in non-English speaking countries who arrived before 1984	9.37	6.85
NGSTD	=	Proportion of students in non-government schools to total students	28.41	9.84
UD[1]	=	1 if university is located in region and 0 elsewhere		
AE	=	Proportion of persons with qualification prior to 1981 to all persons with post-school qualification	47.45	5.08

Table 3.1. (Continued)

		Variable	Mean	Standard Deviation
K_S^7	=	Nominal gross capital stock per person aged 15 years or over by States and Territories ($ '000)	13.69	0.83
Lu	=	Proportion of trade union members to population aged 15 years or over	29.36	5.56
Ln	=	Proportion of non-union members to population aged 15 years or over	29.19	5.94
I	=	Proportion of employees in all industries to population aged 15 or over; I = Lu + Ln	58.55	5.29
PFL	=	Proportion of part-time workers to full-time workers	27.38	6.43
GNG^8	=	Proportion of employees to population aged 15 years or over by industry sector		
		GNG_1 = Division A (agriculture, forestry, fishing, and hunting)	3.24	4.28
		GNG_2 = Division B (mining)	0.79	1.64
		GNG_3 = Division C (manufacturing)	8.80	3.76
		GNG_4 = Division D (electricity, gas, and water)	0.67	0.82
		GNG_5 = Division E (construction)	4.62	1.71
		GNG_6 = Division F (wholesale and retail trade)	10.89	2.12
		GNG_7 = Division G (transport and storage)	3.03	1.24
		GNG_8 = Division H (communication)	0.94	0.64
		GNG_9 = Division I (finance, property, and business services)	6.57	3.38
		GNG_{10} = Division J (public administration and defence)	3.19	2.75
		GNG_{11} = Division K (community services)	10.52	2.71
		GNG_{12} = Division L (recreation, personal and other services)	4.52	1.71
		$\sum_{j=1}^{12} GNG_j = I$		

Notes:
1. For the detailed descriptions see Chapter 4.
2. Defined in 1989-90 National Health Survey Users' Guide, ABS Cat.No.4363.0, 1991, pp.33-41,45.
3. Coded to the 4 digit level of the Australian Standard Classification of Occupations (ASCO), ABS Cat.No.1222.0.
4. ASCO Major groups 7 and 8 noted by Forster (South Australian Health Commission, 1991, p.32).
5. Rate is the number of beds per thousand population.
6. Classified in the Australian Standard Classification of Countries for Social Statistics (ASCCSS), ABS Cat.No.1269.0.
7. A detailed discussion of capital stock and asset types can be found in Australian National Accounts, ABS Cat.No.5221.0.
8. Coded to the Division level of the Australian Standard Industrial Classification (ASIC), ABS Cat.No.1201.0.

All the equations are assumed to be identified. The following equation in its reduced-form can be derived:

$$V_i = V\,(F_i,\ REGRESSORS_i) \qquad \textbf{(6)}$$

where $V_i = (H_{1_i}, H_{2_i}, H_{3_i}, E_i, Y_i)'$ refers to the row vector of the dependent variables and where F_i and $REGRESSORS_i$ refer to the column vector of F_i and all other explanatory variables in a reduced-form.

The linear structural equations on the proportion of persons with each of no recent illness (H_1) and no chronic conditions (H_2), and the proportion of persons aged 18 years or over with self-assessed good health(H_3) in region i are defined, respectively, by:[5]

$$\begin{aligned} H_{1i} = {} & a_0 + a_1E_i + a_2Y_i + \sum_{j=2}^{3} a_{1+j}\,R_{ji} + a_5F_i + a_6OCC_{1i} + a_7C_i + a_8B_{1i} + a_9B_i \\ & + a_{10}G_s{*}VIC_i + a_{11}G_s{*}QLD_i + a_{12}G_s{*}WA_i + a_{13}G_s{*}TAS_i + a_{14}G_s{*}(SA_i{+}NT_i) \\ & + a_{15}METD_i + a_{16}GNG_{8i} + a_{17}GNG_{9i} + a_{18}GNG_{10i} + Error \qquad \textbf{(7)} \end{aligned}$$

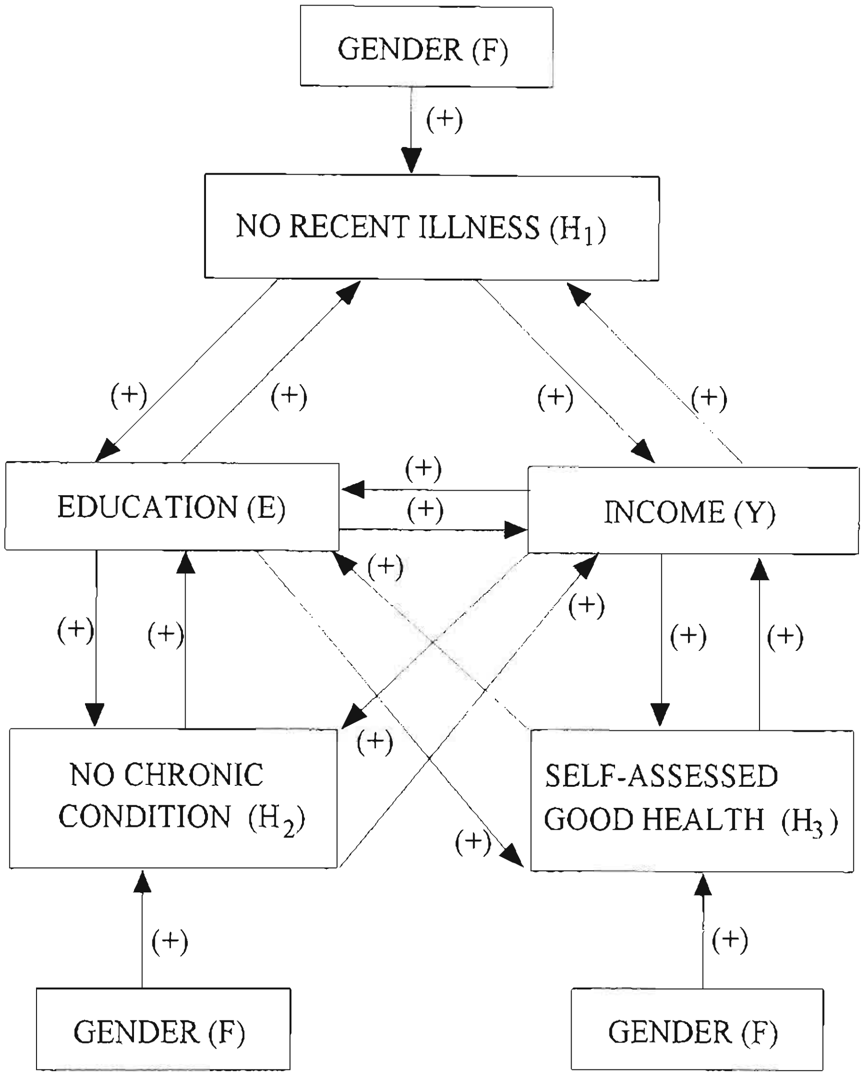
$$\begin{aligned} H_{2i} = {} & b_0 + b_1E_i + b_2Y_i + \sum_{j=2}^{3} b_{1+j}\,R_{ji} + b_5F_i + b_6OCC_{1i} + b_7C_i + b_8B_{1i} + b_9B_i \\ & + b_{10}G_s{*}VIC_i + b_{11}G_s{*}QLD_i + b_{12}G_s{*}WA_i + b_{13}G_s{*}TAS_i \\ & + b_{14}G_s{*}(SA_i{+}NT_i) + b_{15}METD_i + b_{16}GNG_{8i} + b_{17}GNG_{9i} + b_{18}GNG_{10i} \\ & + Error \qquad \textbf{(8)} \end{aligned}$$

$$\begin{aligned} H_{3i} = {} & c_0 + c_1E_i + c_2Y_i + \sum_{j=2}^{3} c_{1+j}\,R_{ji} + c_5F_i + c_6OCC_{1i} + c_7C_i + c_8G_s{*}VIC_i \\ & + c_9G_s{*}QLD_i + c_{10}G_s{*}WA_i + c_{11}G_s{*}TAS_i + c_{12}G_s{*}(SA_i{+}NT_i) + c_{13}METD_i \\ & + c_{14}GNG_{8i} + c_{15}GNG_{9i} + c_{16}GNG_{10i} + Error \qquad \textbf{(9)} \end{aligned}$$

---

5 The linear regression model is extensively used by applied econometricians (MacKinnon and White, 1985, p.305). On the contrary, Parkin, McGuire, and Yule (1987, p.118), and Gerdtham and Jönsson (1991, p.230) suggest that the log-linear form is preferred on econometric grounds. On the other hand, in most of the cases examined by Hitiris and Posnett (1992, p.176) there is little to choose between the two. On that basis, the linear form of a model against the double logarithmic form is tested as reported in Table 5.1, indicating that the null hypothesis of linear model is accepted.

Figure 3.1. Inter-relationships Among Three Indicators of Health, Education, and Income: Effect of Gender



Notes: **1.** Arrows relevant to causation direction from one to the other.

2. Predicted directions in parentheses.

where a_1, a_2, b_1 and $b_2 > 0$ suggest that communities with a relatively higher proportion of persons aged 15 years or over holding post-school qualifications and with relatively higher income may be able to produce better health for children as well as for adults. a_4, b_4, and $c_4 > 0$ suggest that an increase in the proportion of persons aged 18 years or over with moderate alcohol consumption raises the proportion of persons with three indicators of good health (viz. no recent illness, no chronic condition, and self-assessed good health). a_4-a_3, b_4-b_3, and c_4-$c_3 > 0$ suggest that a lower proportion of persons aged 18 years or over with excessive alcohol consumption is associated with a higher proportion of persons with three indicators of good health. a_5, b_5, and $c_5 > 0$ suggest that the higher the proportion of the population which is female in a region, ceteris paribus, the more likely it is that its population will be healthy and able to enjoy a full and long life.

In the structural equations (7), (8), and (9), rate of acute care hospital beds (b) is used as a control variable, since this is equivalent to endogenous variable measured as the proportion.[6]

In order to examine the role of state government intervention in the health sector, a proxy variable for nominal state and local government outlays on health per person by states and territories denoted G_s * $STATES_{ji}$ is introduced in the structural equations. Such treatment reflects the data limitation in the absence of surveys of public health expenditures by the statistical regions and allows us to compare seperately the six different states and territories with New South Wales, the most populous of the six states and two territories,

---

[6] For example, the detrimental effect of excessive drinking on no recent illness is obtained as follows;

$H_1 = a_0 + a_3R_2 + a_4R_3$ where $R_2 + R_3 + R_4 = 100$ (1)

$R_2 = 100 - R_3 - R_4$ (2)

and

$R_3 = 100 - R_2 - R_4$ (3)

Substituting (2) and (3) into the structural equation (1) yields, respectively,

$H_1 = (a_0 + 100a_3) - (a_3 - a_4) R_3 - a_3R_4$ (4)

and

$H_1 = (a_0 + 100a_4) R_2 - a_4R_4$ (5)

Thus, $\partial H_1/\partial R_4 = -(a_3 + a_4)/2$ from (4) and (5)

By the same token,

$\partial H_2/\partial R_4 = -(b_3 + b_4)/2$ and $\partial H_3/\partial R_4 = -(c_3 + c_4)/2$

Note that $\partial H_1/\partial R_2$ and $\partial H_1/\partial R_3$ are a_3 and a_4, respectively.

This is justified as follows,

$\partial H_1/\partial R_2 + \partial H_1/\partial R_3 + \partial H_1/\partial R_4 = a_3 - a_4 - (a_3 - a_4)/2 = (a_3 - a_4)/2 = \partial H_1/\partial R_4$.

Thus, $\partial H_1/\partial R_2 + \partial H_1/\partial R_3 = -2^* \partial H_1/\partial R_4 = a_3 + a_4$.

(South Australia and Northern Territory are combined). For example, a_{10} and $a_{11} < 1$ suggest that the role of state government intervention in the health sector is greater in New South Wales than in Victoria and Queensland. $a_{12} < 4$, $a_{13} < 8$, and $a_{14} < 18$ suggest that the role of government intervention is greater in Queensland than in South Australia, Western Australia, Tasmania, and Northern Territory. Therefore, $a_{10} < 1$, $a_{11} < 1$, $a_{12} < 4$, $a_{13} < 8$, and $a_{14} < 18$ suggest that the role of government intervention is greater in New South Wales than in Victoria, Queensland, South Australia, Western Australia, Tasmania, and Northern Territory. Further, $-(a_{11} + a_{12} + a_{13} + a_{14}) < 2$ suggests that the role of government intervention in the health sector in terms of no recent illness is greater in New South Wales than in Australian Capital Territory.[7]

---

7 In order to examine the role of state government intervention in the health sector, an interaction variable, $G_sSTATES_{ji}$, is introduced [Giannaros (1985, p.15); Let
$H_1 = \beta_0 + \beta_1 G_s{*}SD_1 + \beta_2 G_s{*}SD_2 + \beta_3 G_s{*}SD_3 + \beta_4 G_s{*}SD_4 + \beta_5 G_s{*}SD_5 + \beta_6 G_s{*}SD_6 + \beta_7 G_s{*}SD_7 + \beta_8 G_s{*}SD_8 + \text{Error}$ (1)
where SD_j=state dummies for statistical regions in the eight states and territories; j=1,2,...,8. Including the eight dummies would be introducing perfect multicollinearity. Thus, a dummy for New South Wales, the most populous of the six states and two territories, is treated as a base; SD_1=1. Further, South Australia and Northern Territory are combined [Williams and Carpenter (1990, p.12) and Deeble (1991, p.20)]. Let NSW=SD_1=1. Then the STATES variables are measured as;
VIC=SD_1+SD_2=1+SD_2, QLD=SD_1SD_3=1+SD_3, SA=NSW+QLD+SD_4=2+SD_3+SD_4, WA=NSW+QLD+SA+SD_5, TAS=NSW+QLD+SA+WA+SD_6=8+4SD_3+2SD_4+SD_5+SD_6, NT=NSW+QLD+SA+WA+TAS+SD_7=16+8SD_3+4SD_4+2SD_5+SD_6+SD_7, ACT=NSW+QLD+SA=WA=TAS=NT+SD_8=32 +16SD_3+8SD_4+4SD_5+2SD_6+SD_7+SD_8, and SA+NT=18+9SD_3+5SD_4+2SD_5+SD_6+SD_7 (2)
Note that QLD is arbitrarily chosen to compare with SA, WA, TAS, NT, and ACT. On the other hand, ACT-[QLD+WA+TAS+(SA+NT)]=NSW+SD_8=1+SD_8. If ACT is excluded from the estimations (i.e., ACT=0), then -[QLD+WA+TAS+(SA+NT)]=1+SD_8. From these measurements the structural equation (1) is reformulated as;
$H_1 = a_0 + a_1 G_s{*}VIC + a_2 G_s{*}QLD + a_3 G_s{*}WA + a_4 G_s{*}TAS = a_5 G_s{*}(SA+NT) + \text{Error}$ (3)
This allows one to compare separately the seven different states and territories with New South Wales; substituting (2) into (1) yields
$H_1 = a_0 + (a_1 + a_2 + 4a_3 + 8a_4 + 18a_5)\, G_s{*}SD_1 + a_1 G_s{*}SD_2 + (a_2 + 2a_3 + 4a_4 + 9a_5)\, G_s{*}S + (a_3 + 2a_4 + 5a_5)\, G_s{*}SD_4 + (a_3 + a_4 + 2a_5)\, G_s{*}SD_5 + (a_4 + a_5)\, G_s{*}SD_6 + a_5 G_s{*}SD_7 + \text{Error}$ (4)
Hence, if $a_1 + a_2 = 0.9$, $a_3 = 3$, $a_4 = 7$, and $a_5 = 15$, then $\beta_1 = a_1 + a_2 + 4a_3 + 8a_4 + 18a_5 = 339.8$, $\beta_2 = a_1 = 0.9$, $\beta_3 = a_2 + 2a_3 + 4a_4 + 9a_5 = 169.9$, $\beta_4 = a_3 + 2a_4 + 5a_5 = 92$, $\beta_5 = a_3 + a_4 + 2a_5 = 40$, $\beta_6 = a_4 + a_5 = 22$, and $\beta_7 = a_5$ 15. thus, if a_1 and $a_2 < 1$, $a_3 < 4$, $a_4 < 8$, $a_5 < 18$, then it is suggested that the role of state government intervention in the health sector is greater in New South Wales than in Victoria, Queensland, South Australia, Western Australia, Tasmania, and Northern Territory. On the other hand, the coefficient of -[QLD+WA+TAS+(SA+NT)] less than 2 implies that the coefficient of (1+SD_8) is also less than 2. This suggests that the coefficient of SD_8 is less than 1. For example, let $a_7 G_s{*}(1+SD_8) = a_7 G_s{*}(SD_1+SD_8)$. Then $a_7<2$ implies that $(\partial H_1/\partial G_s{*}SD_1 + \partial H_1/\partial G_s{*}SD_8)<2$. Since $\partial H_1/\partial G_s{*}SD_1=\beta_1$, $\partial H_1/\partial G_s{*}SD_8$, the coefficient of $G_s{*}SD_8$, is less than 1. This implies that the role of the state government intervention in the health sector is greater in New South Wales than in Australia Capital Territory. In order to compare the seven different states and territories with New South Wales, we use the absolute values of the estimated coefficients. For example, suppose a_2 = 3.9523. Its absolute value is 3.9523 which is greater than 1, implying that ceteris paribus, the role of government intervention in the health sector is smaller in New South Wales than in Queensland.

The linear structural equations on the proportion of persons aged 15 years or over with post-school qualifications (E) and on nominal gross annual median income of persons aged 15 years or over (Y) are specified as:

$$
\begin{aligned}
E_i = {} & d_0 + \sum_{j=1}^{3} d_j H_{ji} + d_4 Y_i + \sum_{j=2}^{3} d_{3+j} R_{ji} + \sum_{j=1}^{2} d_{6+j} OCC_{ji} + d_9 C_i + d_{10} NGSTD_i \\
& + d_{11} UD_i + d_{12} AE_i + d_{13} GNG_{1i} + d_{14} GNG_{2i} + d_{15} GNG_{3i} + d_{16} GNG_{4i} \\
& + d_{17} GNG_{6i} + d_{18} GNG_{10i} + \text{Error} \qquad \textbf{(10)}
\end{aligned}
$$

and

$$
\begin{aligned}
Y_i = {} & e_0 + \sum_{j=1}^{3} e_j H_{ji} + e_4 E_i + \sum_{j=2}^{3} e_{3+j} R_{ji} + e_7 OCC_{1i} + e_8 C_i + e_9 METD_i + e_{10} K_s \\
& + e_{11} Ln_i + e_{12} I_i + e_{13} PFL_i + e_{14} GNG_{2i} + e_{15} GNG_{9i} + e_{16} GNG_{10i} \\
& + \text{Error} \qquad \textbf{(11)}
\end{aligned}
$$

where $d_0 > 0$ assumes that families live in regions or communities for several generations so that the post-school educational achievement is related to past family income.[8] d_6 and $e_6 > 0$ reject the hypothesis that differences in education, income and three health indicators are manifestations of moderate drinking.

In order to perform the major objective of this section, a variant of the Hausman's endogeneity tests as described by Dowrick (1993), it is necessary to obtain the predicted values of H_1, H_2, H_3, E, and Y from the estimates of the reduced-form equations for the linear structural equations (7), (8), (9), (10), and (11) specified as:

$$
\begin{aligned}
V_i = {} & \pi_{v0} + \pi_{v1} R_{2i} + \pi_{v2} R_{3i} + \pi_{v3} F_i + \pi_{v4} OCC_{1i} + \pi_{v5} OCC_{2i} + \pi_{v6} C_i + \pi_{v7} B_{1i} \\
& + \pi_{v8} B_i + \pi_{v9} G_s{*}VIC_i + \pi_{v10} G_s{*}QLD_i + \pi_{v11} G_s{*}WA_i + \pi_{v12} G_s{*}TAS_i, \\
& + \pi_{v13} G_s{*}(SA_i + NT_i) + \pi_{v14} METD_i + \pi_{v15} NGSTD_i + \pi_{v16} UD_i + \pi_{v17} AE_i \\
& + \pi_{v18} K_s + \pi_{v19} Ln_i + \pi_{v20} I_i + \pi_{v21} PFL_i + \pi_{v22} GNG1_i + \pi_{v23} GNG2_i \\
& + \pi_{v24} GNG3_i + \pi_{v25} GNG4_i + \pi_{v26} GNG6_i + \pi_{v27} GNG8_i + \pi_{v28} GNG9_i \\
& + \pi_{v29} GNG10_i + \text{Error}_v \qquad \textbf{(12)}
\end{aligned}
$$

8 We should like to thank Phil Lewis for providing this comment.

where $V_i = (H_{1i}, H_{2i}, H_{3i}, E_i, Y_i)'$ refers to the row vector of the dependent variables and the π's are the functions of the structural parameters indicated in the equations (7), (8), (9), (10), and (11).

In equations (7), (8), (9), (10), and (11), the regression coefficient reflect the direct effect. In equation (12), the reduced-form parameter, π, implies the total effect, which is the sum of the direct and indirect effects. For example, moderate drinking affects income directly and indirectly. If drinking entails out-of-pocket or direct expenses, it is negatively related to income. This suggests that the proportion of persons aged 18 years or over with moderate alcohol consumption of a community (R_3) has a negative direct effect on income (Y); $e_6 < 0$ in equation (11).

On the other hand, the indirect effect of moderate drinking on income can analytically be decomposed into the two ways. First, moderate drinking yields beneficial physical and psychological effects which have beneficial health effects. These beneficial health effects of moderate drinking carry over to the labour market, raising productivity and income [Berger and Leigh (1988), and Hamilton and Hamilton (1993)]. Second, since education and income are positively inter-related, the beneficial health effects of moderate drinking raise the level of education, which in turn raises income. In equation (12), $\pi_{Y2} > 0$ suggests that a positive indirect effect of moderate drinking on income offsets a negative direct effect.[9] In brief, the proportion of persons aged 18 years or over with moderate alcohol consumption in a community has a positive indirect effect on income, through an increase in the proportion of persons with each of no recent illness (H_1) and no chronic conditions (H_2) and the proportion of persons aged 18 years or over with self-assessed good health (H_3), and through an increase in the proportion of persons aged 15 years or

---

[9] That is, e_6 implies the direct effect of moderate drinking on income. π_{Y2} implies the total effect of moderate drinking on income. The total effect equals the sum of the direct and indirect effects. Therefore, π_{Y2}-e_6 implies the indirect effect. Given $e_6 < 0$, $\pi_{Y2} > 0$ implies that a positive indirect effect is greater in absolute value than a negative direct effect. For more details see Grossman and Benham (1974, p.227).

over with post-school qualifications (E) caused by an increase in the proportion of persons with each of three indicators of good health.

In equation (12), $\pi_{E1} + \pi_{E2})/2$ and $(\pi_{Y1} + \pi_{Y2})/2 < 0$ imply that the negative indirect effects of excessive drinking on education and income via three health indicators offset the positive direct effects of excessive drinking on education and income.

3.2 LIFESTYLE FACTORS - SMOKING AND DRINKING BEHAVIOURS

Different patterns of behaviour can help to explain differentials in health. While such lifestyle differences make an important contribution to socio-economic differences in health, there may be higher morbidity and mortality in lower socio-economic groups, even after controlling for lifestyle effects. For instance, moderate drinking yields beneficial physical and psychological effects which may have beneficial health effects [Hamilton and Hamilton (1993)]. On the other hand, levels of education and income affect moderate drinking directly and indirectly via good health - an important determinant of the drinking behaviour, implying that two causal relationships potentially exist; moderate drinking affecting health, and health affecting moderate drinking.

However, these beneficial health effects deteriorate as alcohol use increases, together with a negative effect of smoking on health [National Health Strategy of Australia, Research Paper No.1 (1992), and Conway, Pinyopusarerk, Carter, Penm and Stevenson (1993)].

In order to estimate the detrimental effect of excessive drinking on health, the linear structural equations on the proportion of persons with no recent illness and the proportion of persons aged 18 years or over with self-assessed good health in region i, (7) and (9), are modified and take the following forms:

$$H_{1i} = f_0 + f_1R_{3i} + f_2E_i + f_3Y_i + f_4R_{2i} + f_5F_i + f_6OCC_{1i} + f_7G_s*VIC_i + f_8G_s*QLD_i + f_9G_s*WA_i + f_{10}G_s*TAS_i + f_{11}G_s*(SA_i+NT_i) + f_{12}C_i + f_{13}B_{1i} + f_{14}B_i + f_{15}METD_i + f_{16}GNG_{8i} + f_{17}GNG_{9i} + f_{18}GNG_{10i} + \text{Error} \quad \textbf{(13)}$$

$$H_{3i} = g_0 + g_1R_{3i} + g_2E_i + g_3Y_i + g_4R_{4i} + g_5F_i + g_6OCC_{1i} + g_7G_s*VIC_i + g_8G_s*QLD_i + g_9G_s*WA_i + g_{10}G_s*TAS_i + g_{11}G_s*(SA_i+NT_i) + g_{12}C_i + g_{13}METD_i + g_{14}GNG_{8i} + g_{15}GNG_{9i} + g_{16}GNG_{10i} + \text{Error} \quad \textbf{(14)}$$

where $g_4 < 0$ suggests that a decrease in the proportion of persons aged 18 years or over with excessive alcohol consumption of a community increases the proportion of persons with self-assessed good or excellent health.

In equations (13) and (14), it is assumed that the proportion of persons with no chronic condition (H_2) and the proportion of persons aged 18 years or over with moderate alcohol consumption (R_3), display no causal relationships. If good health in terms of no chronic condition has a negative effect on moderate drinking, it can be expected that an increase in the proportion of persons with no chronic condition in a community increases the proportion of persons aged 18 years or over with excessive alcohol consumption (R_4).[10] For this reason, the proportion of persons with no chronic condition variable (H_2) is introduced into the linear structural equation of the proportion of moderate drinkers (R_3) specified as:[11]

---

10 In Appendix Table A-38, the tests of independence and contingency table indicate that no chronic condition and moderate drinking are negatively dependent on one another. Assume that no chronic condition has a negative effect on moderate drinking; $h_5 < 0$. If a higher proportion of persons aged 18 years or over with moderate alcohol consumption is associated with a lower proportion of persons aged 18 years or over with excessive alcohol consumption, one can suggest that no chronic condition has a positive effect on excessive drinking. Given the detrimental effect of excessive drinking on no chronic condition in the first specification of the model, this also suggests that no chronic condition and excessive drinking have no causal relationships.

11 The linear form of a model against the double logarithmic form is tested as reported in Table 5.13, indicating that the null hypothesis of linear model is accepted.

$$R_{3i} = h_0 + h_1H_{1i} + h_2H_{3i} + h_3E_i + h_4Y_i + h_5H_{2i} + h_6R_{1i} + \sum_{j=3}^{5} h_{4+j}AGE_{ji}$$

$$+ \sum_{j=3}^{5} h_{7+j}FAGE_{ji} + h_{13}G_s{*}VIC_i + h_{14}G_s{*}QLD_i + h_{15}G_s{*}WA_i$$

$$+ h_{16}G_s{*}TAS_i + h_{17}G_s{*}(SA_i{+}NT_i) + \text{Error} \qquad \textbf{(15)}$$

where $h_5 < 0$ suggests that the proportion of persons with no chronic condition and the proportion of persons aged 18 years or over with moderate alcohol consumption, display no causal relationships.

The equations (13), (14) and (15) examine the possibility that the positive correlations among education, income, and moderate drinking develop only after the hypothesis that moderate drinking yields beneficial physical and psychological effects which have beneficial health effects. This implies that the mechanism behind the education-income-moderate drinking relationships may also give rise to the education-income relationship.

The smoking variable (R_1) is introduced into the equation (15) in order to estimate the detrimental effect of smoking on health, together with the detrimental effect of excessive drinking on health, from the following reduced-form equations of the pairs of linear structural equations (13), (14), and (15):

$$V_i = \pi_{v0} + \pi_{v1}E_i + \pi_{v2}Y_i + \pi_{v3}H_{2i} + \pi_{v4}R_{1i} + \pi_{v5}R_{2i} + \pi_{v6}R_{4i} + \pi_{v7}AGE_{3i}$$

$$+ \pi_{v8}AGE_{4i} + \pi_{v9}AGE_{5i} + \pi_{v10}FAGE_{3i} + \pi_{v11}FAGE_{4i} + \pi_{v12}FAGE_{5i}$$

$$+ \pi_{v13}F_i + \pi_{v14}OCC_{1i} + \pi_{v15}G_s{*}VIC_i + \pi_{v16}G_s{*}QLD_i + \pi_{v17}G_s{*}WA_i$$

$$+ \pi_{v18}G_s{*}TAS_i + \pi_{v19}G_s{*}(SA_i{+}NT_i) + \pi_{v20}C_i + \pi_{v21}B_{1i} + \pi_{v22}B_i$$

$$+ \pi_{v23}METD_i + \pi_{v24}GNG_{8i} + \pi_{v25}GNG_{9i} + \pi_{v26}GNG_{10i} + \text{Error}_v \qquad \textbf{(16)}$$

where $V_i = (H_{1i}, H_{3i}, R_{3i})'$ refers to the row vector of the dependent variables and the π's are the functions of the structural parameters indicated in the equations (13), (14), and (15).

In equation (16), the coefficient of multiple determination (R^2), the ratio of explained variance to the total variance, for the reduced-form estimates on the proportion of persons aged 18 years or over with moderate alcohol consumption (R_3), will be equal to 1.000 since the sum of the proportion of persons aged 18 years or over without alcohol consumption (R_2), the proportion of persons aged 18 years or over with moderate alcohol consumption, and the proportion of persons aged 18 years or over with excessive alcohol consumption (R_4) is 100 per cent. Therefore, Hausman's form of the tests for endogeneity is not applied.

On the basis of this consideration, the Beggs (1988, p.96) tests for exogeneity are performed. For example, the test for exogeneity of the proportion of persons with no recent illness (H_1) and the proportion of persons with self-assessed good health (H_3) in the proportion of persons aged 18 years or over with moderate alcohol consumption equation (15) involves augmenting the ordinary least squares (OLS) regression for the proportion of persons aged 18 years or over with moderate alcohol consumption with the residual of each of the suspected regressors of endogeneity (H_1 and H_3) from each of the equations (13) and (14), and then testing the joint significance of the added variables in the original OLS regression. Therefore, two-stage least squares (TSLS) can be used for the proportion of persons aged 18 years or over with moderate alcohol consumption only, replacing the proportion of persons with no recent illness and the proportion of persons with self-assessed good health in equation (15) by their predicted values obtained from the reduced-form estimates for each of the proportion of persons with no recent illness and the proportion of persons aged 18 years or over with self-assessed good health in equation (16).[12]

For this reason, estimates are restricted to the reduced-form equations for the proportion of persons with no recent illness and the proportion of persons aged 18 years or over with self-assessed good health. In equation (16), π_{H14}, π_{H16}, π_{H34}, and $\pi_{H36} < 0$

---

[12] The coefficient of multiple determination (R^2) of the reduced-form estimate on moderate drinking is unity so that the estimated values of no recent illness and self-assessed good health are the same as the actual values. This implies that the TSLS estimates on two health indicators will be same as the OLS estimates. Note that the no recent illness equation (13) has been taken from the first specification of the model. Thus, there is no specific reason that the no drinking variable (R_2) is controlled for.

suggest that smoking and excessive drinking have total detrimental effects on health. π_{H14} and $\pi_{H16} < 0$ also suggest that on average, smoking and excessive drinking deteriorate health in terms of no recent illness for children as well as for adults.

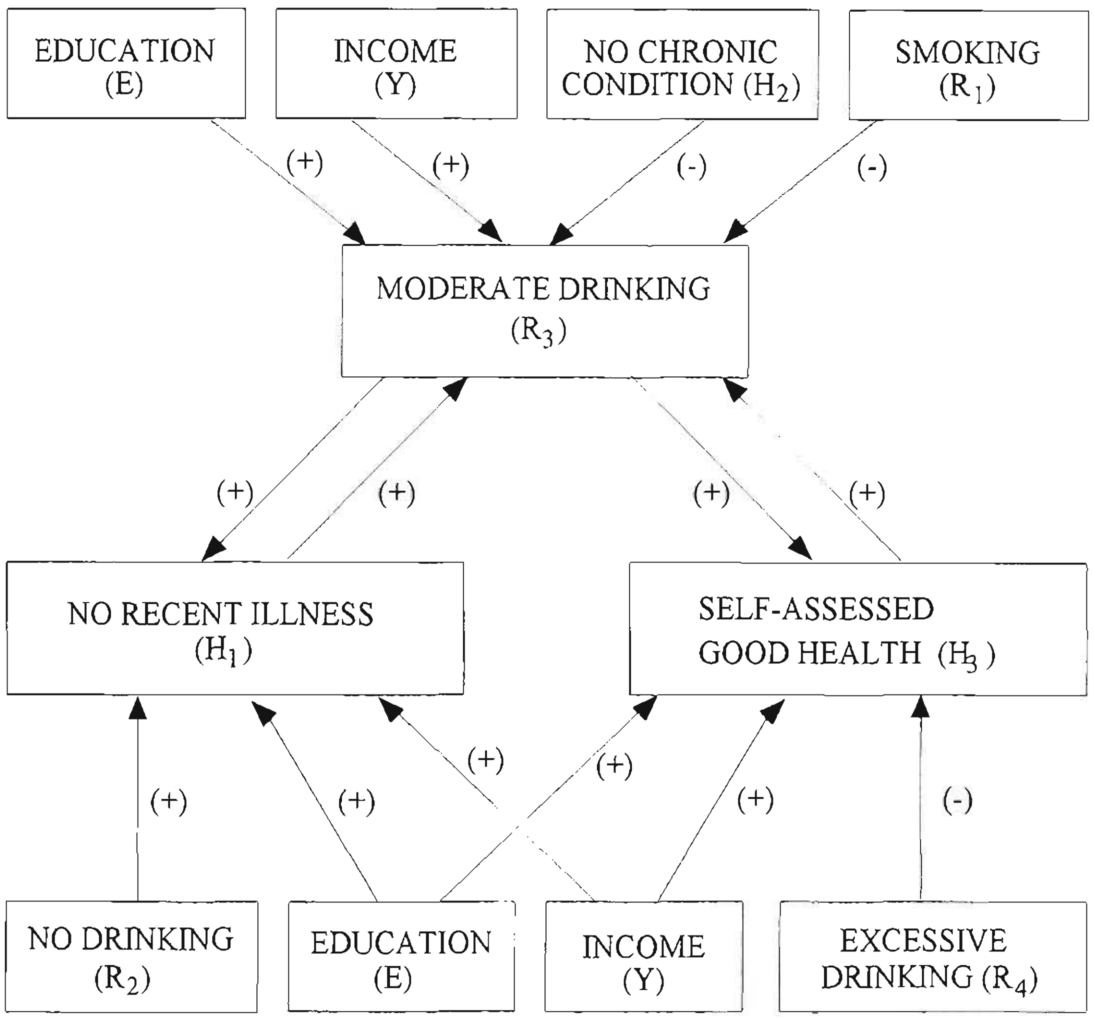
A schematic view (Figure 3.2) presents the logical linkages among the null hypotheses given in the equations (13), (14), (15), and (16). For example, an increase in either education or income affects moderate drinking directly and indirectly, through good health in terms of no recent illness and self-assessed good health.

So far it has been assumed that smoking and excessive drinking have detrimental effects on health. Then the analysis is extended to allow for changes in smoking and excessive drinking in response to changes in education and income. Addicts with lower rates of discount respond more to changes in harmful future consequences of addictive goods, such as a negative effect on health of smoking and excessive drinking [Becker, Grossman, and Murphy (1991)]. It follows that a higher expected cost (due perhaps to greater information about health hazards) is likely to reduce the demand for cigarettes and alcohol use, particularly among the rich. Since individuals with more education and higher income have lower rates of discount, communities with a higher proportion of persons with post-school qualifications and higher income have lower rates of discount.[13] This suggests that a reduction in the demand for cigarettes and alcohol use rather than their higher prices is likely to have a substantial positive effect on health. Thus, a community with lower rates of discount would be better off if it substantially eliminated tobacco smoking by reducing the proportion of the population who smoke on a regular basis and if it reduced the proportion of persons who are drinking regularly at hazardous or harmful levels [Conway, Pinyopusarerk, Carter, Penm, and Stevenson (1993)]. For example, the campaign against smoking and excessive drinking appears to be potentially beneficial to health in the community [Collins and Lapsley (1993)].

---

[13] The approach can be justified under the assumption that relationships among aggregates depend on characteristics and resources among individuals [Auster, Leveson, and Sarachek (1969, p.418)].

Figure 3.2. Inter-relationships Among No Recent Illness, Self-assessed Good Health, and Moderate Drinking: Effects of Education, Income, No Chronic Condition, Smoking, No Drinking, and Excessive Drinking



Notes: 1. Arrows relevant to causation direction from one to the other.
2. Predicted directions in parentheses.

In order to test the effects of education and income on smoking and excessive drinking, the linear structural equations on the proportion of persons aged 18 years or over with cigarettes consumption (R_1) and the proportion of excessive persons aged 18 years or over with excessive alcohol consumption (R_4) in statistical region i are defined, respectively, by:[14]

$$R_{1i} = k_0 + k_1R_{4i} + k_2E_i + k_3Y_i + k_4AGE_{8i} + k_5R_{3i} + k_6NES_i + k_7Ln_i + k_8OCC_{1i} + k_9GNG_{1i} + k_{10}GNG_{3i} + k_{11}GNG_{9i} + k_{12}GNG_{11i} + Error \tag{17}$$

$$R_{4i} = \ell_0 + \ell_1R_{1i} + \ell_2E_i + \ell_3Y_i + \ell_4AGE_{8i} + \ell_5R_{3i} + \ell_6NES_i + \ell_7Ln_i + \ell_8GNG_{1i} + \ell_9GNG_{3i} + \ell_{10}GNG_{4i} + \ell_{11}GNG_{5i} + \ell_{12}GNG_{6i} + \ell_{13}GNG_{7i} + \ell_{14}GNG_{8i} + Error \tag{18}$$

where each of k_2, k_3, ℓ_2, and ℓ_3 is hypothesised to be positive, and where k_5 and $\ell_5 < 0$ suggest that an increase in the proportion of persons aged 18 years or over with moderate alcohol consumption (R_3) of a community reduces the proportion of persons aged 18 years or over with cigarettes consumption and the proportion of persons aged 18 years or over with excessive alcohol consumption.

Given that moderate drinking and smoking are negatively associated with one another, the equations (17) and (18) suggest that the negative correlations between moderate drinking and excessive drinking develop only after the observation that moderate drinking is negatively associated with smoking.[15] This implies that the mechanism behind excessive drinking-moderate drinking relationships may also give rise to the excessive drinking-smoking relationship.

---

[14] The linear form of a model against the double logarithmic form is tested as reported in Table 5.19, indicating that the null hypothesis of linear model is accepted.

[15] In Appendix Table A-39, the tests of indpendence and contingency table indicate that moderate drinking and smoking are negatively dependent on one another.

Equations (17) and (18) reflect that smoking and excessive drinking complement each other; communities with relatively higher proportion of persons aged 18 years or over with cigarettes consumption have a higher proportion of persons aged 18 years or over with excessive alcohol consumption, and vice versa.[16] Given the importance of the population's health, this suggests that smoking and excessive drinking might be considered jointly in policy formulation. For this purpose, a variant of the Hausman's tests for exogeneity as described by Beggs (1988, p.96) is used. Therefore, it is necessary to obtain the residuals of R_1 and R_4 from the reduced-form equations for the pairs of the linear structural equations (17) and (18) specified as:

$$\begin{aligned} V_i = \; & \pi_{v0} + \pi_{v1}E_i + \pi_{v2}Y_i + \pi_{v3}AGE8_i + \pi_{v4}R3_i + \pi_{v5}NES_i + \pi_{v6}Ln_i \\ & + \pi_{v7}OCC1_i + \pi_{v8}GNG1_i + \pi_{v9}GNG3_i + \pi_{v10}GNG4_i + \pi_{v11}GNG5_i \\ & + \pi_{v12}GNG6_i + \pi_{v13}GNG7_i + \pi_{v14}GNG8_i + \pi_{v15}GNG9_i + \pi_{v16}GNG11_i \\ & + Error_v \qquad \textbf{(19)} \end{aligned}$$

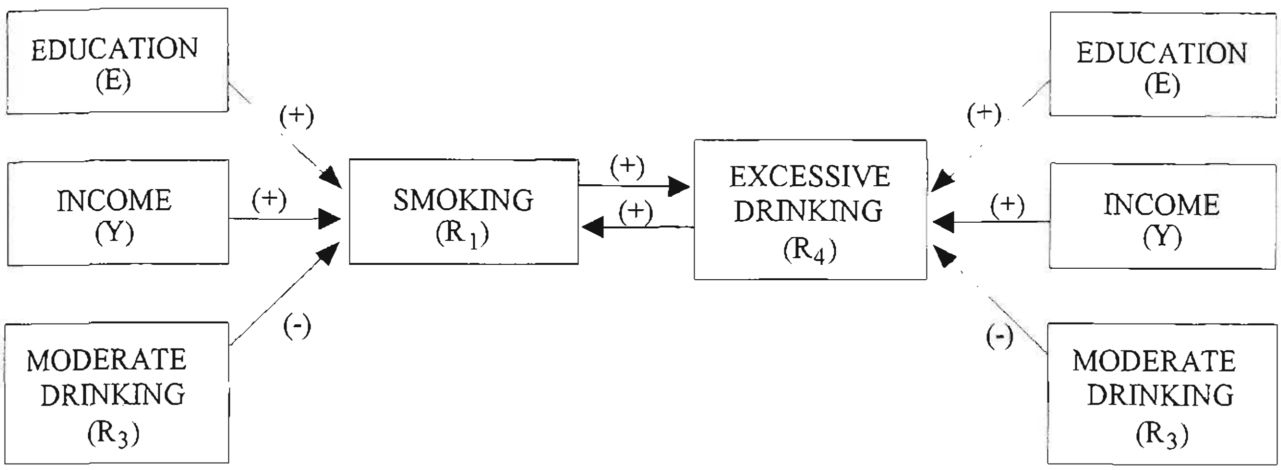
where $V = (R_1, R_4)'$ refers to the row vector of the dependent variables and the π's are the functions of the structural parameters indicated in the equations (17) and (18).

A schematic view (Figure 3.3) presents the logical linkages among the null hypotheses given in the equations (17) and (18).

---

16 In the first specification of the model, smoking is not controlled for. In Appendix Table A-41, A-42, and A-43, the tests of independence and contingency tables indicate that smoking and each of no recent illness and no chronic conditions are positively dependent on one another, whereas smoking and self-assessed good health are negatively dependent on one another. The former is not consistent with the spread of information that smoking is a serious health hazard [Fuchs (1986, p.245)]. Thus, the second specification of the model examines if smoking as well as excessive drinking has a detrimental effect on each of no recent illness and self-assessed good health. Furthermore, the positive inter-relationship between smoking and excessive drinking suggests that differences in smoking and excessive drinking are causal to differentials in three indicators of health.

Figure 3.3. Inter-relationships Between Smoking and Excessive Drinking: Effects of Education, Income, and Moderate Drinking



Notes: **1.** Arrows relevant to causation direction from one to the other.
2. Predicted directions in parentheses.

Assume that excessive drinking is a causal factor for more education and higher income in the equations (10) and (11). Assume further that the negative indirect effects of excessive drinking on education and income via three indicators of health offset the positive direct effects of excessive drinking on education and income in equation (12). Then, the inter-relationship between smoking and excessive drinking rejects the hypothesis that differences in both education and income are causal to excessive drinking differences in favour of the existence of smoking.

4.0 DATA

The data for this investigation comes from the 1989-90 National Health Survey (NHS) and the 1991 Commonwealth of Australia Census that have been conducted by the Australian Bureau of Statistics (ABS). Details of the data sources are given in Table 4.1.

Table 4.1. Source of Variables[1]

Variable	Source[2]
H, E, Y, R, F, OCC, C, AGE, NES, I, PFL, GNG	1989-90 National Health Survey, ABS (4363.0)
B	June 1989 State and Territory Health Authorities[3]
G, K	1990-91 Australian National Accounts, ABS (5221.0)
Ln	Trade Union Members Australia, ABS (6325.0)
NGSTD, AE	1991 Census of Population and Housing, ABS (2722.0)
UD	Australian Education Directory[4]. Universities in Australia[5]

Notes: **1.** ABS Catalogue Numbers in parentheses.
2. Provided by the ABS Information Consultancy Service for in-depth data investigations.
3. For details of data source see A Social Health Atlas of Australia, Vol.2, ABS Cat.No.4385.0, 1992, p.303.
4. Edited by J. Badger and P. Mathews, Australian Council for Educational Research, 1992.
5. Edited by M. Dwyer and S. Lewis, The Australian Financial Review Library, 1992.

The following persons living in Australia but not usually considered part of the Australia resident population were excluded from the scope of the NHS survey; diplomatic personnel of overseas governments and non-Australian members of their households, non-Australian service personnel stationed in Australia and their dependents, and overseas visitors whose usual place of residence is outside Australia. Other than these persons, non-Australians working in Australia, or in Australia as students or settlers, and their dependents were included in the survey.

Each person aged 18 years or over in the selected households was interviewed personally, with the exception of persons too old or sick. Persons aged 15 to 17 years were interviewed with the consent of a parent or guardian: otherwise, a parent or guardian was interviewed on their behalf. For persons aged less than 15 years, information was obtained from a person responsible for the child, usually the mother. Although the survey was conducted over a twelve-month period, October 1989 to September 1990, selected households with an equal chance of selection were interviewed only once in that period.[17] The data for the 1989-90 NHS are based on a representative sample of 22,200 households (or 57,000 persons) covering about one third of one percent of the Australian resident population selected at random.

Estimates obtained from the NHS were derived using a complex ratio estimation procedure which ensures that the survey estimates conform to an independently estimated distribution of the total population by age, sex and area. The age-sex-area population benchmarks were derived from an average of the four mid-quarterly population estimates for the period, adjusted for the scope of the study.

The Australian Bureau of Statistics makes data available from a number of its sample surveys for sixty-one Statistical Regions which are subregions below the state level, for all but Tasmania and the two territories, the Northern Territory and the Australian Capital Territory.

In general, the Statistical Region is large enough for the production of reliable estimates from sample surveys. The Statistical Regions are listed in Appendix Table A-1. Here the Statistical Regions are ordered as they appear in the data package of the 1989-90 NHS. Maps for Statistical Regions are shown in Glover and Woollacott (1992b, p.178).

---

[17] For a complete description of the sample, see the 1989-90 National Health Survey Users' Guide (1991, pp.1-11).

Using Statistical Regions, variations in reporting practices and in the accuracy of information, which may be serious at the individual level, tend to be averaged out. Another advantage of using regions as the unit of observation is that health is likely to vary less relative to variations in education and income across regions than across individuals. However, the approach could be justified by the assumption that the community's preferences for health, education and lifestyle factors versus their consumption goods represent an aggregation of individual preferences. For example, in examining the production model of health measured by mortality rates in logarithmic form across 51 states of the United States as the unit of observation in 1960 (sample size of white population in the labor force), Auster, Leveson and Sarachek (1969, p.418) note that relationships among aggregates may depend on characteristics and resources among individuals.

With few exceptions, the variables used in the model are expressed as proportions of an appropriate denominator. The respective measurements of three health status indicators referred to as H_1, H_2 and H_3 are as follows; first, H_1 is the proportion of persons without medical conditions such as illness, injury or disability experienced in the two weeks prior to interview; second, H_2 is the proportion of persons without medical conditions such as illness, injury or disability which have lasted or are expected to last for a period of six months or more including long-term and permanent impairment or disabilities; and third, H_3 is the proportion of persons aged 18 years and over with self-assessed good or excellent health status as opposed to poor or fair status at the survey interview. Note that each of three health status indicators implies a positive health effect, so that a higher proportion in each will be associated with relatively better health in a community.

The proportion of persons aged 15 years or over with post-school qualifications, E, is expressed as the proportion of those who obtained a qualification such as a bachelor degree or higher, trade/apprenticeship certificate/diploma since leaving school.

Nominal gross annual median income of persons aged 15 years or over, Y, is used as an alternative to the mean for a frequency distribution, with open-ended class interval (hence, $60,001 or more).[18] The measured median income is before taxes and government transfers.

The proportion of persons aged 18 years or over with cigarettes consumption, R_1, is measured as the proportion of persons aged 18 years or over with the average daily consumption of one or more cigarettes (or pipes or cigars) at the time of interview in 1989-90.

The proportion of persons aged 18 years or over with moderate alcohol consumption, R_3, is measured as the proportion of persons aged 18 years or over with the average daily consumption of alcohol greater than zero but less than 50 millilitres or 35 grams for males and 25 millilitres or 17.5 grams for females on a regular basis over the seven days prior to interview. This definition, referred to as a low risk alcohol intake, is based on recommendations of the National Health and Medical Research Council (NH & MRC) of Australia. Therefore, the proportion of persons aged 18 years or over with excessive alcohol consumption, R_4, is measured as the proportion of persons aged 18 years or over with the average daily consumption of alcohol greater than 50 millilitres or 35 grams for males and 25 millilitres or 17.5 grams for females [National Health Strategy of Australia, Research Paper No.1 (1992, p.127)].

In the proportion of skilled workers amongst all employees, OCC_1, skilled workers consist of Australian Bureau of Statistics occupation groups such as managers, administrators, and professionals. Unskilled workers are defined as plant

---

[18] See Mansfield (1983, p.37).
The median in grouped data may be derived as follows:
Median = $L + (f_s/2 - f_c)*W_m/f_m$
where L is the lower class boundary of the median, f_s is the sum of the frequencies, f_c is the cumulative frequency up to the median class, W_m is the width of the median class interval, and f_m is the frequency of the median class.

and machine operators and drivers, and labourers and related workers. This definition is based on the argument of Foster (South Australian Health Commission, 1991, p.32).

The total state and local government outlays on health per person aged 15 years or over, G, is measured as public health expenditures per person aged 15 years or over in current prices at June 30, 1991 by States and Territories. Such treatment allows us to examine the role of government intervention in the health sector and to separately compare Victoria, Queensland, South Australia, Western Australia, Tasmania, Northern Territory, Australian Capital Territory with New South Wales, treated as a base for comparisons in the analysis. Commonwealth funding data by statistical regions and states are not available in the ABS Information Consultancy Service for in-depth data investigation. Thus, Commonwealth outlays on health are excluded from the measurement. The total current outlays on health include hospital and other institutional services, hospital benefits, dental clinics and practitioners, clinics and other non-institutional services, medical benefits, public health, pharmaceuticals, and health research.

Data on nominal public health expenditure per person aged 15 years or over by States and Territories (G) are only available for the Census Years of 1990-91. However, under the assumption that a change of G is approximately proportional to a change in the age-adjusted population, the size and make-up of the populations of States and Territories will not have been such as to result in a shift in health status pattern of an area from low to high, or vice versa: that is, the relativities between States and Territories will not have changed substantially [Glover and Woollacott (1992b)].

The mix of part-time workers and full-time workers (PFL) is directly computed as persons who worked less than 35 hours per week divided by persons who worked 35 hours or more per week and measures the rate of part-time workers per person aged 15 years or over to full-time workers per person aged 15 years or over. It should be noted that this definition differs from that used in ABS Labour Force Surveys.[19]

Regions with a university campus containing both the chancellor and the admission offices are regarded as 1 in the university dummy variable, UD (see Appendix Table A-7). Generally, this assumption produces a useful measure for access to university courses since a variety of courses may be offered at this university campus. The 39 universities in Australia are located in 28 statistical regions. Note that the Australian Maritime College and the Australian Defence Force Academy are included in the measurement.

All the structural equations are estimated using both ordinary least squares (OLS) and two stage least squares (TSLS) methods, since all the structural equations are identified by the rank condition as presented in Appendix Tables A-2, A-3, and A-4. A list of data used in this study is available on request.

---

[19] See the 1989-90 National Health Survey Users' Guide (1991, p.45).

5.0 ESTIMATION RESULTS

5.1 THE INTER-RELATIONSHIPS AMONG HEALTH STATUS, EDUCATION AND INCOME

The major objective in this section is to test Lewis' propositions (1994) of the multiple inter-relationships among a set of three indicators of good health (no recent illness, no chronic condition, and self-assessed good health), education and income; ceteris paribus, an increase in each of these variables causes increases in the other two. For example, health status may be improved directly by increased expenditure on medical care or indirectly through increased expenditure on education, and through a more healthy environment, better diet and access to better medical care as a result of higher income. Education may be facilitated directly by increased current outlays on additional years of schooling or indirectly through increased expenditure on health and through higher income. Income may be raised directly by an increased level of skill or indirectly through higher productivities due to better health, and increases in skills and knowledge as a result of more education. An implication of this is that the cross effects among a set of three indicators of good health, education and income are positive. It turns out that they complement each other and therefore a set of three indicators of good health and education can be viewed as normal goods.[20]

5.1.1 Diagnostic Testing of the Hypothesis

Diagnostic testing is concerned with establishing whether an estimated model is an adequate description of an economic phenomenon. Beggs (1988, p.99) argues that diagnostic testing of econometric models is a positive activity which stimulates recourse to improved economic and statistical modelling in the sense that it can help create a clearer

---

[20] No recent illness (H_1) and no chronic condition (H_2) are measured as the proportions of persons with no recent illness and no chronic condition, respectively. Self-assessed good health (H_3) is measured as the proportion of persons aged 18 years or over with self-assessed good or excellent health to population in the same age scope. Education (E) is measured as the proportion of persons aged 15 years or over with post-school qualifications to population in the same age scope. Income (Y) is measured as nominal gross annual median income of persons aged 15 years or over.

picture of where the problems lie in existing models; the more tests that are carried out, the less the chance of accepting a poor model. In this subsection, diagnostic testing of the inter-relationships among a set of three indicators of good health, education, and income is conducted. In the first stage of testing, the linear and double natural logarithmic versions of the function are contrasted to each other using non-nested tests to determine which model is a better representation of behaviour. Then only the diagnostic tests of the linear functions which have survived stage one are reported in the following subsection.

Non-nested specific alternative models can be tested using variable addition methods. The MacKinnon-White-Davidson PE (1983, pp.54-56) test to compare the linear and double natural logarithmic models involves augmenting the linear model with the difference between the logarithm of the predictions of the linear model and the predictions of the natural logarithmic model from the OLS estimates, and augmenting the natural logarithmic equation with the difference between the predictions of the linear equation and the exponential of the prediction of the natural logarithmic equation from the OLS estimates.[21]

On the other hand, Beggs (1988, p.95) proposes testing natural logarithmic versus linear models by augmenting the linear equation with the exponential of the prediction of the

---

[21] The values of $\ln B_1$, $\ln GNG_1$, $\ln GNG_2$, $\ln GNG_4$, $\ln GNG_8$, and $\ln GNG_{10}$ are less than zero. Therefore, the prediction of the double natural logarithmic model for the non-nested tests is obtained by relying upon the first-order Taylor series approximation $\ln(1+x) \approx x$. For example, $\ln B_1 = \ln(1+B_1-1) \approx B_1-1$. Expressing the linear structural equations (7), (8), (9), (10), and (11) in natural logarithms results in the following functional forms:

$\ln H_1 \approx H(\ln E, \ln Y, \ln R_2, \ln R_3, \ln F, \ln OCC_1, \ln C, B_1\text{-}1, \ln B, \ln G_s{*}STATES_j, METD, GNG_8\text{-}1, \ln GNG_9, GNG_{10}\text{-}1)$

$\ln H_2 \approx H(\ln E, \ln Y, \ln R_2, \ln R_3, \ln F, \ln OCC_1, \ln C, B_1\text{-}1, \ln B, \ln G_s{*}STATES_j, METD, GNG_8\text{-}1, \ln GNG_9, GNG_{10}\text{-}1)$

$\ln H_3 \approx H(\ln E, \ln Y, \ln R_2, \ln R_3, \ln F, \ln OCC_1, \ln C, \ln G_s{*}STATES_j, METD, GNG_8\text{-}1, \ln GNG_9, GNG_{10}\text{-}1)$

$\ln E \approx E(\ln H_1, \ln H_2, \ln H_3, \ln Y, \ln R_2, \ln R_3, \ln OCC_1, \ln OCC_2, \ln C, \ln NGSTD, UD, \ln AE, GNG_1\text{-}1, GNG_2\text{-}1, \ln GNG_3, GNG_4\text{-}1, \ln GNG_6, GNG_{10}\text{-}1)$

$\ln Y \approx Y(\ln H_1, \ln H_2, \ln H_3, \ln E, \ln R_2, \ln R_3, \ln OCC_1, \ln C, METD, \ln K, \ln Ln, \ln I, \ln PFL, GNG_2\text{-}1, \ln GNG_9, GNG_{10}\text{-}1)$

where j = VIC, QLD, WA, TAS and (SA+NT), and where the subscript i denoting the individual statistical region is omitted from the functions for the sake of convenience. Table 3.1 in Chapter 3 provides a description of the variables as well as their means and standard deviations.

dependent variable from the natural logarithmic equation and the natural logarithmic model with the prediction of the dependent variable from the linear model run by OLS.

The null hypothesis (H_0) of the linear form of a model against the double natural logarithmic model is not rejected if the absolute value of the t-statistic on the added regressor of the linear model is less than the critical t-value, whereas the alternative hypothesis (H_1) of the double natural logarithmic model against the linear model is not rejected if the t-statistic on the added regressor of the double natural logarithmic model is less than the critical value. If the null hypothesis is not rejected, it is assumed that the linear model is preferred to the double natural logarithmic model. The no rejection of the null hypothesis suggests that the added regressor of the linear model is asymptotically uncorrelated with the disturbances of the model and the associated estimated regression coefficient is asymptotically zero. If the alternative hypothesis is not rejected, it is assumed that the double natural logarithmic model is preferred to the linear model. If both the null and alternative hypotheses are not rejected, the tests are said to be inconclusive.

The regression results presented in Table 5.1 indicate that the linear model is preferred to the double natural logarithmic model because the linear model t-statistic is less than the critical value. However for both the MacKinnon-White-Davidson PE and Beggs tests of the education function under the columns "Test I" and "Test II", respectively, the estimated t-statistics on the added regressor of each of the two models are less than the critical t-values of 1.683 with forty-one degrees of freedom at the 0.10 level of significance using a two-tailed test, implying that the tests are inconclusive. Therefore, the Box-Cox procedure as described by Maddala (1977, p.317) in Table 5.1 under the column "Box-Cox" is reported. For example, the Box-Cox procedure for the education function (E) involves dividing each E_i by the geometric mean of the E's; the exponential of the mean of the natural logarithm of E. Then we estimate the two equations and choose the one with the smaller residual sum of squares (RSS). In each of the two education equations, the double natural logarithmic model is chosen.

Table 5.1. Non-Nested Test of Double Logarithmic Versus Linear Models of Each of Three Indicators of Health, Education, and Income[1]

	Test I[2]		**Test II[3]**		**Box-Cox[4]**	
Equation	Logarithmic model (H_1)	Linear model (H_0)	Logarithmic model (H_1)	Linear model (H_0)	Logarithmic model (H_1)	Linear model (H_0)
No recent illness	t = 3.544***	t = 1.513 Accept H_0	t = 3.418***	t = 1.608 Accept H_0	RSS = 0.289	RSS = 0.245 Accept H_0
No chronic condition	t = 2.199**	t = 0.883 Accept H_0	t = 3.422***	t = 0.397 Accept H_0	RSS = 0.610	RSS = 0.597 Accept H_0
Self-assessed good health	t = 3.561***	t = 1.054 Accept H_0	t = 3.257***	t = 1.176 Accept H_0	RSS = 0.060	RSS = 0.053 Accept H_0
Education	t = 1.602 Reject H_0	t = 0.166 Accept H_0	t = 0.243 Reject H_0	t = 0.436 Accept H_0	RSS = 0.228 Reject H_0	RSS = 0.231
Income	t = 3.053***	t = 1.384 Accept H_0	t = 2.610**	t = 0.271 Accept H_0	RSS = 0.125	RSS = 0.116 Accept H_0

Notes: **1.** The t-statistic denotes the estimated absolute value. *** and ** indicate significance at the 1% and 5% levels on a two-tailed test, respectively. By "Accept H_0" we strictly mean "cannot reject H_0".

2. For the test procedure see MacKinnon, White, and Davidson (1983, pp.54-56).

3. For the test procedure see Beggs (1988, p.95).

4. The Box-Cox procedure as described by Maddala (1977, p.317). The use of the Box-Cox procedure has been suggested by Pat Wilson, thesis examiner.

However, the linear model is preferred to the double natural logarithmic model in the other four equations. Furthermore, Lewis' propositions of the multiple inter-relationships among three indicators of health, education, and income suggest a full simultaneous equations model. The Theil maximum adjusted multiple determination (Adj.R^2) criterion as described by Maddala (1992, p.497) indicates that the estimated value of Adj.R^2 for each of the two education functions is larger in the linear model (Adj.R^2 = 0.752) than in the double natural logarithmic model (Adj.R^2 = 0.744), suggesting that the linear model cannot be rejected. Thus, we choose the linear model in the education equation.

In order to test the null hypothesis of independence, each dependent variable is classified as the top third, the middle third and the bottom third of regions in terms of the sample proportions as introduced by Lewis, O'Brien and Thampapillai (1990, pp.197-201). The null hypothesis of independence between each pair of variables is rejected if the calculated value for Chi-Square (χ^2) is larger than the critical value. If the null hypothesis is rejected, there is strong evidence that one is dependent on one another. The 95 per cent critical values for $\chi^2(1)$ and $\chi^2(4)$ are 3.841 and 9.488, respectively.

The results reported in Table 5.2 reveal that the null hypothesis of independence between no recent illness and income, between no chronic conditions and education, and between no chronic conditions and income is not rejected.[22]

On the other hand, the null hypothesis of independence between no recent illness and education, between self-assessed good health and education, between self-assessed good health and income, and between education and income is rejected. For example, this implies that the proportion of statistical regions from each of three categories of self-

---

[22] The "acceptance" of the null hypothesis may simply reflect that sixty one statistical regions are not sufficient numbers of observations to allow a full evaluation for the test of independence between each pair of variables.

Table 5.2. Tests of Independence Among Three Indicators of Health, Education, and Income[1]

	Value of Chi-Square[2]			
Cross-classified Variables[3]	Pearson (4)	Likelihood ratio (4)	Mantel-Haenszel (1)	Decision
No recent illness-Education	5.603	5.853	5.436**	Reject H_0
No recent illness-Income	3.908	4.037	1.302	Accept H_0
No chronic condition-Education	4.052	3.955	2.342	Accept H_0
No chronic condition-Income	3.081	3.162	2.366	Accept H_0
Self-assessed good health-Education	11.664**	12.203**	9.157***	Reject H_0
Self-assessed good health-Income	13.025**	14.721***	11.693***	Reject H_0
Education-Income	37.292***	39.227***	23.476***	Reject H_0

Notes: **1.** For the test procedure see Lewis, O'Brien, and Thampapillai (1990, pp.197-201). By "accept H_0" we strictly mean "cannot reject H_0".

2. The number of degrees of freedom in parentheses. *** and ** indicate significance at the 1% and 5% levels, respectively.

3. The contingency tables between each pair of variables are given in Appendix Tables A-9, A-10, A-11, A-12, A-13, A-14, and A-15..

assessed good health is not the same for the proportion of statistical regions from each of three categories of income.[23]

We can suggest from the independence tests that the considerable associations between each pair of variables primarily result from a system of multiple causations. Therefore, this study attempts to carry out the tests for endogeneity, one of the major objectives of this study, by augmenting the OLS regression with the predicted values of the suspected regressors of endogeneity from the reduced-form estimations. Dowrick (1993, p.2) proposes that the null hypothesis of no endogeneity among the suspected regressors is rejected if the F-statistic for the joint significance of each equation is larger than the critical value. The 99 per cent critical values for F(2,38), F(2,40), F(4,34) and F(4,36) are 5.21, 5.18, 3.93 and 3.89, respectively. The regression results presented in Table 5.3 under the column "Endogeneity Tests" reveal that there is strong statistical evidence of endogeneity among different health indicators, education and income. For example, the endogeneity tests for education reject the null hypothesis that a set of three indicators of good health and income are not endogenous to the dependent variable, since the F-statistic for the joint significance is larger than the 99 per cent critical value for F(4,34) of 3.93. The numerator degrees of freedom are four suspected regressors. The denominator degrees of freedom subtract the intercept term, eighteen control variables, four suspected regressors and their predicted values for the endogeneity tests from the reduced-form estimations from the number of observations with sixty-one.

It can be concluded from the independence and endogeneity tests that ceteris paribus, a set of three indicators of good health, education and income are inter-related.

---

[23] The contingency table given in Appendix Table A-14 indicates that the percentages of the top third of regions in terms of self-assessed good health are 19.7 per cent and 1.6 per cent for the top third of regions and the bottom third of regions in terms of income. The table also implies that the percentages of the top third of regions in terms of income are 19.7 per cent and 6.6 per cent for the top third of regions and the bottom third of regions in terms of self-assessed good health, respectively. Therefore, self-assessed good health and income are positively dependent on one another.

The RESET2 test for functional form misspecification involves augmenting the regression with the square of the predicted value of the model and applying the t-test to the added coefficient. The null hypothesis of functional form misspecification is rejected if the absolute value of the t-statistic on the added regressor is less than the tabulated t-value. The regression results presented in Table 5.3 under the column "RESET2" suggest that the null hypothesis of functional form misspecification is rejected in all the estimated ordinary least squares (OLS) regressions with one exception. The estimated OLS regression for no chronic condition, indicating that the assumption of zero expected value of residuals is violated.

On the other hand, Lewis, O'Brien and Thampapillai (1990, p.296) argue that another indication of possible misspecification is low R^2 (the coefficient of multiple determination) and F-statistic; that is, the proportion of variability in the dependent variable that is explained by the specified model is less than that which would be explained by the correct model, implying that the R^2 and F-statistic are lower than those for the correctly specified model. In the estimated OLS regression for no chronic condition in Table 5.3, the value of R^2 is 0.776.

The observed value of R^2 is high; that is, 77.6 per cent of the total variation in the dependent variable is explained by variation in the full set of independent variables. The estimated F-statistic is 8.067, which is larger than the 99 per cent critical value for $F(18,42)$ of 2.40.

On the other hand, the Durbin-Watson (D.W.) statistic is a test for correlation of the residuals; this may reflect autoregressive disturbances or it may result from misspecifications [Gerdtham and Jönsson (1992)]. The null hypothesis of non-autoregressive residuals is rejected if the estimated D.W. value is less than the tabulated lower boundary value (d_L) or if it is larger than the value of $4\text{-}d_L$. The null hypothesis cannot be rejected if the estimated D.W. value is larger than the tabulated upper boundary value (d_U) or if it lies between d_U and the value of $4\text{-}d_U$. The tests are said to be inconclusive if the estimated D.W. value lies

Table 5.3. Diagnostic Evaluation of Each Equation for Three Indicators of Health, Education, and Income: The OLS Estimates

Equation	Summary and Diagnostic Statistics[1]							
	Endogeneity Tests[2]		Test of Functional Form Misspecification[3]					
	Adj.R^2	Joint Test	RESET2	R^2	Durbin-Watson Statistic	F-Stat	Hetero-skedasticity[4]	Homo-skedasticity[5]
No recent illness (H_1)	0.640	F(2,38) = 6.342***	t = 0.664	0.759	2.429	F(18,42) = 7.344***	F(22,39) = 0.936	χ^2 (18) = 12.708
No chronic condition (H_2)	0.720	F(2,38) = 8.733***	t = 2.895***	0.776	2.315	F(18,42) = 8.067***	F(22,39) = 1.297	χ^2 (18) = 23.192
Self-assessed good health (H_3)	0.567	F(2,40) = 5.357***	t = 0.330	0.630	1.796	F(16,44) = 4.674***	F(20,41) = 0.942	χ^2 (16) = 17.243
Education (E)	0.742	F(4,34) = 8.836***	t = 0.906	0.826	1.625	F(18,42) = 11.107***	F(24,37) = 0.852	χ^2 (18) = 19.330
Income (Y)	0.924	F(4,36) = 37.574***	t = 1.165	0.947	1.728	F(16,44) = 49.198***	F(22,39) = 1.355	χ^2 (16) = 19.461

Notes: **1.** *** indicates significance at the 1% level.
2,4 For the test procedure see Dowrick (1993, p.2).
3. For the test procedure see Dowrick (1993, p.2), and Lewis, O'Brien and Thampapillai (1990, p.296 and pp.302-307). The third power of the prediction of the dependent variable for RESET3 test has too low tolerance (1.00E-04 limits reached) and therefore the test is not performed.
5. For the test procedure see Breusch and Pagan (1979, p.1288).

between the tabulated lower and upper boundary values or between the values of 4-d_U and 4-d_L. If the null hypothesis is rejected, it is assumed that there is evidence of autocorrelated disturbances or misspecifications. If the null hypothesis is not rejected, it is assumed that the model is correctly specified. If the test is inconclusive, it is assumed that misspecification is not obvious.

The regression results presented in Table 5.3 reveal that the null hypothesis is not rejected at the 0.05 level of significance for all the regression equations estimated by OLS, suggesting that each of the five regression equations is correctly specified.[24] For example, the regression results presented in Table 5.3 indicate that the value of D.W. of the estimated OLS regression for no chronic condition is 2.315. The tabulated upper boundary value (d_U) and the value of 4-d_U with 18 explanatory variables at the 5 per cent level of significance are 2.382 and 1.618, respectively. The observed D.W. value lies between d_U and 4-d_U. Therefore, the null hypothesis is not rejected, suggesting that the model is correctly specified.

On the basis of RESET2 test, and the estimated values of R^2, F and D.W., it can be suggested that correct specifications are implied in all the five estimated OLS regressions, indicating that the assumption of zero expected values of residuals is not violated.

The tests for heteroscedasticity are executed by regressing the squared residual of the model from the reduced-form estimation on a set of explanatory variables. Dowrick (1993, p.2) proposes a test procedure where a set of explanatory variables consists of the control variables, the suspected regressors of endogeneity, their predicted values for the endogeneity tests and the added regressors for RESET2 tests. For example, the income equation has twenty-one explanatory variables; twelve control variables, four suspected

---

[24] With small cross-section data sets with the sixty one statistical regions and the large number of explanatory variables in the model the power of the test will depend on the way in which the observations are ordered. Here the sixty one statistical regions are ordered as they appear in the data package of the 1989-90 National Health Survey provided by the Australian Bureau of Statistics (ABS); the external data source organisation. For more detials see Gerdtham and Jönsson (1992, p.191), Maddala (1977, p.287), and Neter, Wasserman and Whitmore (1982, pp.4-5).

regressors of endogeneity, their predicted values for the endogeneity tests and the squared prediction of the model for RESET2. Including the intercept term, the numerator degrees of freedom are twenty-two. Since the number of observations is sixty-one, the denominator degrees of freedom are thirty-nine; that is, $F(22,39)$. The null hypothesis of no heteroscedasticity is rejected if the F-statistic is larger than the critical value. The 95 per cent critical value for $F(24,40)$ is 1.79. Consequently, it can be seen from the OLS regressions in Table 5.3 that the assumption that the variances of the disturbances are approximately constant for all of the sixty-one statistical regions cannot be rejected.

In order to reinforce this conclusion the Breusch and Pagan (1979, p.1288) tests for homoscedasticity are also adopted; under the null hypothesis of homoscedasticity, then the number of sample observations times the coefficient of multiple determination (R^2) from the secondary regression stated above has a Chi-Squared distribution with degrees of freedom equal to the number of non-constant explanatory variables in the secondary regression. The 90 per cent critical values for $\chi^2(18)$ and $\chi^2(16)$ are 25.989 and 23.542, respectively. From the reported regressions provided in Table 5.3, heteroskedasticity could not be detected.

In brief, diagnostic testing in this subsection reaches the following conclusions:

1. With one exception (the education equation) the non-nested tests of the linear form of a model versus the double natural logarithmic model indicate that the linear model is preferred to the double natural logarithmic model in all the equations for a set of three indicators of good health (viz. no recent illness, no chronic condition and self-assessed good health), and income. The Mackinnon-White-Davidson PE test (1983), the Beggs test (1988), and the Box-Cox procedure (Maddala, 1977) are conducted. The Box-Cox procedure indicates that the double natural logarithmic model is chosen in the equation for education. However, on the basis of Lewis' propositions of a full simultaneous equations model and the Theil maximum adjusted multiple determination criterion (Maddala, 1992) we choose the linear model. Therefore, only the linear models of the five dependent variables are used for the application of diagnostic tests.

2. Both the independence (Lewis, O'Brien and Thampapillai, 1990) and the endogeneity (Dowrick, 1993) tests demonstrate the multiple inter-relationships among a set of three indicators of good health, education and income as proposed by Lewis; education and income affecting a set of three indicators of good health, a set of three indicators of good health and income affecting education, and a set of three indicators of good health and education affecting income.

3. Results from the RESET2 test, and the estimated values of R^2 and F (Lewis, O'Brien and Thampapillai, 1990), and the observed value of D.W. (Gerdtham and Jönsson, 1992) indicate that correct specifications are implied in the five estimated OLS regressions, indicating that the assumptions of zero expected values of residuals are not violated.

4. Both the Breusch-Pagan (1979) and Dowrick (1993) tests indicate that heteroskedasticity could not be detected in any of the equations estimated by OLS. Therefore, the assumption that the variances of the disturbances are approximately constant for all of the sixty-one statistical regions cannot be rejected.

5.1.2 Estimates of the Inter-relationships Among Health Status, Education and Income

It has been suggested from the previous subsection that the estimated ordinary least squares (OLS) models may adequately describe the behavioural relationships. Moreover, each of the estimated regression equations run by OLS in Table 5.3, taken as a whole, does significantly explain the variation in the dependent variable, since the estimated F-statistics are greater than the 99 per cent critical values for $F(16,44)$ of 2.44 and $F(18,42)$ of 2.40. Therefore, it can be concluded that each of the five regression equations provides a statistically significant linear relationship between the dependent variable and the set of independent variables.

This subsection concentrates on an analysis of the estimated regression equations in Table 5.4 which suggest that a set of three indicators of good health, education, and income are positively inter-related.[25] It also reviews the tests of coefficient instability for both the estimated two-stage least squares (TSLS) regressions on education and income possibly caused by the university location dummy on each of three indicators of health. The test results are summarised in Tables 5.5 and 5.6.

An analysis of the estimated linear regression equations in Table 5.4 allows us to make the following comments. The proportion of persons aged 15 years or over with post-school qualifications and nominal gross annual median income of persons aged 15 years or over are positively related to the proportions of persons with a set of three indicators of good health, ceteris paribus. It can also be observed from this evidence that communities with relatively higher levels of education and income produce, on average, better health in terms of no recent illness and no chronic condition for children as well as for adults.[26]

---

[25] The control variables of the OLS and two-stage least squares (TSLS) regressions for H_1 and H_2 are R_2, R_3, F, OCC_1, C, B_1, B, G_s*$STATES_j$, METD, GNG_8, GNG_9, GNG_{10} where j = VIC, QLD, WA, TAS and (SA+NT).
The control variables of the OLS and TSLS regressions for H_3 are R_2, R_3, F, OCC_1, C, G_s*$STATES_j$, METD, GNG_8, GNG_9, GNG_{10}.
The control variables of the OLS and TSLS regressions for E are R_2, R_3, OCC_1, OCC_2, C, NGSTD, UD, AE, GNG_1, GNG_2, GNG_3, GNG_4, GNG_6, GNG_{10}.
The control variables of the OLS and TSLS regressions for Y are R_2, R_3, OCC_1, C, METD, K, Ln, I, PFL, GNG_2, GNG_9, GNG_{10}.
Their estimated results are reported in Table 5.11 and Table 5.12.

[26] In his comments, Phil Lewis questions the appropriateness of the estimation results by using the 10% level of significance. The problem with a preassigned significance level is that if the sample size is large enough, every null hypothesis can be rejected. Therefore, the significance level should depend on the sample size. For example, Leamer (1978, p.106) argues that the significance level must be made a decreasing function of sample size. Maddala (1992, p.32) also argues that the significance levels to be used should be much higher for small sample sizes (sometimes 25 to 50% and even 99% levels). Lewis, O'Brien and Thampapillai (1990, p.145) note that as the sample size gets larger, on average, the mean square error gets smaller. In this study, small cross-sectional data sets with sixty one statistical regions have been used. Furthermore, they note that the consequence of imperfect multicollinearity is that the standard errors of the regression coefficients are large (p.309). For example, in Appendix Table A-8, the simple bivariate correlation coefficients between education and self-assessed good health, and between income and self-assessed good health are 0.469 and 0.469, respectively. The estimated t-statistic for each regression coefficient is therefore not necessarily a good indication of the role that the independent variable (H_3) plays in explaining the variations in the dependent variables (E and Y).

In an attempt to further examine the important issues of the multiple inter-relationships among health status, education and income, we review the independence tests as reported in Table 5.2. Given small cross-sectional data sets with sixty one statistical regions, the null hypothesis of independence between each pair of variables is rejected, since the calculated value of Chi-Square of each pair is larger than the 50% critical values with 1 degree of freedom and 4 degrees of freedom of 0.46 and 3.36,

Table 5.4 presents the standard errors of the OLS and TSLS estimates (S.E. of estimates) for a set of three indicators of good health, education, and income. The standard errors of the estimates indicate that the smaller the variance of the sampling distribution, the greater is the precision of the estimator, that is, the greater is the chance of a sample estimate lying within some specified interval about the true value. The standard errors of the OLS and TSLS estimates of the five equations are small.

In a comparison of the OLS and TSLS estimates, the standard errors in the former are smaller than in the latter for no recent illness, education, and income. Therefore, we concentrate on an analysis of the econometric results from the OLS estimates for these three regressions.[27]

On the other hand, the standard errors of the TSLS estimates for either no chronic conditions or self-assessed good health are smaller than those of the OLS estimates. Furthermore, the D.W. for these two regression equations estimated by TSLS are 2.255 and 1.963, respectively, suggesting that the null hypothesis of no misspecification is not rejected at the 0.05 level of significance [Gerdtham and Jönsson (1992)]. Therefore, the TSLS regressions are correctly specified. On the basis of these considerations, we concentrate on an analysis of the econometric results from the TSLS estimates for either no chronic conditions or self-assessed good health.

---

respectively. Therefore, we conclude that health status (H_1, H_2 and H_3), education (E) and income (Y) are significant in explaining one another. We should like to thank Phil Lewis for pointing this out to us.

27 Sixty one statistical regions do not generate a sufficient number of observations given that the model includes a large number of explanatory variables. Therefore, we use the standard errors of estimates (SEE) in choosing between models. For example, see Maddala (1977, p.287), and Neter, Wasserman and Whitmore (1982, pp.208-210). The use of SEE is also based upon overall model performance, as suggested by Associate Professor John Mangan of Lancaster University.

Table 5.4. Estimates of the Inter-relationships Among Three Indicators of Health, Education, and Income

Suspected Regressors of Endogeneity[2]	Dependent Variables[1]: No recent illness (H_1)		No chronic condition (H_2)		Self-assessed good health (H_3)		Education (E)		Income (Y)	
	OLS	TSLS	OLS	TSLS	OLS	TSLS	OLS	TSLS[3]	OLS	TSLS[4]
H_1							0.0190	-0.0787	0.0922	0.0446
							(0.079)	(0.201)	(1.746)*	(0.533)
H_2							0.0979	0.2212	0.0100	-0.0016
							(0.783)	(1.167)	(0.338)	(0.033)
H_3							0.0444	0.3620	0.0001	-0.0588
							(0.195)	(0.981)	(0.016)	(0.789)
E	0.1065	0.1300	0.5711	1.0078	0.2618	0.5382			0.0919	0.1049
	(1.048)	(0.892)	(3.097)***	(4.334)***	(2.116)**	(3.633)***			(2.832)***	(2.085)**
Y	0.7771	0.7247	0.3790	0.0590	0.6523	0.5928	1.4550	1.3467		
	(3.705)***	(2.886)***	(0.997)	(0.147)	(2.490)**	(2.207)**	(4.008)***	(2.894)***		
R^2	0.759	0.726	0.776	0.803	0.630	0.692	0.826	0.822	0.947	0.937
S.E. of estimates	2.200	2.344	3.989	3.740	2.753	2.512	3.450	3.489	0.818	0.894
D.W.	2.429	2.366	2.315	2.255	1.796	1.963	1.625	1.752	1.728	1.698
ξ^5	0.431	0.402	0.180	0.028	0.134	0.121	0.501	0.463		

Notes:
1. Values in parentheses are the estimated absolute t-values. ***, **, and * indicate significance at the 1%, 5%, and 10% levels on a two-tailed test, respectively.
2. The estimated results of control variables are given in Tables 5.11 and 5.12.

3, 4. The coefficient stability tests and analyses are performed in Table 5.6.

5. Is the income elasticity at mean. Mean values of the predicted values of H_1, H_2, H_3, E, and Y are 29.16, 34.07, 79.00, 47.06, and 16.19, respectively.

The OLS estimates on no recent illness suggest that a 10 percentage point and $1,000 increase in the proportion of persons aged 15 years or over with post-school qualifications and nominal gross annual median income of persons aged 15 years or over in a community increases the proportion of persons with no recent illness by 1.065 percentage points and 0.777 percentage points, respectively.[28]

The TSLS estimates on no chronic condition suggest that a 10 percentage point increase in the proportion of persons aged 15 years or over with post-school qualifications and a $1,000 increase in the nominal gross annual median income of persons aged 15 years or over in a community increases the proportion of persons with no chronic condition by 10.078 percentage points and 0.059 percentage points, respectively.[29]

The TSLS estimates on self-assessed good health suggest that a 10 percentage point increase in the proportion of persons aged 15 years or over with post-school qualifications and a $1,000 increase in the nominal gross annual median income of persons aged 15 years or over in a community increases the proportion of persons aged 18 years or over with self-assessed good health by 5.382 percentage points and 0.593 percentage points, respectively.

The estimations of the linear form of the education regression imply that the level of education tends to be higher in communities with higher levels of any of the three indicators of good health and income. The OLS estimates suggest that a 10 percentage point increase in the proportions of persons with no recent illness, no chronic condition, and self-assessed

---

[28] Note that education is not significantly related to no recent illness when using the individual t-test. This neutrality effect of education on no recent illness may be attributed to a high correlation between education and income; $\gamma = 0.774$ as shown in Appendix Table A-8. Therefore, the income variable has been excluded from the equation. Then, in specification I in Appendix Table A-17, the education variable was found to be statistically significant at the 5% level (a t-statistic of 2.334 with the coefficient of 0.2499 in the OLS estimates). However, the standard error of the OLS estimates increased from 2.200 to 2.533.

[29] Note that income is not significantly related to no chronic conditions when using the individual t-test. For the same reason of the above footnote, the education variable has been excluded from the equation. Then, in specification II in Appendix Table A-18, the income variable was found to be statistically significant at the 10% level (a t-statistic of 1.803 with the coefficient of 0.8190 and the income elasticity at mean of 0.389 in the TSLS estimates). However, the standard error of the TSLS estimates was increased from 3.740 to 4.433.

good health (age 18 or over) in a community increases the proportion of persons aged 15 years or over with post-school qualifications by 0.190 percentage points, 0.979 percentage points, and 0.444 percentage points, respectively. The OLS estimates also suggest that a $1,000 increase in nominal gross annual median income of persons aged 15 years or over increases the proportion of persons aged 15 years or over with post-school qualifications by 1.455 percentage point.[30] A positive coefficient of the constant term suggests that past family income also has a positive effect on education. The implication here is that families have lived in communities for several generations.[31]

The estimations of the linear form of the income regression reveal that ceteris paribus, the higher levels of any of the three indicators of good health and education, the higher the income. The OLS estimates suggest that a 10 percentage point increase in the proportion of persons with each of no recent illness, no chronic condition, self-assessed good health (age 18 or over), and post-school qualifications (age 15 or over) in a community increases the nominal gross annual median income of persons aged 15 years or over by $922, $100, $1, and $919 respectively.[32]

---

[30] Note that a set of three indicators of good health is not significantly related to education when using the individual t-test. These neutrality effects of a set of three indicators of health on education may be attributed to a high correlation between no recent illness and no chronic condition; $\gamma = 0.721$ as shown in Appendix Table A-8. Therefore, the no recent illness and income variables have been excluded from the equation. Then, in specification III in Appendix Table A-20, each of no chronic condition and self-assessed good health is significantly related to education at the 5% and 1% levels, respectively (t-statistics of 2.082 and 2.903 with the coefficients of 0.2537 and 0.9612 in the TSLS estimates, respectively). Note that the standard error (S.E.) of the estimates is increased from 3.450 in the OLS estimates to 3.747 in the TSLS estimates. No recent illness is always insignificantly related to education as shown in Appendix Table A-20. However, the results reported in Table 5.2 reveal that no recent illness and education are dependent on one another.

[31] The estimated coefficient of the constant term is 27.4308 with a t-statistic of 1.165 in the OLS regression on education in Table 5.12. We should like to thank Phil Lewis for providing this comment.

[32] Note that two indicators of health status (H_2 and H_3) are not significantly related to income by using the individual t-test. For the same reason of the footnote 26, in specification I in Appendix Table A-21, the no recent illness, self-assessed good health, and education variables have been excluded from the equation. Then no chronic condition is significantly related to income in both the OLS and TSLS estimates (a t-statistic of 2.340 with the coefficient of 0.0592 in the OLS estimates and a t-statistic of 1.745 with the coefficient of 0.0567 in the TSLS estimates). Note that the standard errors (S.E.) of the estimates are increased from 0.818 to 0.886 in the OLS estimates and from 0.894 to 0.908 in the TSLS estimates. Self-assessed good health is always insignificantly related to income as shown in Appendix Table A-21. However, the results reported in Table 5.2 and Appendix Table A-14 reveal that self-assessed good health and income are positively dependent on one another.

The above evidence along with the independence tests in Table 5.2 and the endogeneity tests in Table 5.3 indicate that there are the multiple positive inter-relationships among a set of three health indicators, education and income in accordance with Lewis' propositions, suggesting a full simultaneous equations model. In short, an increase in each of these variables causes increases in the other two, ceteris paribus. Therefore, the instrument choice should not be based upon stability analysis of a single final target variable. For example, health status can be improved directly by increased expenditure on medical care or indirectly through increased expenditure on education and through healthy environment, better diet and access to the better medical care as a result of higher income. Education levels can be enhanced directly by increased outlays for schooling or indirectly through increased expenditure on health and through higher income. Income can be increased directly by increased levels of skill or indirectly through higher productivities due to better health, and increases in the skills and knowledge as a result of more education. It can also be suggested that the cross effects among a set of three health indicators, education and income are positive, implying that they complement each other and that health and education are normal goods.

Under the assumption that community's preferences for health and education versus consumption goods represent an aggregation of individual preferences,[33] Table 5.4 presents the ceteris paribus mean income elasticities (denoted ξ). As the magnitude of their observed income elasticities are less than one, both health and education are normal necessities [Layard and Walters (1978), and Olsen (1993)]. The OLS estimates suggest that a 10 per cent increase in nominal gross annual median income of persons aged 15 years or over increases the proportions of persons with no recent illness and post-school qualifications (age 15 or over) by 4.31 per cent and 5.01 per cent, respectively.[34] The TSLS estimates

33 In examining the production model of health measured by mortality rates in logarithmic form across 51 states of the United States as the unit of observation in 1960 (sample size of white population in the labour force), Auster, Leveson, and Sarachek (1969, p.418) note that relationships among aggregates may depend on characteristics and resources among individuals.

34 Utilising data from the National Longitudinal Survey of Youth in 1979 managed by the Human Resource Research Centre at Ohio State University with sponsorship from the United States Department of Labor and other federal agencies with the high school subsample of 753 youths who were enrolled in twelfth grade at the time of the 1982 interview, Cook and Moore (1993, p.423) provide

suggest that a 10 per cent increase in nominal gross annual median income of persons aged 15 years or over increases the proportions of persons with no chronic condition and self-assessed good health (age 18 or over) by 0.28 per cent and 1.21 per cent, respectively. Given these observations, government expenditures on three indicators of health and education, financed by a progressive income tax, are progressive and equalising.[35]

The observed R^2 values in the second stage of TSLS estimates presented in Table 5.4 are 0.726 for the proportion of persons with no recent illness 0.803 for the proportion of persons with no chronic conditions, 0.692 for the proportion of persons aged 18 years or over with self-assessed good health, 0.822 for the proportion of persons aged 15 years or over with post-school qualifications, and 0.937 for the nominal gross annual median income of persons aged 15 years or over. An implication of this is that the two-stage least squares (TSLS) method works well. Table 5.4 shows that the econometric results from the TSLS estimates are consistent with those from the OLS estimates with two exceptions; the estimated education and income regression equations. These differences may be due to the effect of university education on the three indicators of health. That is, it is implicitly suggested that universities make substantial contributions to education and thereby health in communities in which they are located. In Table 5.5, estimates of the indirect effect of university on three indicators of health suggest that universities yield beneficial health

---

the estimated two-stage least squares regressions for a long form; other things being equal, a coefficient of the log of income on the highest year completed is 0.042 with a standard error of 0.084 and the mean value of the high school completed variable of 12.619. From his estimations, it is observed that the income elasticity of schooling is 0.0033, suggesting that schooling is a normal necessity.

35 For example, consider a government expenditure on health financed by a progressive income tax:

$B = aY^{\beta}$ (due to a government expenditure on health) **(1)**

$C = bY^{\gamma}$ (due to a progressive income tax) **(2)**

where B stands for benefits by a government expenditure on health, Y for income, C for costs such as tax, β for the benefit elasticity, and γ for the cost elasticity. The empirical observation that health is a normal necessity implies that $0 < \beta < 1$ and a progressive income tax that $\gamma > 1$. Thus, the policy is said to be progressive and equalising.

From the equations (1) and (2),

$(B-C)/Y = aY^{\beta-1} - bY^{\gamma-1}$ **(3)**

A sufficient condition for progressivity is satisfied for all values of Y:

$dB/dY - dC/dY = a\beta Y^{\beta-1} - b\gamma Y^{\gamma-1} < 0$ **(4)**

Since $0 < \beta < 1$ and $\gamma > 1$, benefits and costs are both progressive, so is the package. As a consequence, the policy is equalising. For more details see Layard and Walters (1978, pp.100-101, p.138, and p.443).

effects indirectly, through beneficial education and income effects to communities in which they are located. For example, a coefficient of 2.4478 implies that communities with universities, other things being equal, have higher a proportion of persons with no chronic conditions by 2.4478 percentage points than communities without universities, due to a higher proportion of persons aged 15 years or over with post-school qualifications and nominal gross annual median income of persons aged 15 years or over.[36]

Table 5.5. Estimates of Total (Indirect) Effect of University on Three Indicators of Health through Education and Income[1]

Dependent Variables		
No Recent Illness (H_1)	No Chronic Condition (H_2)	Self-assessed Good Health (H_3)
0.7594	2.4478	0.9074

Notes: 1. No direct effect. Therefore, the indirect effect equals the total effect; an estimate of the reduced-form parameter of the university location dummy on three indicators of health given in Appendix Table A-16.

Given these observations, the university location dummies (UD) may be regarded as influential observations. In general, these observations belong to positive outliers for the three indicators of good health from the TSLS estimates. An outlier is an observation that is far removed from the rest of the observations. Positive outliers are extremely healthy communities, i.e., those communities with extremely high proportions of persons with three indicators of good health. This topic is important in regional comparisons based on cross sectionals of only 61 observations. Using the least squares method this outlying observation can produce substantial changes in the estimated regression equation [Beggs (1988) and

[36] Table 5.12 shows the direct effect of university on education with a coefficient of 1.0205, suggesting that communities with universities (UD) have higher proportion of persons aged 15 years or over with post-school qualifications by 1.0205 percentage points than communities without universities. The direct effect of university on education is greater than the total indirect effect of university on no recent illness. This is due to an estimate of the reduced-form income parameter of the university location dummy (hence, the indirect effect only) with a negative coefficient of -0.7204 in Appendix Table A-16. However, given the greater proportion of the population who are students in communities with university than in those without it, the university may have had a positive effect on income per permanent residents of a community. Permanent residents are defined as persons whose current usual residence is the same for many years except for those whose current usual residence is only a recent address and Australian students as well as non-Australians in Australia as students and their dependents who were included in the 1989-90 National Health Survey [The 1989-90 NHS Users' Guide (1991, p.3)].

Maddala (1992)]. For example, in their regression of per capita health care expenditure on per capita GDP based on cross-sectionals of 22 OECD countries in 1985, Gerdtham and Jönsson (1992) introduce a dummy variable that has a value of one for the outlying countries and zero otherwise, assuming that outliers are caused by unusual events. They find that outliers play some role in the regression.

On the other hand, the specification II in Appendix Table A-27 presents the results of both the OLS and TSLS regressions when no recent illness is omitted from the regression model on education and when no chronic condition and self-assessed good health are omitted from the regression model on income, since their interacted variables of the university location dummies make substantial contributions to education and income.

Following this approach, it is crucial that influential observations (i.e., the university location dummies) should be reliable and comparable in different communities. In comparisons of communities, difficulties arise when comparing the levels of education and income in different communities. Moreover, as shown in Table 5.12 and in specification II in Appendix Table A-27, the standard errors (S.E.) of the estimates are increased from 3.448 to 3.450 in the regression for education and from 0.872 to 0.881 in the income regression. Further, our interest lies with the notions that a set of three indicators of good health is an important determinant of the values of education and income [Lewis, O'Brien and Thampapillai (1990)] and that outlying observations can contribute a lot of information about the coefficient estimates [Gerdtham and Jönsson (1992)].

On the basis of these arguments, we test the null hypothesis of no coefficient instability by the university location dummies in the TSLS estimates. Both the Breusch-Pagan (1979) and Dowrick (1993) tests indicate that heteroskedasticity could not be detected in each of the estimated OLS regressions on education and income, suggesting that there is no noticeable instability. The null hypothesis of no coefficient instability is rejected if the F-statistic is larger than the critical F. If the null hypothesis is rejected, it is assumed that there is strong evidence of coefficient instability in both education and income, possibly caused by the university location dummies on three indicators of health. Therefore, in order to test for stability, this study introduces an interaction variable and a dummy distinguishing the two groups; the university-and no university-locations of 61 statistical regions [Giannaros (1985)].

Both the Beggs (1988) and Dowrick (1993) tests for stability are employed for the same purpose. Hence, Beggs uses the variable addition methods, whereas the test suggested by Dowrick augments the regression with the added regressors from the previous tests, the interaction variable and the dummies. The introduction of the interaction variable and the dummy allows us to estimate the differences in the coefficients of the slope and intercept between the two groups.[37] On an a priori basis, the coefficient of the interaction variable is positive depending on the directions of the slopes of both the education and income equations due to the university location dummies. Because the dummy has units for communities with universities, a positive sign of the coefficient of the interaction variable would mean that the slope of both education and income functions increased due to universities. In the two-stage least squares (TSLS) regression for education, the university location dummy variable (UD) interacts with no recent illness.

The Beggs test involves augmenting the regression with the new interaction variable, and then testing the joint significance of the added coefficient by the F-test. The

---

[37] Both the Beggs and Dowrick approaches also allow us to suggest the source of the coefficient instability. From the Beggs approach, the two-stage least squares (TSLS) regression for each of education and income is modified as follows; in a functional form,

$E = E(H_1, H_2, H_3, Y, H_1{*}UD, UD, R_2, R_3, OCC_1, OCC_2, C, NGSTD, AE, GNG_1, GNG_2, GNG_3, GNG_4, GNG_6, GNG_{10})$

$Y = Y(H_1, H_2, H_3, E, H_2{*}UD, H_3{*}UD, UD, R_2, R_3, OCC_1, C, METD, K, Ln, I, PFL, GNG_2, GNG_9, GNG_{10})$

where the subscript i denoting the individual statistical region is omitted from the functions for the sake of convenience.

numerator degree of freedom is one. The denominator degrees of freedom are obtained by subtracting the eighteen original explanatory variables, the new interaction variable, and the constant term from the number of observations with the sixty-one statistical regions. The observed F-statistic presented in Table 5.6 under the column "Test I" is greater than the 99 per cent critical value for F(1,41) of 7.36. Therefore, the null hypothesis of no coefficient instability is rejected. The D.W. statistic suggests that the model is correctly specified.

In the TSLS regression for income, the university location dummy variable interacts with no chronic condition and self-assessed good health. The Beggs test involves augmenting the regression with the two new interaction variables and the university location dummy variable, and then testing the joint significance of the added variables by the F-test. The numerator degrees of freedom are three. The number of original explanatory variables is sixteen. Including the constant term the denominator degrees of freedom are forty-one. In Table 5.6 under the column "Test I", the observed F-statistic is greater than the 99 per cent critical value for F(3,41) of 4.36. Therefore, the null hypothesis of no coefficient instability is rejected. The D.W. statistic suggests that the model is correctly specified.

In order to reinforce the above conclusions, the Dowrick test is also conducted. The test involves augmenting the TSLS regression with the new interaction variable(s), the predictions of the suspected regressors of endogeneity for the endogeneity test, the squared prediction of the model for the RESET2 test, and the constant term. For example, in the TSLS regression for education, the numerator degrees of freedom consist of the one interaction variable, the eighteen original explanatory variables, the predictions of the four suspected regressors of endogeneity, the squared prediction of the model, and the constant term. The suspected regressors of endogeneity are no recent illness, no chronic condition, self-assessed good health, and income. Therefore, the numerator degrees of freedom are twenty-five. The denominator degrees of freedom are obtained by subtracting the numerator degrees of freedom, the eighteen original explanatory variables, and the constant term from the number of observations with the sixty-one statistical regions. The observed F-statistic in Table 5.6 under the column "Test II" is greater than the 99 per cent critical value for F(25,17) of 3.00, suggesting that the null hypothesis of no coefficient instability is rejected. The null hypothesis of no coefficient instability is also rejected for the estimated income regression equation since the observed F-statistic is greater than the 99 per cent critical value for F(25,19). Thus, the Dowrick test reinforces the Beggs test.

It can be observed from the reported regression on education that there was a downward change in the intercept of the regression line (from 13.9231 for no university down to 3.9483) and an increase in the regression line's slope. This indicates a substantial increase in the coefficient of no recent illness due to the university location dummy, pointing towards an upward rotation in the regression line.

With no university, a 10 percentage point increase in the proportion of persons with no recent illness caused a 1.923 percentage point decrease in the proportion of persons aged 15 years or over with post-school qualifications versus a 1.640 percentage point increase with a university, ceteris paribus. The regression estimates on income also indicate that the slope and intercept of the regression line change significantly between the university - and no university - locations. There was a downward change in the intercept of the regression line due to university (from 0.5356 for no university down to -6.2325). Other things being equal, with no university a 10 percentage point increase in the proportion of persons with each of no chronic condition and self-assessed good health (age 18 or over) caused \$119 and \$341 decreases in nominal gross annual median income of persons aged 15 years or over versus \$159 and \$322 increases for university, respectively. These results imply that universities positively effect education and income in communities in which they are located. The number of university places may also have positive effects on health, education, and income.

Table 5.6. Summary Statistics for Changes in Intercept and Slope of Three Indicators of Health in Education and Income Functions Due to the University Dummy: Two-Stage Least Squares (TSLS) Estimates

Dependent Variable/ Coefficient of	Constant	Interacted Variable: No recent illness (H_1)	No chronic condition (H_2)	Self-assessed good health (H_3)	R^2	D.W.	The parameter stability test: Test I[1]	Test II[2]
Education (E)[3]					0.831	1.759	F(1,41)	F(25,17)
No university	13.9231	-0.1923					= 10.585	= 8.067
University	3.9483	0.1640						
Change in coefficient	-9.9748	0.3563						
Income (Y)[4]					0.944	1.750	F(3,41)	F(25,19)
No university	0.5356		-0.0119	-0.0341			= 36.387	= 34.103
University	-6.2325		0.0159	0.0322				
Change in coefficient	-6.7681		0.0278	0.0663				

Notes:

1. For the test procedure see Beggs (1988, p.97).
2. For the test procedure see Dowrick (1993, p.2 and pp.26-27).
3. The estimated results of the predicted values of H_1, H_2, H_3, Y, the university dummy (1 for regions with a university and zero elsewhere), its interacted variable with H_1, and the constant term are -0.1923 (0.486), 0.1811 (0.956), 0.2857 (0.775), 1.4869 (3.160), -9.9748 (1.343), 0.3563 (1.411), and 13.9231 (0.467) with the absolute t-values in parentheses, respectively.
4. The estimated results of the predicted values of H_1, H_2, H_3, E, the university dummy, its interacted variable with each of H_2 and H_3, and the constant term are 0.1092 (1.260), -0.0119 (0.244), -0.0341 (0.440), 0.0900 (1.804), -6.7681 (1.269), 0.0278 (0.759), 0.0663 (1.080), and 0.5356 (0.072) with the absolute t-values in parentheses, respectively.

3,4. The estimated results of the control variables are given in Table 5.12.

The results and analysis of all equations estimated allow for the following summary remarks to be made regarding the propositions tested:

1. Other things being equal, a set of three indicators of good health (no recent illness, no chronic condition, and self-assessed good health), education, and income are positively inter-related. By the definitions, communities with relatively higher levels of education and income produce, on average, better health in terms of no recent illness and no chronic condition for children as well as for adults.

2. The standard errors of the ordinary least squares (OLS) and two-stage least squares (TSLS) estimates of the five production functions are small. In a comparison of the OLS and TSLS estimates, the former is smaller than the latter for the proportion of persons with no recent illness, the proportion of persons aged 15 years or over with post-school qualifications, and nominal gross annual median income of persons aged 15 years or over. Therefore, the OLS estimates are preferred for these three regressions. On the other hand, the standard errors of the TSLS estimates for the proportion of persons with no chronic condition and the proportion of persons aged 18 years or over with self-assessed good health are smaller than those of the OLS estimates. Therefore, the TSLS estimates are preferred for these two regressions. Furthermore, the Durbin-Watson statistics for the proportion of persons with no chronic condition and the proportion of persons aged 18 years or over with self-assessed good health estimated by TSLS indicate that the TSLS regressions are correctly specified.

3. The OLS estimates suggest that a higher proportion of persons aged 15 years or over with post-school qualifications and a higher level of nominal gross annual median income of persons aged 15 years or over within a community are associated with a higher proportion of persons with no recent illness.

4. The TSLS estimates suggest that a higher proportion of persons aged 15 years or over with post-school qualifications and a higher level of nominal gross annual median income of persons aged 15 years or over in a community are associated with a higher proportion of persons with no chronic condition.

5. The TSLS estimates suggest that a higher proportion of persons aged 15 years or over with post-school qualifications and a higher level of nominal gross annual median income of persons aged 15 or over within a community are associated with a higher proportion of persons aged 18 years or over with self-assessed good health.

6. The OLS estimates suggest that a higher proportion of persons having no recent illness, no chronic condition and self-assessed good health (age 18 or over), and a higher level of nominal gross annual median income of persons aged 15 or over within a community are associated with a higher proportion of persons aged 15 years or over with post-school qualifications. It also suggests that the level of education within a community is positively related to past family income.

7. The OLS estimates suggest that a higher proportion of persons with no recent illness, no chronic condition, self-assessed good health (age 18 or over), and post-school qualifications (age 15 or over) within a community is associated with higher levels of nominal gross annual median income of persons aged 15 years or over.

8. The regression estimates indicate that the three indicators of health and education are normal necessities. Given the empirical observations, government expenditures on three indicators of health and education financed by a progressive income tax are progressive and such a policy would reduce health and educational inequalities. Note that the approach could be applied under the assumption that the community's preferences for health and education versus their consumption goods represent an aggregation of individual preferences.

9. The observed high values of multiple determination (R^2) in the second stage of TSLS estimates imply that the TSLS method works well. The econometric results from the TSLS estimates are consistent with those from the OLS estimates with two exceptions: the estimated education and income regression equations. The differences may be due to the university dummy variable on three indicators of good health. Furthermore, it has been observed from the reduced-form equations that universities affect three indicators of good health (no recent illness, no chronic condition, and self-assessed good health) indirectly, through education and income. The coefficient stability analysis indicates that universities have positive effects on education and income to communities. Given the regression results, increasing the number of university places may have positive effects on health, education, and income.

5.1.3 Impact of the Gender Distribution

This subsection examines the inter-relationship between differentials in health and the gender distribution. It has been suggested that universities yield beneficial health effects to communities in which they are located, due to beneficial education and income effects. On the other hand, it is also suggested that communities with a university campus have many service industries such as finance and business services, recreation, personal and other services, and community services. Given that the service sector is female-intensive, women move from a community without a major service sector to a community with it. An implication of this is that the proportion of the population which is female is higher in communities with a university than in communities without a university. Since women generally have better health than men, the combined effects of a university on health and gender imply that health status has an impact on the gender distribution.

The regression results presented in Table 5.7 suggest that an increase in the proportion of the population which is female (F) in a community, ceteris paribus, may increase the proportions of persons with no recent illness, and no chronic condition, and the proportion of persons aged 18 years or over with self-assessed good health.[38] For example, the estimates suggest that a 10 percentage point increase in the proportion of the population which is female in a community, all other things being equal, increases the proportion of persons aged 18 years or over with self-assessed good health by 7.312 to 8.213 percentage points.

Table 5.7. Estimates of the Impact of Gender on Three Indicators of Health

Suggested Regressor of Endogeneity[2,5]	Dependent Variables[1]					
	No recent illness (H_1)		No chronic condition (H_2)		Self-assessed good health (H_3)	
	OLS	TSLS	OLS	TSLS	OLS	TSLS
Gender (F)	0.4956 (1.744)*	0.4861 (1.588)	0.2337 (0.453)	0.3065 (0.628)	0.7312 (2.119)**	0.8213 (2.586)**
Independence Test[3]	$\chi^2(4)$ = 4.945		$\chi^2(4)$ = 15.823***		$\chi^2(4)$ = 11.800**	
Exogeneity (joint) Test[4]	F(1,41) = 7.344***		F(1,41) = 8.067***		F(1,43) = 4.674**	
S.E. of estimates	2.200		3.989		2.753	

Notes: 1. Values in parentheses are the estimated absolute t-values. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

2. The estimated results of all other control variables are given in Table 5.11.

3. The test of Independence is a likelihood ratio test between each of dependent variables and gender with the number of degrees of freedom in parentheses. The contingency tables between each pair of variables are given in Appendix Tables A-29, A-30, and A-31. For the test procedure see Lewis, O'Brien, and Thampapillai (1990, pp.197-201).

4. The exogeneity test is a variant of the Hausman test as described by Beggs (1988, p.96) applied to the suspected regressor of endogeneity. The G*(SA+NT) variable has too low tolerance (1.00E-04 limits reached) in each of the H_1 and H_2 equations for the exogeneity test. The G*TAS variable has too low tolerance (1.00E-04 limits reached) in the H_3 equation for the test. Therefore, they are not entered into the respective equations for the test.

5. The reduced-form estimates on gender and three indicators of health are given in Appendix Table A-32.

---

[38] These results are consistent with the Schultz (1993, p.727) argument that because the decline in mortality has been generally more favourable for females than for males, other objective indicators of health will confirm a similar pattern.

In order to test the null hypothesis of independence between a set of three indicators of good health and gender distribution, each variable is classified as the top third, the middle third, and the bottom third of regions in terms of the sample proportions as introduced by Lewis, O'Brien, and Thampapillai (1990). The null hypothesis of independence between each pair of variables is rejected if the calculated value for Chi-Square (χ^2) is larger than the critical value. The 90 per cent and 99 per cent critical values for $\chi^2(4)$ are 7.779 and 13.277, respectively.

The results reported in Table 5.7 reveal that the null hypothesis of independence between no recent illness and gender distribution is not rejected.[39] On the other hand, the null hypothesis of independence between a set of the other two indicators of good health and gender distribution is rejected. For example, this implies that the proportion of statistical regions from each of three categories of the proportion of persons aged 18 years or over with self-assessed good health is not the same for the proportion of statistical regions from each of three categories of the proportion of the population which is female.[40]

In order to reinforce this conclusion, a variant of the Hausman test for exogeneity as described by Beggs (1988, p.96) is also conducted. The test involves augmenting the OLS regression with residuals of the suspected regressors of endogeneity from the reduced-form estimations, and then testing the joint significance of the added variables in the original OLS regression by the F-test.[41] The null hypothesis of exogeneity between the suspected

---

39 The acceptance of the null hypothesis may simply reflect that sixty one statistical regions are not sufficient numbers of observations to allow a full evaluation for the test of independence between the two.

40 The contingency table given in Appendix Table A-29 indicates that the percentage of the top third of regions in terms of self-assessed good health are 16.4 per cent and 8.2 per cent for the top third of regions and the bottom third of regions in terms of gender distribution (F). The table also implies that the percentages of the top third of regions in terms of gender distribution are 16.4 per cent and 3.3 per cent for the top third of regions and the bottom third of regions in terms of self-assessed good health, respectively. Therefore, self-assessed good health and gender distribution are positively dependent on one another.

41 For this purpose, the linear structural equations on the proportion of persons with each of three indicators of health (7), (8), and (9) are modified as follows; in a functional form,
$H_1 = H(F, E, Y, R_2, R_3, OCC_1, C, B_1, B, G_s{*}STATES_j, METD, GNG_8, GNG_9, GNG_{10})$
$H_2 = H(F, E, Y, R_2, R_3, OCC_1, C, B_1, B, G_s{*}STATES_j, METD, GNG_8, GNG_9, GNG_{10})$
$H_3 = H(F, E, Y, R_2, R_3, OCC_1, C, G_s{*}STATES_j, METD, GNG_8, GNG_9, GNG_{10})$
where j = VIC, QLD, WA, TAS and (SA+NT) and where the subscript i denoting the individual statistical region is omitted from the functions for the sake of convenience.

regressors is rejected if the F-statistic is larger than the critical value. The 99 per cent critical values for $F(1,40)$ and $F(1,42)$ are 7.31 and 7.27, respectively. The 95 per cent critical value for $F(1,42)$ is 4.07. The regression results presented in Table 5.7 suggest that there is strong statistical evidence of endogeneity between a set of three indicators of good health and gender. For example, the exogeneity test for the proportion of persons with self-assessed good health rejects the null hypothesis that gender distribution is exogenous to the dependent variable, since the observed F-statistic for the joint significance is larger than the 95 per cent critical value for $F(1,43)$ of 4.03.

From the independence and exogeneity tests, it can be suggested that a set of three indicators of good health and gender distribution are inter-related. Thus, it is suggested that universities yield beneficial health effects to communities in which they are located and that communities with a university campus have many service industries. Given that the service sector is female-intensive, the regression results suggest that the proportion of the population which is female is higher in communities with a university than in communities without a university.

In order to obtain the residuals of the suspected regressors of endogeneity from the reduced-form estimations, the above functions are expressed as the following reduced-form of the model for specification I,
$V = V(E, Y, R_2, R_3, OCC_1, C, B_1, B, G_s{*}STATES_j, METD, GNG_8, GNG_9, GNG_{10})$
where $V = (H_1, H_2, F)'$ refers to the row vector of the suspected regressors of endogeneity in parenthesis. Note that the H_1 and H_2 functions consist of the same explanatory variables.
For specification II,
$V = V(E, Y, R_2, R_3, OCC_1, C, G_s{*}STATES_j, METD, GNG_8, GNG_9, GNG_{10})$
where $V = (H_3, F)'$ refers to the row vector of the suspected regressors of endogeneity in parenthesis.
The reduced-form estimates for specification I and II are given in Appendix Table A-32.

On the other hand, Table 5.8 shows that different levels of education and income are influenced by the gender distribution, due to differences in health.[42] More specifically, REG(TSLS) estimates in the table suggest that communities with a university have a higher proportion of the population which is female than communities without a university due to beneficial health effects, which in turn reflect higher levels of education and income (education, in particular).

Table 5.8. Estimates of Total (Indirect) Effect of Gender on Education and Income through Three Indicators of Health[1]

Dependent Variables					
Education (E)			Income (Y)		
OLS	TSLS	REG (TSLS)[2]	OLS	TSLS	REG (TSLS)[2]
0.1924	0.6036	1.9417	0.0744	0.0196	0.3810

Notes: 1. No direct effect and not the coefficient of the reduced-form estimates.
2. REG (TSLS) performs the coefficient stability tests due to communities with a university in the two-stage least squares (TSLS) estimates as analysed in Table 5.6.

42 The indirect effect of the gender distribution on education and income are obtained from the linear structural equations (7), (8), (9), (10), and (11):

$$H_{1i} = a_1E_i + a_2Y_i + a_5F_i \quad (7)'$$
$$H_{2i} = b_1E_i + b_2Y_i + b_5F_i \quad (8)'$$
$$H_{3i} = c_1E_i + c_2Y_i + c_5F_i \quad (9)'$$
$$E_i = d_1H_{1i} + d_2H_{2i} + d_3H_{3i} + d_4Y_i \quad (10)'$$
$$Y_i = e_1H_{1i} + e_2H_{2i} + e_3H_{3i} + e_4E_i \quad (11)'$$

where all other explanatory variables and constant terms are not shown and where the subscript i denotes the individual statistical region. Substituting (7)', (8)', and (9)' into each of (10)' and (11)' yields

$$\alpha_0E_i = \alpha_1Y_i + \alpha_2F_i \quad (12)$$
$$\beta_0Y_i = \beta_1E_i + \beta_2F_i \quad (13)$$

where $\alpha_0 = 1 - d_1a_1 - d_2b_1 - d_3c_1$, $\alpha_1 = d_1a_2 + d_2b_2 + d_3c_2 + d_4$,
$\alpha_2 = d_1a_5 + d_2b_5 + d_3c_5$, $\beta_0 = 1 - e_1a_2 - e_2b_2 - e_3c_2$,
$\beta_1 = e_1a_1 + e_2b_1 + e_3c_1 + e_4$, $\beta_2 = e_1a_5 + e_2b_5 + e_3c_5$
From (12) and (13),

$$dE_i/dF_i = (\alpha_1\beta_2/\alpha_0\beta_0 + \alpha_2/\alpha_0)/(1 - \alpha_1\beta_1/\alpha_0\beta_0) \quad (14)$$
$$dY_i/dF_i = (\alpha_2\beta_1/\alpha_0\beta_0 + \beta_2/\beta_0)/(1 - \alpha_1\beta_1/\alpha_0\beta_0) \quad (15)$$

The expressions (14) and (15) are the total indirect effects of the proportion of females (F) on the proportion of persons aged 15 years or over with post-school qualifications (E) and nominal gross annual median income for persons aged 15 years or over (Y), respectively.

In brief, the estimations suggest that the higher the proportion of the population which is female in a community, ceteris paribus, the more likely it is that its population will be healthy and able to enjoy a full and long life. Given that communities with a university have many service industries and that the service sector is female-oriented, the estimations also suggest that communities with a university have a higher proportion of the population which is female than communities without a university, due to beneficial health effects, which in turn yield a higher proportion of persons aged 15 years or over with post-school qualifications and a higher nominal gross annual median income of persons aged 15 years or over. Given the estimations, we can suggest if universities offered more places to females, there would be community benefits in terms of health and income.

5.1.4 Impact of the Skill Level

The positive association between education and income is one of the most consistent empirical findings of the human capital literature. The conventional view, the productivity augmenting view, is that education enhances earnings via the production of marketable skills - a fundamental way of increasing productivity [Lewis (1991)]. It is suggested implicitly from this view that the level of skill is influenced by income, through education. Further, skilled workers are more attracted to communities where income is expected to be higher. Therefore, communities with higher income tend to have people with higher levels of skill.

On the other hand, lack of self-esteem raises the perceived costs of education or training, while a lack of optimism or a present orientation limits the expected benefits [Ehrenberg and Smith (1985)]. This suggests that communities with better health are probably much more likely to undertake human capital investments to become more skilled even though the payoffs are distant. Thus, communities with better health also tend to have people with higher levels of skill.

Given these issues, the main objectives of this subsection are to determine direction of causation between both education and income, and the skill level and to analyse the indirect effect of health on the skill level via education and income.

The coefficient estimates in Table 5.9 under the column "REG(TSLS)" on income suggest that communities with a university have a higher proportion of skilled workers to all employees aged 15 years or over (OCC_1) than communities without a university, due to beneficial health and education effects, which in turn generate a higher nominal gross annual median income of persons aged 15 years or over.

In order to observe the direction of causation between both education and income, and the skill level, the null hypothesis of independence between each pair of variables is conducted. Each variable is classified as the top third, the middle third, and the bottom third of regions in terms of the sample proportions [Lewis, O'Brien, and Thampapillai (1990)]. The null hypothesis is rejected because the calculated value for Chi-Square (χ) is larger than the 95 per cent critical value for $\chi(4)$ of 9.488.[43]

In order to reinforce this conclusion and analyse the indirect effects of health on skill levels via education and income, a variant of the Hausman test for exogeneity as described by Beggs (1988, p.96) is also conducted.[44] The null hypothesis of exogeneity is rejected if the F-statistic is larger than the critical value. If the null hypothesis is rejected, it is assumed that there is evidence of endogeneity among the suspected regressors. The regression

---

43 The contingency table given in Appendix Table A-29 indicates that the percentages of the top third of regions in terms of education are 19.7 per cent and 3.3 per cent for the top third of regions and the bottom third of regions in terms of skill level. This suggests that communities with higher level of education have people with higher levels of skill.

44 For this purpose, the linear structural equations on education and income, (10) and (11), are modified as follows; in a functional form,

$E = E(OCC_1, Y, H_1, H_2, H_3, R_2, R_3, OCC_2, C, NGSTD, UD, AE, GNG_1, GNG_2, GNG_3, GNG_4, GNG_6, GNG_{10})$

$Y = Y(OCC_1, E, H_1, H_2, H_3, R_2, R_3, C, METD, K_s, Ln, I, PFL, GNG_2, GNG_9, GNG_{10})$

where the subscript i denoting the individual statistical region is omitted from the functions for the sake of convenience.

In order to obtain the residuals of the suspected regressors of endogeneity from the reduced-form estimations, the above functions are expressed as the following reduced-form of the model:

$V = V(H_1, H_2, H_3, R_2, R_3, OCC_2, C, NGSTD, UD, AE, METD, K_s, Ln, I, PFL, GNG_1, GNG_2, GNG_3, GNG_4, GNG_6, GNG_9, GNG_{10})$

where $V = (E, Y, OCC_1)'$ refers to the row vector of the suspected regressors of endogeneity in parenthesis.

Table 5.9. Estimates of the Impact of Skill Level on Education and Income

Suspected Regressors of Endogeneity[2]	Dependent Variables[1]					
	Education (E)			Income (Y)		
	OLS	TSLS	REG (TSLS)[3]	OLS	TSLS	REG (TSLS)[3]
Skill Level (OCC_1)	0.0209 (0.157)	0.0166 (0.120)	0.0025 (0.018)	0.0838 (3.766)***	0.0826 (3.318)***	0.1004 (3.896)***
Education (E)				0.0919 (0.2832)***	0.1049 (2.085)**	0.0900 (1.804)*
Income (Y)	1.4550 (4.008)***	1.3467 (2.894)***	1.4869 (3.160)***			
Independence Test[4]	$\chi^2(4)$ = 10.801**			$\chi^2(4)$ = 21.314***		
Exogeneity (joint) Test[5]	F(2,38) = 10.112***			F(2,40) = 45.706***		
S.E. of estimates	3.447			0.802		

Notes: 1. Values in parentheses are the estimated absolute t-values. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

2. The estimated results of all other control variables are given in Table 5.12.

3. REG (TSLS) performs the coefficient stability tests due to regions with a university in the two-stage least squares (TSLS) estimates as analysed in Table 5.6.

4. The test of independence is a likelihood ratio test between each of dependent variables and skill level with the number of degrees of freedom in parentheses. The contingency tables between each pair of variables are given in Appendix Tables A-29 and A-30. For the test procedure see Lewis, O'Brien, and Thampapillai (1990, pp.197-201).

5. The exogeneity test is a variant of the Hausman test as described by Beggs (1988, p.96) applied to the suspected regressors of endogeneity.

results presented in Table 5.9 suggest that there is strong evidence of endogeneity among education, income, and the level of skill, since the observed F-statistics for the joint significance are larger than the 99 per cent critical values for F(2,38) of 5.21 and F(2,40) of 5.18, respectively.

From the independence and exogeneity tests, it can be suggested that an increase in the proportion of persons aged 15 years or over with post-school qualifications within a community increases the proportion of skilled workers to all employees aged 15 years or over, which in turn enhances nominal gross annual median income of persons aged 15 years or over. The results also suggest that nominal gross annual median income of persons aged 15 years or over has a positive effect on the proportion of skilled workers to all employees aged 15 years or over, due to a positive effect on the proportion of persons aged 15 years or over with post-school qualifications, and that communities with higher income have people with a higher level of skill because skilled workers are more attracted to communities where income is expected to be higher.

Table 5.10. Estimates of Total (Indirect) Effect of Three Indicators of Health on Skill Level through Education and Income[1]

No recent illness (H_1)	No chronic condition (H_2)	Self-assessed good health (H_3)
-0.1060	0.0380	0.1812

Notes: 1. No direct effect. Therefore, the indirect effect equals the total effect; an estimate of the reduced-form skill level parameter of three indicators of health given in Appendix Table A-31.

On the other hand, Table 5.10 shows that communities with better health have more people with higher levels of skill, due to the higher levels of education and income. For example, a coefficient of 0.0380 suggests that a 10 percentage point increase in the proportion of persons with no chronic condition (H_2), other things being equal, increases the proportion of skilled workers to all employees aged 15 years or over by 0.380

percentage points, due to increases in the proportion of persons aged 15 years or over with post-school qualifications and nominal gross annual median income of persons aged 15 years or over. The estimated negative reduced-form coefficient of the proportion of persons with no recent illness (H_1) may be due to the bias from the reduced-form equation that results when no functional form definition for the proportion of skilled workers to all employees aged 15 years or over is made.

In brief, the estimations suggest that education enhances earnings via the production of marketable skills -- a fundamental way of increasing productivity; an increase in the proportion of persons aged 15 years or over with post-school qualifications within a community increases the proportion of skilled workers to all employees aged 15 years or over, which in turn enhances nominal gross annual median income of persons aged 15 years or over.

The estimations also suggest that an increase in nominal gross annual median income of persons aged 15 years or over of a community increases the proportion of skilled workers to all employees aged 15 years or over, due to an increase in the proportion of persons aged 15 years or over with post-school qualifications. On the other hand, the estimations suggest that communities with better health and higher income attract people with higher levels of skill in the sense that skilled workers are more attracted to communities where better health and higher income are expected.

5.1.5 Impact of the Control Variables

For each statement in this subsection the qualifications 'ceteris paribus' and 'on average' are to be understood. Table 5.11 and Table 5.12 present both the ordinary least squares (OLS) and the two-stage least squares (TSLS) estimates of each of five production functions for three indicators of health status, education and income. The S.E. of the OLS and TSLS estimates of the five production functions are small. In a comparison of the OLS and TSLS estimates, the former is smaller than the latter for the proportion of persons with no recent illness, the proportion of persons aged 15 years or over with post-school qualifications, and nominal gross annual median income of persons aged 15 years or over. Therefore, we concentrate on an analysis of the econometric results from the OLS estimates for these three regressions.

On the other hand, the standard errors of the TSLS estimates for the proportions of persons with either no chronic conditions or self-assessed good health (H_2 and H_3) are smaller than those of the OLS estimates. Furthermore, the D.W. for these two regression equations estimated by TSLS are 2.255 and 1.963, respectively, suggesting that the null hypothesis of no misspecification is not rejected at the 0.05 level of significance [Gerdtham and Jönsson (1992)]. Therefore, the TSLS regressions are correctly specified. On the basis of these considerations, we concentrate on an analysis of the econometric results from the TSLS estimates for the proportion of persons with either no chronic conditions or self-assessed good health (age 18 or over).

From the proportions of persons with no recent illness, no chronic condition, and self-assessed good or excellent health aged 18 years or over equations presented in Table 5.11, the positive signs of the estimated coefficients of the proportion of persons aged 18 years or over with moderate alcohol consumption (R_3) suggest that a 10 percentage point increase in the proportion of persons aged 18 years or over with moderate alcohol consumption in a community increases the proportion of persons with no recent illness, no chronic condition, and self-assessed good health (age 18 or over) by 4.625 percentage points, 1.677 percentage points, and 2.253 percentage points, respectively. By their definitions, these results reveal that communities with a relatively higher proportion of persons aged 18 years or over with moderate alcohol consumption produce better health for children as well as for adults. Therefore, moderate drinking yields beneficial physical and psychological effects which have beneficial health effects [Hamilton and Hamilton (1993)].

Since the no drinking (R_2), moderate drinking, and excessive drinking (R_4) variables sum to unity, the positive signs of the estimated coefficients of no drinking and moderate drinking in Table 5.11 imply that the beneficial health effects of moderate drinking deteriorate as alcohol use increases [National Health Strategy, Research Paper No.1 (1992)]. In Appendix Table A-16, estimates of the reduced-form three health indicators parameters of no drinking and moderate drinking also suggest that excessive drinking has total detrimental effects on good health. For example, the direct, indirect, and total effects of excessive drinking on no chronic condition can be summarised as follows;

No Chronic Conditions (H_2)[1]

Direct	Indirect[2]	Total
-0.5172	-0.3102	-0.8274

Note: **1.** See Footnote 6
2. Subtracted the direct effect from the total effect.

The above estimates suggest that a 10 percentage point decrease in the proportion of persons aged 18 years or over with excessive alcohol consumption (R_4) within a community increases the proportion of persons with no chronic conditions by 8.274 percentage points; by 5.172 percentage points directly and by 3.102 percentage points indirectly, due to increases in the proportion of persons aged 15 years or over with post-school qualifications and nominal gross annual median income of persons aged 15 years or over via increases in the proportion of persons with no recent illness and the proportion of persons aged 18 years or over with self-assessed good health. The estimates imply that communities with a relatively lower proportion of persons aged 18 years or over with excessive alcohol consumption produce better health for children as well as for adults.

The bias that results when differences in endowment of health at the beginning of the period are unobserved and when no correction is made is obvious for the estimated coefficients of both the proportion of skilled workers aged 15 years or over to all employees

(OCC_1) and the proportion of persons with doctor consultations (C).[45] For example, the negative signs of the estimated coefficients of the proportion of persons with doctor consultations in the health functions suggest the higher proportion of persons who see a doctor more often when they are in poor health at the beginning of the period.[46]

The estimated coefficients for the rate of acute care private hospital beds per thousand population (B_1) carry the expected positive signs. A 10 percentage point increase in acute care private hospital beds per thousand population increases proportions of persons with no recent illness and no chronic condition by 16.725 percentage points and 16.875 percentage points, respectively. The estimated coefficients of the rate of acute care hospital beds per thousand population (B) on proportions of persons with no recent illness and no chronic condition carry the negative and positive signs, respectively. The unexpected negative sign of the estimated coefficient suggests the higher proportion of persons who go to the acute care public hospital when they are in poor health in terms of recent illness at the beginning of the period.[47]

---

45 There is evidence that the lower down the socio-economic scale one goes, the higher the incidence of sickness. For example, Narendranthan, Nickell, and Metcalf (1985, p.258) use information from the National Training Survey conducted on behalf of the Manpower Services Commission which provides a unique retrospective longitudinal data set for 17,708 British males in the labour force over the period 1965-75. They find from a logit analysis that the probability of sickness spells is negatively correlated with schooling and that sickness spells reduce dramatically as people rise through the social class, with those unskilled in 1975 being ten times more prone to lengthy sickness spells as professionals in the previous ten years. Their results are consistent with the Culyer's result (1976, p.21) from data of the 1973 General Household Survey - Introductory Report London that semi and unskilled manual workers have a substantially higher incidence of sickness than any other social class in England and Wales for 1971 (males only).

46 For more details on this point see Kemna (1987, p.201).

47 It is suggested from these estimations that the rate of acute care public hospital beds per 1,000 persons has a positive effect on the proportion of persons with each of no recent illness and no chronic condition. Evidence is given in Newhouse and Friedlander (1980, p.211).

Table 5.11. Estimates of the Control Variables on Three Indicators of Health[1]

Control Variables	Dependent Variables[2]					
	No recent illness (H_1)		No chronic condition (H_2)		Self-assessed good health (H_3)	
	OLS	TSLS	OLS	TSLS	OLS	TSLS
R_2	0.6666 (5.504)***	0.6723 (5.039)***	0.7216 (3.286)***	0.8666 (4.071)***	0.0301 (0.213)	0.1031 (0.786)
R_3	0.4625 (3.322)***	0.4597 (3.093)***	0.1643 (0.651)	0.1677 (0.707)	0.2478 (1.579)	0.2253 (1.568)
F	0.4956 (1.744)*	0.4861 (1.588)	0.2337 (0.453)	0.3065 (0.628)	0.7312 (2.119)**	0.8213 (2.586)**
OCC_1	-0.3454 (4.140)***	-0.3455 (3.745)***	-0.4222 (2.791)***	-0.5074 (3.448)***	-0.0398 (0.466)	-0.0807 (1.017)
C	-0.7272 (4.403)***	-0.7375 (4.124)***	-1.4848 (4.958)***	-1.6281 (5.707)***	-0.3710 (1.995)*	-0.4216 (2.457)**
B_1	1.6725 (2.834)***	1.6717 (2.650)**	1.5209 (1.421)	1.6875 (1.677)		
B	-0.0288 (0.176)	-0.0252 (0.144)	0.2271 (0.765)	0.2877 (1.031)		
METD	-4.6082 (3.412)***	-4.5578 (3.087)***	-2.3346 (0.953)	-3.0716 (1.304)	-1.0378 (0.631)	-1.8150 (1.185)
G*VIC	-0.3134 (0.688)	-0.2878 (0.582)	-0.7152 (0.866)	-0.2854 (0.362)	0.6240 (1.216)	0.9591 (1.989)*
G*QLD	-3.9523 (3.443)***	-3.9781 (3.186)***	-4.6144 (2.217)**	-5.6175 (2.820)***	1.0478 (0.739)	0.2284 (0.173)
G*WA	-2.8824 (2.586)**	-2.8625 (2.398)**	-5.4764 (2.710)***	-5.7320 (3.009)***	1.6033 (1.193)	1.4630 (1.190)
G*TAS	1.4274 (0.767)	1.4285 (0.703)	-5.7016 (1.691)*	-4.1349 (1.275)	2.0195 (0.882)	3.6629 (1.697)*
G*(SA+NT)	0.1084 (0.155)	0.1102 (0.145)	3.7025 (2.913)***	3.2786 (2.712)***	-1.0699 (1.269)	-1.5848 (2.012)*
GNG_8	0.5869 (0.977)	0.6507 (0.963)	2.7012 (2.479)**	3.6810 (3.414)***	0.6225 (0.841)	1.2099 (1.704)*
GNG_9	-0.3704 (1.834)*	-0.3687 (1.635)	-1.1537 (3.150)***	-1.3730 (3.816)***	-0.4009 (1.627)	-0.5957 (2.543)**
GNG_{10}	-0.2673 (1.900)*	-0.2645 (1.678)*	-0.5342 (2.094)**	-0.6821 (2.712)***	-0.2504 (1.426)	-0.4054 (2.403)**
Constant	-27.6073 (1.532)	-27.6287 (1.347)	14.9927 (0.459)	-9.9519 (0.304)	7.3012 (0.331)	-12.6340 (0.596)
R^2	0.759	0.726	0.776	0.803	0.630	0.692
S.E. of estimates	2.200	2.344	3.989	3.740	2.753	2.512
D.W.	2.429	2.366	2.315	2.255	1.796	1.963

Notes: 1. The reduced-form estimates are given in Appendix Table A-16.

2. Values in parentheses are the estimated absolute t-values. ***, **, and * indicate significance at the 1%, 5%, and 10% levels on a two-tailed test, respectively.

As expected, metropolitan communities (METD) have lower proportions of persons with no recent illness, no chronic condition, and self-assessed good health (age 18 or over) by 4.608 percentage points, 3.072 percentage points, and 1.815 percentage points over non-metropolitan communities. These results are consistent with the hypothesis that urbanisation has an adverse effect on health because of such factors as air and water pollution and congestion.

Given the positive effect of public health expenditure on better health,[48] this subsection examines the role of state government intervention in the health sector with a proxy variable (G*STATES) for nominal state and local governments outlays on health per person aged 15 years or over by states and territories (G). Such treatment allows one to separately compare the six different states and territories with New South Wales, the most populous of the six states and two territories, treated as a base; South Australia and Northern Territory are combined as one group. The regression results suggest that with two exceptions (Queensland and Australian Capital Territory) the role of state government intervention in the health sector in terms of no recent illness is greater in New South Wales than in all other states and territories.. Its role in the health sector in terms of no chronic condition is greater in New South Wales than in Victoria, South Australia and Tasmania, while its role is smaller in New South Wales than in Queensland, Western Australia, Northern Territory and Australian Capital Territory. With one exception (Australian Capital Territory) the role of state government intervention in the health sector in terms of self-assessed good health is greater in New South Wales than in all other states and territories. Note that Commonwealth funding data by statistical regions or by states and territories are not available in the Australian Bureau of Statistics (ABS) Information Consultancy Service for in-depth data investigation. Thus, Commonwealth outlays on health are excluded from the measurements.

---

[48] For example, Leu (1986, p.59), utilising cross-sectional data for OECD countries in 1974, finds that after per capita GDP and education are controlled for, public financing has a negative effect on neonatal mortality. Hitiris and Posnett (1992, p.179) also report the results for 20 OECD countries during the 28 years 1960-87 that mortality rates are negatively related to per capita total health expenditure when GDP per capita, the proportion of the population aged 65 and over, and the shift dummy for the U.K. are held constant. Moreover, Anand and Ravallion (1993, p.141) employ data for 22 developing countries from the 1990 World Development report on poverty (World Bank, 1990) and present the results that even when holding average income and poverty incidence constant, life expectancy is positively related to public health spending.

The estimated OLS regressions presented in Table 5.12 indicate that a 10 percentage point increase in the proportion of students in non-government schools to total students (NGSTD) increases the proportion of persons aged 15 years or over with post-school qualifications by only 0.011 percentage points.

The estimated coefficient on the university location dummy variable (UD) suggests that communities with a university have higher proportion of persons aged 15 years or over with post-school qualifications by 1.021 percentage points than communities without a university.

In Table 5.12, the proportion of persons aged 15 years or over with post-school qualifications and nominal gross annual median income of persons aged 15 years or over are negatively related to the proportion of persons aged 18 years or over with moderate alcohol consumption. However, the health economics literature suggests that moderate drinking yields beneficial health effects, which in turn carry over to the labour market, raising productivity and income.[49] Therefore, estimates of the reduced-form education and income parameters of moderate drinking given in Appendix Table A-16 are reported below:

Education (E)	Income (Y)
0.2797	0.0400

[49] For example, see Berger and Leigh (1988, p.1346), and Hamilton and Hamilton (1993, p.1).

Table 5.12. Estimates of the Control Variables on Education and Income[1]

Control Variables	Dependent Variables[2]					
	Education (E)			Income (Y)		
	OLS	TSLS	REG(TSLS)[3]	OLS	TSLS	REG(TSLS)[4]
R_2	-0.1980 (0.904)	-0.2725 (1.121)	-0.2467 (1.024)	-0.0404 (0.614)	-0.0024 (0.040)	-0.0442 (0.727)
R_3	-0.1550 (0.688)	-0.1419 (0.532)	-0.1406 (0.533)	-0.0421 (0.793)	-0.0042 (0.062)	-0.0458 (0.656)
OCC_1	0.0209 (0.157)	0.0166 (0.120)	0.0025 (0.018)	0.0838 (3.766)***	0.0826 (3.318)***	0.1004 (3.896)***
OCC_2	-0.2030 (0.921)	-0.0819 (0.344)	-0.0824 (0.350)			
C	0.0988 (0.350)	0.2759 (0.841)	0.1713 (0.515)	0.0838 (1.408)	0.0404 (0.526)	0.0885 (1.122)
NGSTD	0.0011 (0.012)	0.0518 (0.497)	0.0432 (0.418)			
UD	1.0205 (0.782)	0.2918 (0.193)	-9.9748 (1.343)			
K				0.0779 (0.459)	0.1476 (0.750)	0.0297 (0.141)
Ln				0.0858 (2.761)***	0.0902 (2.559)**	0.0818 (2.360)**
I				0.1273 (3.989)***	0.1533 (3.747)***	0.1326 (3.240)***
PFL				-0.0994 (3.871)***	-0.1095 (3.719)***	-0.1026 (3.494)***
GNG_1	-0.3183 (1.545)	-0.2911 (1.324)	-0.3757 (1.666)			
GNG_2	-0.7696 (1.818)*	-0.8777 (1.950)*	-0.9200 (2.064)**	0.3913 (4.507)***	0.3638 (3.723)***	0.4227 (4.203)***
GNG_3	-0.3341 (1.630)	-0.3525 (1.491)	-0.4028 (1.704)*			
GNG_4	0.9303 (1.424)	0.6854 (1.003)	0.9247 (1.328)			

Table 5.12. Continued[1]

Control Variables	Dependent Variables[2]					
	Education (E)			Income (Y)		
	OLS	TSLS	REG(TSLS)[3]	OLS	TSLS	REG(TSLS)[4]
GNG_6	-0.3926	-0.5562	-0.5569			
	(1.351)	(1.666)	(1.687)*			
GNG_9				0.1936	0.1553	0.1985
				(2.416)**	(1.516)	(1.944)*
GNG_{10}	-0.2996	-0.3038	-0.3606	0.1580	0.1245	0.1687
	(1.283)	(1.241)	(1.469)	(2.898)***	(1.833)*	(2.415)**
METD				1.1907	1.1192	1.2332
				(2.787)***	(2.375)**	(2.644)**
AE	0.3065	0.2302	0.2400			
	(2.037)**	(1.285)	(1.354)			
Constant	27.4308	3.4875	13.9231	-3.3321	-1.4514	0.5356
	(1.165)	(0.119)	(0.467)	(0.569)	(0.211)	(0.072)
R^2	0.826	0.822	0.831	0.947	0.937	0.944
S.E. of estimates	3.450	3.489	3.448	0.818	0.894	0.872
D.W.	1.625	1.752	1.759	1.728	1.698	1.750

Notes: **1.** The reduced-form estimates are given in Appendix Table A-16.

2. Values in parentheses are the estimated absolute t-values. ***, **, and * indicate significance at the 1%, 5%, and 10% levels on a two-tailed test, respectively.

3,4. REG(TSLS) performs the coefficient instability analysis due to regions with a university. The estimated results of the predicted values of H_1, H_2, H_3, Y, the university dummy (1 for regions with a university and zero elsewhere), its interacted variable with H_1 and the constant term are -0.1923 (0.486), 0.1811 (0.956), 0.2857 (0.775), 1.4869 (3.160), -9.9748 (1.343), 0.3563 (1.411), and 13.9231 (0.467), respectively. Those of H_1, H_2, H_3, E, the university dummy, its interacted variable with each of H_2 and H_3, and the constant term are 0.1092 (1.260), -0.0119 (0.244), -0.0341 (0.440), 0.0900 (1.804), -6.7681 (1.269), 0.0278 (0.759), 0.0663 (1.080), and 0.5356 (0.072) with the absolute t-values in parentheses, respectively.

The above estimates of the reduced-form education and income parameters suggest that a 10 percentage point increase in the proportion of persons aged 18 years or over with moderate alcohol consumption within a community increases the proportion of persons aged 15 years or over with post-school qualifications and nominal gross annual median income of persons aged 15 years or over by 2.797 percentage points and $400, respectively, due to an increase in the proportion of persons with no recent illness, no chronic condition, and self-assessed good health (age 18 or over).

On the other hand, it can also be suggested from Table 5-12 and Appendix Table A-16 that the negative indirect effects of excessive drinking (R_4) on education and income via three indicators of good health offset the positive direct effects of excessive drinking on education and income;

Education (E)			Income (Y)		
Direct	Indirect	Total	Direct	Indirect	Total
0.1765	-0.3377	-0.1612	0.0413	-0.0593	-0.0180

The above estimates turn out that a 10 percentage point decrease in the proportion of persons aged 18 years or over with excessive alcohol consumption increases the proportion of persons aged 15 years or over with post-school qualifications and nominal gross annual median income of persons aged 15 years or over by 1.612 percentage points and $180, respectively, due to increases in the proportion of persons with each of no recent illness and no chronic conditions, and an increases in the proportion of persons aged 18 years or over with self-assessed good health.

The estimated coefficients of both the proportion of persons aged 15 years or over without union membership (Ln) and the proportion of persons aged 15 years or over who are employed in the twelve industries (I) carry the expected positive signs. The coefficients suggest that a 10 percentage point increase in the proportion of persons aged 15 years or over without union membership within a community increases nominal gross annual median income of persons aged 15 years or over by \$2,131, while a 10 percentage point increase in the proportion of persons aged 15 years or over with union membership (Lu) increases nominal gross annual median income of persons aged 15 years or over in a community by \$1,273.[50] This suggests that non-union members receive a wage earnings premium over union members in the community. These results are inconsistent with the Brown-Medoff hypothesis that unionisation has a substantial positive effect on wage.[51] One possible explanation is that non-union members in Australia may be over-represented in industries with higher earnings.

The estimated income equation suggests that the proportion of part-time workers to full-time workers (PFL) has a negative effect on nominal gross annual median income of persons aged 15 years or over. A coefficient implies that a 10 percentage point decrease in the proportion of part-time workers to full-time workers within a community increases the nominal gross annual median income of persons aged 15 years or over by \$99.40.

The regression results in Table 5.12 reveal that nominal gross annual median income of persons aged 15 years or over rises with the proportion of persons aged 15 years or over who are employed in the mining industry (GNG_2), in the finance and business services industry (GNG_9), and in the public administration and defence industry (GNG_{10}). The

---

50 Let

$Y = a_0 + a_1 Ln + a_2 I$ (1)

where all other explanatory variables are not shown for the sake of convenience and where $a_1 = 0.0858$ and $a_2 = 0.1273$ in Table 5.12.

By the definition,

$I = Ln + Lu$ (2)

Substituting (2) into (1) yields

$Y = a_0 + (a_1 + a_2) Ln + a_2 Lu$

Therefore, the coefficients of Ln and Lu are 0.2131 and 0.1273, respectively.

51 Brown and Medoff (1978, p.364) estimate the earnings functions for 38,065 female workers and 57,067 male workers in private sector with merged 1973-75 Current Population Survey (CPS) files. The coefficients estimates of the union membership dummies (1 for yes and 0 for no) to the log of the hourly wage in real terms for females as well as males imply significantly positive unionisation impacts of wages. The estimated coefficients on the union membership variables for females and males are 0.241 with its standard error of 0.007 and 0.211 with its standard error of 0.004, respectively. Additional variables in the earnings functions are the years of schooling, age, age squared, and a vector of region dummies.

estimated coefficients suggest that a 10 percentage point increase in the proportion of persons aged 15 years or over who are employed in each of the mining, finance and business services, and public administration and defence increases nominal gross annual median income of persons aged 15 years or over by $391.30, $193.60 and $158.00, respectively.

An analysis of the econometric results conducted in this subsection allows us to make the following comments regarding the impact of the control variables on each of five production functions for the proportions of persons with no recent illness, no chronic condition, self-assessed good health (age 18 or over) and post-school qualifications (age 15 or over), and nominal gross annual median income of persons aged 15 years or over:

1. The standard errors of the ordinary least squares (OLS) and two-stage least squares (TSLS) estimates of the five production functions are small. In a comparison of the OLS and TSLS estimates, the former is smaller than the latter for the proportion of persons with no recent illness, the proportion of persons aged 15 years or over with post-school qualifications, and nominal gross annual median income of persons aged 15 years or over. Therefore, the OLS estimates are preferred for these three regressions. On the other hand, the standard errors of the TSLS estimates for the proportion of persons with no chronic condition and the proportion of persons aged 18 years or over with self-assessed good health are smaller than those of the OLS estimates. Therefore, the TSLS estimates are preferred for these two regressions. Furthermore, the D.W. statistics for the proportion of persons with no chronic condition and the proportion of persons aged 18 years or over with self-assessed good health estimated by TSLS indicate that the TSLS regressions are correctly specified.

2. The proportion of persons aged 18 years or over with moderate alcohol consumption in a community has a positive effect on the proportion of persons with a set of three indicators of good health (viz., no recent illness, no chronic condition, and self-assessed good health). This suggests that moderate drinking yields beneficial physical

and psychological effects which have beneficial health effects. By the definition, it also suggests that communities with a relatively higher proportion of persons aged 18 years or over with moderate alcohol consumption produce better health for children as well as for adults.

3. It can also be suggested from the estimations that excessive drinking has detrimental effects on three indicators of good health. By their definitions, communities with a relatively lower proportion of persons aged 18 years or over with excessive alcohol consumption produce better health for children as well as for adults.

4. The bias that results when differences in endowment of health at the beginning of the period are unobserved and when no correction is made is obvious for the estimated coefficients of both the proportion of skilled workers to all employees aged 15 years or over and the proportion of persons with doctor consultation. Given the bias and the previous evidence, we could suggest the positive effect of the proportion of skilled workers to all employees aged 15 years or over in a community on the proportion of persons with a set of three indicators of good health, and a higher proportion of persons who see a doctor more often when they are in poor health at the beginning of the period.

5. The greater the rates of both private and public acute care public hospital beds per thousand population, the greater the proportion of persons with either no recent illness or no chronic condition.

6. Metropolitan communities have a lower proportion of persons with a set of three indicators of good health than non-Metropolitan communities.

7. The role of state government intervention in the health sector is greater in New South Wales than in Victoria, Queensland, South Australia, Western Australia, Tasmania, and Northern Territory. The role of state government intervention in the health sector

in terms of self-assessed good health is greater in New South Wales than in Australian Capital Territory, while the role of state government intervention in the health sector in terms of no recent illness and no chronic condition is smaller in New South Wales than in Australian Capital Territory.

8. The effect of non-government schools on education is positive.

9. Communities with a university have higher proportion of persons aged 15 years or over with post-school qualifications than communities without a university.

10. An increase in the proportion of persons aged 18 years or over with moderate alcohol consumption in a community increases the proportion of persons aged 15 years or over with post-school qualifications and nominal gross annual median income of persons aged 15 years or over, due to an increase in the proportion of persons with a set of three indicators of good health. Thus, an empirical analysis is consistent with the health economics literature that the beneficial health effects of moderate drinking carry over to the labour market, raising productivity and income. Given that moderate drinking reduces the levels of education and income, we reject the hypothesis that differences in education, income, and three indicators of good health are manifestations of moderate drinking.

11. The negative indirect effects of excessive drinking on education and income via a set of three indicators of good health offset the positive direct effects of excessive drinking on education and income. It turns out that communities with a relatively lower proportion of persons aged 18 years or over with excessive alcohol consumption have higher proportion of persons aged 15 years or over with post-school qualifications and nominal gross annual median income of persons aged 15 years or over, due to higher proportions of persons with no recent illness and no chronic condition and higher proportion of persons aged 18 years or over with self-assessed good health.

12. Communities with a relatively higher proportion of persons aged 15 years or over without union membership have higher nominal gross annual median income of persons aged 15 years or over, compared to communities with relatively higher proportion of persons aged 15 years or over with union membership. This suggests that non-union members receive wage earnings premium over union members in the community. Given the Brown-Medoff hypothesis that unionisation has a substantial positive effect on wage, these results must be held to show that non-union members may be over-represented in industries with higher earnings.

13. An increase in the proportion of part-time workers to full-time workers has a negative effect on nominal gross annual median income of persons aged 15 years or over.

14. Nominal gross annual median income of persons aged 15 years or over rises with the proportion of persons aged 15 years or over who are employed in the mining industry, finance and business services industry, and public administration and defence industry.

5.2 LIFESTYLE FACTORS - SMOKING AND DRINKING BEHAVIOUR

5.2.1 Moderate Drinking

Different patterns of behaviour explain differentials in health; moderate drinking yields beneficial physical and psychological effects which have beneficial health effects. On the other hand, levels of education and income affect moderate drinking directly and indirectly through good health, implying that two causal relationships potentially exist; moderate drinking affecting health, and health affecting moderate drinking. However, these beneficial health effects deteriorate as alcohol use increases, together with a negative effect of smoking on health. Thus, the main objectives of the present subsection are to carry out tests to estimate empirically the inter-relationship between a set of two indicators of good health (viz., no recent illness and self-assessed good health) and moderate drinking, examine the impact of education and income on moderate drinking, and analyse the detrimental impact of smoking and excessive drinking on health.[52] For this purpose, the health model for self-assessed good health used is modified appropriately to allow for such empirical analysis.

5.2.1.1 Diagnostic Testing of the Hypothesis

Beggs (1988, p.99) argues that the more tests that are carried out, the less the chance there is of accepting a poor model. On the basis of the Beggs argument, diagnostic testing of the inter-relationships between a set of two indicators of good health and moderate drinking is conducted. It is suggested implicitly that no chronic condition and moderate drinking have no inter-relationships. In the first stage of testing, the linear and natural logarithmic versions of the function are contrasted to each other using non-nested tests to determine which model is a better representation of behaviour. Then only the diagnostic tests of the linear functions which have survived stage one are reported.

Table 5.13 shows a set of results for non-nested tests of linear versus double natural logarithmic models in a set of two indicators of good health and moderate drinking. Each of the MacKinnon-White-Davidson PE (1983, pp.54-56) test, the Beggs (1988, p.95) test, and the Bera-McAleer BM test as described by Maddala (1992, p.222) is reported as a t-test on the added regressor.

The MacKinnon-White-Davidson PE test is executed by augmenting the linear model with the difference between the natural logarithm of the prediction of the linear model and the prediction of the natural logarithmic model from the ordinary least squares

[52] Moderate drinking (R_3) is measured as the proportion of persons aged 18 years or over with the average daily consumption of alcohol greater than zero but less than 50 millilitres or 35 grams for males and 25 millilitres or 17.5 grams for females on a regular basis. Therefore, excessive drinking (R_4) is measured as the proportion of persons aged 18 years or over with the average daily consumption of alcohol greater than 50 millilitres or 35 grams for males and 25 millilitres or 17.5 grams for females. Smoking (R_1) is measured as the proportion of persons aged 18 years or over with the average daily consumption of one or more cigarettes (or pipes or cigars).

Table 5.13. Non-Nested Test of Double Logarithmic versus Linear Models of Each of No Recent Illness, Self-assessed Good Health, and Moderate Drinking[1]

Equation	Test I[2]		Test II[3]		Test III[4]		Box-Cox[5]	
	Logarithmic model (H_1)	Linear model (H_0)	Logarithmic model (H_1)	Linear model (H_0)	Logarithmic model (H_1)	Linear model (H_0)	Logarithmic model (H_1)	Linear model (H_0)
No recent illness	t = 3.544***	t = 1.513 Accept H_0	t = 3.418***	t = 1.608 Accept H_0			RSS = 0.289	RSS = 0.245 Accept H_0
Self-assessed good health	t = 3.835***	t = 0.827 Accept H_0	t = 3.485***	t = 1.017 Accept H_0			RSS = 0.060	RSS = 0.053 Accept H_0
Moderate drinking	t = 1.371 Reject H_0	t = 0.279 Accept H_0	t = 1.048 Reject H_0	t = 0.350 Accept H_0	t = 0.650 Reject H_0	t = 0.346 Accept H_0	RSS = 0.186	RSS = 0.177 Accept H_0

Notes:
1. The t-statistic denotes the estimated absolute value. *** and ** indicate significance at the 1% and 5% levels, respectively. By "accept H_0" we strictly mean "cannot reject H_0".
2. For the test procedure see MacKinnon, White and Davidson (1983, pp.54-56).
3. For the test procedure see Beggs (1988, p.95).
4. The Bera-McAleer BM test as described by Maddala (1992, p.222).
5. The Box-Cox procedure as described by Maddala (1977, p.317).

(OLS) estimates, and augmenting the natural logarithmic equation with the difference between the predictions of the linear equation and the exponential of the prediction of the natural logarithmic equation from the OLS estimates.[53]

The Beggs test involves augmenting the linear equation with the exponential of the prediction of the dependent variable from the natural logarithmic equation and the natural logarithmic model with the prediction of the dependent variable from the linear model run by OLS.

The Bera-McAleer BM test involves augmenting the linear equation with the residual of the natural logarithm of the prediction of the linear equation from the OLS estimates and the natural logarithmic equation with the residual of the exponential of the prediction of the natural logarithmic equation from the OLS estimates.

The null hypothesis (H_0) of the linear form of a model against the double natural logarithmic model is not rejected if the absolute value of the t-statistic on the added regressor of the linear model is less than the critical t-value, whereas the alternative hypothesis (H_1) of the double natural logarithmic model against the linear model is not rejected if the t-statistic on the added regressor of the double natural logarithmic model is less than the critical value. If the null hypothesis is not rejected, it is assumed that the linear model is preferred to the double natural logarithmic model. The no rejection of the null

---

[53] The values of $\ln B_1$, $\ln GNG_1$, $\ln GNG_2$, $\ln GNG_4$, $\ln GNG_8$, and $\ln GNG_{10}$ are less than zero. Therefore, the prediction of the double natural logarithmic model for the non-nested tests is obtained by relying upon the first-order Taylor series approximation $\ln(1+x) \approx x$. For example, $\ln B_1 = \ln(1+B_1-1) \approx B_1-1$. Expressing the linear structural equations (13), (14), and (15) in natural logarithms results in the following functional forms:

$\ln H_1 \approx H(\ln R_3, \ln E, \ln Y, \ln R_2, \ln F, \ln OCC_1, \ln G_s{*}STATES_j, \ln C, B_1\text{-}1, \ln B, METD, GNG_8\text{-}1, \ln GNG_9, GNG_{10}\text{-}1)$

$\ln H_3 \approx H(\ln R_3, \ln E, \ln Y, \ln R_4, \ln F, \ln OCC_1, \ln G_s{*}STATES_j, \ln C, METD, GNG_8\text{-}1, \ln GNG_9, GNG_{10}\text{-}1)$

$\ln R_3 \approx R(\ln H_1, \ln H_3, \ln E, \ln Y, \ln H_2, \ln R_1, \ln AGE_3, \ln AGE_4, \ln AGE_5, \ln FAGE_3, \ln FAGE_4, \ln FAGE_5, \ln G_s{*}STATES_j)$

where j = VIC, QLD, WA, TAS and (SA+NT), and where the subscript i denoting the individual statistical region is excluded from the model for the sake of convenience. Table 3.1 in Chapter 3 provides a description of the variables as well as their means and standard deviations.

hypothesis suggests that the added regressor of the linear model is asymptotically uncorrelated with the disturbances of the model and the associated estimated regression coefficient is asymptotically zero. If the alternative hypothesis is not rejected, it is assumed that the double natural logarithmic model is preferred to the linear model. If both the null and alternative hypotheses are not rejected, the tests are said to be inconclusive.

In comparing a set of two indicators of good health functions, the linear model is preferred to the double natural logarithmic model because the estimated t-statistic on the added regressor of the linear model is less than the critical t-value at the 0.05 level of significance. In the moderate drinking function, the estimated t-statistic on the added regressor of both the models is less than the critical value of 2.020 for forty-two degrees of freedom at the 0.05 level of significance on a two-tailed test, implying that the tests are inconclusive. Therefore, the Box-Cox procedure as described by Maddala (1977, p.317) in Table 5.13 under the column "Box-Cox" is reported. The Box-Cox procedure for the moderate drinking function (R_3) involves dividing each R_{3i} by the geometric mean of the R_3's (the exponential of the mean of the natural logarithm of E) and estimating the two equations. Given the results between the two moderate drinking equations, we choose the linear model in the moderate drinking equation because the estimated value of the residual sum of squares is smaller in the linear model than in the double natural logarithmic model.

In order to test the null hypothesis of independence, each dependent variable is classified as the top third, the middle third and the bottom third of regions in terms of the sample proportions as suggested by Lewis, O'Brien and Thampapillai (1990, pp.197-201). The null hypothesis of independence between each pair of variables is rejected if the calculated value for Chi-Square (χ^2) is larger than the critical value. If the null hypothesis is rejected, there is strong evidence that one is dependent on the other.

The results reported in Table 5.14 reveal that the null hypothesis of independence between a set of two indicators of good health and moderate drinking is rejected. For example, the test rejects the null hypothesis of independence between self-assessed good health and moderate drinking because the calculated value for χ^2 of Pearson (4), Likelihood ratio (4), and Mantel-Haenszel (1) are 15.717, 15.111, and 11.848, respectively, with the number of degrees of freedom in parentheses. The 99 per cent critical values for this statistic with four degrees of freedom and with one degree of freedom are 13.277 and 6.635, respectively. Therefore, the test rejects the null hypothesis of independence, concluding that the proportion of statistical regions from each of three categories of self-assessed good health is not the same for the proportion of statistical regions for each of three categories of moderate drinking.[54]

Table 5.14. Tests of Independence Between Each of No Recent Illness and Self-assessed Good Health, and Moderate Drinking[1]

Cross-classified Variables[3]	Value of Chi-Square[2] Pearson (4)	Likelihood ratio (4)	Mantel-Haenszel (1)	Decision
No recent illness and Moderate drinking	13.658***	13.270**	7.663***	Reject H_0
Self-assessed good health and Moderate drinking	15.717***	15.111***	11.848***	Reject H_0

Notes: **1.** For the test procedure see Lewis, O'Brien, and Thampapillai (1990, pp.197-201).
2. The number of degrees of freedom in parentheses. *** and ** indicate significance at the 1% and 5% levels, respectively.
3. The contingency tables between each pair of variables are given in Appendix Tables A-36 and A-37.

[54] It appears from the contingency table given in Appendix Table A-37 that the percentages of the top third of regions in terms of self-assessed good health are 19.7 per cent and 4.9 per cent for the top third of regions and the bottom third of regions in terms of moderate drinking respectively. The table also implies that the percentages of the top third of regions in terms of moderate drinking are 19.7 per cent and 4.9 per cent for the top third of regions and the bottom third of regions in terms of self-assessed good health, respectively. Therefore, self-assessed good health and moderate drinking are positively dependent on one another.

This conclusion is reinforced by the tests for exogeneity as proposed by Beggs (1988, p.96). The Beggs test is executed by augmenting the ordinary least square (OLS) regression with the residuals of the suspected regressors of endogeneity from the OLS estimates, and then testing the joint significance. For example, the tests for exogeneity of no recent illness and self-assessed good health in the moderate drinking equation involves augmenting the OLS regression for moderate drinking with the residual of each of the suspected regressors of endogeneity (H_1 and H_3) from each of the regressions.[55] Individual t-tests and the joint F-test are reported in Table 5.15 under the column "Exogeneity Tests".

For example, in the moderate drinking equation, the absolute values of the t-statistic on the residuals of a set of two indicators of good health (residual H_1 and residual H_3) are 1.991 and 5.050, respectively. The explanatory variables consist of the fifteen control variables, the two suspected regressors of endogeneity and their residuals. Including the intercept term the degrees of freedom are forty-one. The absolute values of the t-statistics are larger than the critical values. Further, the F-statistics for joint significance is larger than the 99 per cent critical value for F(2,39) of 5.20. Therefore, the tests reject the null hypothesis of exogeneity.

In sum, it can be concluded from the independence and exogeneity tests that, ceteris paribus, a set of two indicators of good health and moderate drinking are inter-related.

The RESET2 test for functional form misspecification involves augmenting the regression the square of the predicted value of the model and applying the t-test to the added coefficient. The null hypothesis of functional form misspecification is rejected if the absolute value of the t-statistic on the added regressor is less than the critical t-value. The

[55] In order to observe the detrimental effect of excessive drinking on health, the excessive drinking variable is included in the self-assessed good health equation. With the no-drinking variable in the no recent illness equation, the reduced-form equation for moderate drinking include both no-drinking and excessive drinking. Since the sum of the three drinking variables is 100 per cent, the coefficient of multiple determination for the reduced-form estimates on moderate drinking will be equal to 1.000. For this reason, the Hausman's form of the tests for exogeneity is not applied. Further, the simultaneity can be estimated by two-stage least squares (TSLS) for moderate drinking only.

regression results presented in Table 5.15 under the column headed by RESET2 indicate that the t-statistic of the added regressors in the three equations are less than the critical t-values at the 0.05 level of significance on a two-tailed test, respectively. Therefore, the null hypothesis of functional form misspecification is rejected in all the estimated OLS regressions for a set of two indicators of good health and moderate drinking, indicating that the assumption of zero expected value of residuals is not violated in any of the three OLS regressions.

On the other hand, Lewis, O'Brien, and Thampapillai (1990, p.296) argue that another indication of possible misspecification is low R^2 and F-statistic; that is, the proportion of variability in the dependent variable that is explained by the specified model is less than that which would be explained by the correct model, implying that the R^2 and F-statistic are lower than those for the correctly specified model. In the reported regressions, the observed values of R^2 are high and the F-statistics are larger than the 99 per cent critical values. Therefore, the models appear to be correctly specified.

In order to reinforce the above conclusions, the D.W. statistics are reported in Table 5.15 under the column headed by D.W. The regression results reveal that the null hypothesis of non-autoregressive residuals or no misspecification is not rejected at the 0.05 level of significance for all the regression equations estimated by OLS, suggesting that each of the three regression equations is correctly specified. For example, the value of D.W. of the estimated OLS regression for moderate drinking is 2.139. The tabulated upper boundary value (d_U) and the value of 4-d_U with 17 explanatory variables at the 5 per cent level of significance are 2.330 and 1.670, respectively. The observed D.W. value lies between d_U and 4-d_U. Therefore, the null hypothesis is not rejected, suggesting that the model is correctly specified.

Table 5.15. Diagnostic Evaluation for Each Equation for No Recent Illness, Self-assessed Good Health, and Moderate Drinking: The OLS Estimates

Equation	Summary and Diagnostic Statistics[1]										
		Exogeneity Tests[2]				Tests of Functional Form Misspecification[3]				Hetero-skedasticity[4]	Homo-skedasticity[5]
	S.E. of estimates	residual H_1	residual H_3	residual R_3	joint test	RESET 2	R^2	D.W.	F-Stat		
No recent illness (H_1)	2.223			t = 0.343	F(1,40) = 6.817**	t = 0.664	0.759	2.429	F(18,42) = 7.344***	F(21,40) = 0.783	χ^2 (18) = 12.708
Self-assessed good health (H_3)	2.472			t = 3.406***	F(1,42) = 6.142**	t = 0.330	0.630	1.796	F(16,44) = 4.674***	F(19,42) = 1.375	χ^2 (16) = 17.243
Moderate drinking (R_3)	2.623	t = 1.991*	t = 5.050***		F(2,39) = 8.982***	t = 0.375	0.685	2.139	F(17,43) = 5.511***	F(21,40) = 0.603	χ^2 (17) = 9.026

Notes:
1. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.
2. The tests for exogeneity are carried out by augmenting the regression with residuals of each of the suspected regressor(s) of endogeneity. Individual t-tests and the joint F test are reported. For the test procedure see Beggs (1988, p.96).
3. For the test procedure see Dowrick (1993, p.2), and Lewis, O'Brien and Thampapillai (1990, p.296 and pp.302-307)
4. For the test procedure see Dowrick (1993, p.2).
5. For the test procedure see Breusch and Pagan (1979, p.1288).

On the basis of RESET2 test, and the observed values of R^2, F and D.W., it can be concluded that correct specifications are implied in all the three estimated OLS regressions, indicating that the assumptions of zero expected values of residuals are not violated.

The tests for heteroskedasticity[56] proposed by Dowrick (1993, p.2) are performed by regressing the squared residual of the model on the explanatory variables. For example, the moderate drinking equation has twenty-one explanatory variables; fifteen control variables, the two suspected regressors of endogeneity, their residuals for the exogeneity test, and the squared prediction of the model for RESET2. Including the intercept term, the numerator degrees of freedom are 21 and hence the denominator degrees of freedom are 40. The estimated F-statistic equals 0.603, which is less than the 95 per cent critical value for $F(21,40)$ of 1.83. Thus the test cannot reject the null hypothesis of no heteroskedasticity. It can be concluded from the tests for heteroskedasticity that the assumption that the variances of the disturbances are approximately constant for all of the sixty-one statistical regions in the three equations cannot be rejected.

In order to reinforce this conclusion, the Breusch and Pagan (1979, p.1288) test for homoskedasticity is also conducted; under the null hypothesis of homoskedasticity, the number of sample observations times R^2 from the secondary regression executed by regressing the squared residual of the model has a Chi-Squared distribution with degrees of freedom equal to the number of non-constant explanatory variables in the secondary regression. From the reported regressions provided in Table 5.15 under the column "Homoskedasticity", the calculated Chi-Square values with eighteen, sixteen, and seventeen non-constant explanatory variables are 12.708, 17.243, and 9.026, respectively. The 95 per cent critical values for this statistic are 28.869, 26.296, and 27.587, respectively. Therefore, heteroskedasticity could not be detected in the three equations.

---

[56] Hetero meaning 'different'; homo meaning 'same'; skedastic from the Greek 'skedastos' meaning capable of being scattered. See Lewis, O'Brien, and Thampapillai (1990, p.297).

In summary, on the basis of the diagnostic statistical tests we reach the following conclusions:

1. The non-nested tests of the linear form of a model versus the double natural logarithmic model indicate that the linear model is preferred to the double natural logarithmic model in all the equations for a set of two indicators of good health and moderate drinking. Therefore, only the linear models of the three dependent variables are used for the application of all the diagnostic tests except the independence tests. Four tests are conducted; the MacKinnon-White-Davidson PE (1983) test, the Beggs (1988) test, the Bera-McAleer BM (Maddala, 1992) test, and the Box-Cox procedure (Maddala, 1977).

2. Both the independence (Lewis, O'Brien, and Thampapillai, 1990) and the exogeneity (Beggs, 1988) tests indicate that two causal relationships exist: moderate drinking affecting a set of two indicators of good health, and a set of two indicators of good health affecting moderate drinking.

3. Results from the RESET2 tests, the estimated values of R^2 and F (Lewis, O'Brien and Thampapillai, 1990), and the observed value of D.W. (Gerdtham and Jönsson, 1992) indicate that correct specifications are implied in the three estimated OLS regressions, indicating that the assumptions of zero expected value of residuals are not violated.

4. Both the Dowrick (1993) and Breusch-Pagan (1979) tests indicate that heteroskedasticity could not be detected in all equations estimated by OLS, implying that the assumption that the variances of the regression disturbances are approximately constant for all of the sixty-one statistical regions cannot be rejected.

5.2.1.2 Estimates of the Impact of Health Status, Education, and Income

It has been suggested that the estimated ordinary least squares (OLS) models may adequately describe the behavioural relationships. Moreover, each of the estimated regression equations run by OLS in Table 5.15, taken as a whole, significantly explains the variation in the dependent variable, since the estimated F-statistics are greater than the 99 per cent critical values for $F(18,42)$ of 2.40, $F(16,44)$ of 2.44, and $F(17,43)$ of 2.46. Therefore, it can be concluded that each of the three regression equations provides a statistically significant linear relationship between the dependent variable and the set of independent variables.

The standard errors of the estimates (S.E. of estimates) indicate that the smaller the variance of the sampling distribution, the greater is the precision of the estimator, that is, the greater is the chance of a sample estimate lying within some specified interval about the true value. From Table 5.16, it can be seen that the standard errors of the OLS estimates for a set of two indicators of good health and moderate drinking are small.

On the other hand, the standard error (S.E.) of the two-stage least squares (TSLS) estimates on moderate drinking is smaller than that of the OLS estimates, implying that the estimated TSLS regression is preferred.[57] Furthermore, the D.W. statistics for the moderate drinking equation estimated by TSLS is 2.292, suggesting that the null hypothesis of no misspecification is not rejected at the 0.05 level of significance [Gerdtham and Jönsson (1992)]. Therefore, the TSLS regression appears to be correctly specified. On the basis of these considerations, we concentrate on an analyses of the econometric results from the TSLS estimates for moderate drinking.

---

[57] The control variables of the OLS and TSLS regressions for moderate drinking are E, Y, H_2, R_1, AGE_3, AGE_4, AGE_5, $FAGE_3$, $FAGE_4$, $FAGE_5$, G_s*$STATES_j$ where J = VIC, QLD, WA, TAS and (SA+NT).
The control variables of the OLS regression for no recent illness are E, Y, R_2, F, OCC_1, G_s*$STATES_j$, C, B_1, B, METD, GNG_8, GNG_9, GNG_{10}.
The control variables of the OLS regression for self-assessed good health are E, Y, R_4, F, OCC_1, G_s*$STATES_j$, C, METD, GNG_8, GNG_9, GNG_{10}.
Their estimated results are reported in Table 5.16 and Table 5.22.

Table 5.16. Estimates of Each Function for No Recent Illness, Self-assessed Good Health, and Moderate Drinking

Explanatory Variables[2]	Dependent Variables[1] No recent illness (H_1)	Self-assessed good health (H_3)	Moderate drinking (R_3)	
	OLS	OLS	OLS	TSLS
No recent illness (H_1)			0.3640 (1.840)*	0.4957 (1.953)*
Self-assessed good health (H_3)			0.3296 (2.044)**	0.5831 (2.624)**
Moderate drinking (R_3)	0.4625 (3.322)***	0.2176 (1.917)*		
Education (E)	0.1065 (1.048)	0.2618 (2.116)**	0.3385 (2.688)**	0.2925 (2.194)**
Income (Y)	0.7771 (3.705)***	0.6523 (2.490)**	0.1504 (0.483)	0.0649 (0.218)
No chronic condition (H_2)			-0.2543 (2.143)**	-0.3363 (2.651)**
R^2	0.759	0.630	0.685	0.722
S.E. of estimates	2.200	2.753	3.264	3.067
D.W.	2.429	1.796	2.139	2.292

Notes: **1.** Values in parentheses are the estimated absolute t-values. *** and **, and * indicate significance at the 1%, 5%, and 10% levels on a two-tailed test, respectively.

2. The estimated results of other explanatory variables on each of no chronic condition, self-assessed good health, and moderate drinking are given in Table 5.22.

An analysis of the estimated regression equations allows us to make the following comments.

A set of two indicators of good health and moderate drinking are positively inter-related, ceteris paribus.

The OLS estimates on no recent illness suggest that a 10 percentage point increase in the proportion of persons aged 18 years or over with moderate alcohol consumption increases the proportion of persons with no recent illness by 4.625 percentage points. The OLS estimates also suggest that 10 percentage point increase in the proportion of persons aged 18 years or over with post-school qualifications and a $1,000 increase in the nominal gross annual median income of persons aged 15 years or over increase the proportion of persons with no recent illness by 1.065 percentage points and 0.777 percentage points, respectively.[58] This suggests that the higher the proportion of persons aged 18 years or over with moderate alcohol consumption, the higher the proportion of persons aged 15 years or over with post-school qualifications, and the higher the level of nominal gross annual median income of persons aged 15 years or over within a community, the healthier the region in terms of no recent illness for children as well as for adults.

The OLS estimates on self-assessed good health suggest that a 10 percentage point increase in the proportion of persons aged 18 years or over with moderate alcohol consumption in a community increases the proportion of persons aged 18 years or over with self-assessed good health by 2.176 percentage points. The OLS estimates also suggest that a 10 percentage point increase in the proportion of persons aged 15 years or over with post-school qualifications and a $1,000 increase in the nominal gross annual median income of

[58] Note that education is not significantly related to no recent illness when using the individual t-test. This neutrality effect of education on no recent illness may be attributed to a high correlation between education and income; $\gamma = 0.774$ as shown in Appendix Table A-8. Therefore, the income variable has been excluded from the equation. Then, in specification I in Appendix Table A-17, the education variable was found to be statistically significant at the 5% level (a t-statistic of 2.334 with the coefficient of 0.2499 in the OLS estimates). However, the standard error of the OLS estimates was increased from 2.200 to 2.533.

persons aged 15 years or over within a community increase the proportion of persons aged 18 years or over with self-assessed good health by 2.168 percentage points and 0.652 percentage points, respectively.

A 10 percentage point increase in the proportion of persons with either no recent illness or self-assessed good health (age 18 or over) increases the proportion of persons aged 18 years or over with moderate alcohol consumption by 4.957 percentage points and 5.831 percentage points, respectively. The TSLS estimates suggest that a 10 percentage point increases in the proportion of persons aged 15 years or over with post-school qualifications and a $1,000 increase in the nominal gross annual median income of persons aged 15 years or over within a community increase the proportion of persons aged 18 years or over with moderate alcohol consumption by 2.925 percentage points and 0.065 percentage points, respectively.

On the other hand, the proportion of persons with no chronic condition (H_2) within a community has a negative effect on the proportion of persons aged 18 years or over with moderate alcohol consumption. Thus the estimations suggest that communities with a higher proportion of persons with no chronic condition have a higher proportion of persons aged 18 years or over with excessive alcohol consumption (R_4).

Given the estimations of two causal relationships between a set of two indicators of good health and moderate drinking, we could expect that education and income influence moderate drinking directly and indirectly, through either no recent illness or self-assessed good health. The regression results presented in Table 5.17 suggest that a 10 percentage point increase in the proportion of persons aged 15 years or over with post-school qualifications and a $1,000 increase in the nominal gross annual median income of persons aged 15 years or over within a community increase the proportion of persons aged 18 years or over with moderate alcohol consumption by 2.521 percentage points and 7.027

percentage points, respectively, due to an increase in the proportion of persons with no recent illness and self-assessed good health (age 18 or over).

Table 5.17. Estimates of Direct, Indirect, and Total Effects of Education and Income on Moderate Drinking

Explanatory Variables	Direct effect[1]	Indirect effect	Total effect[2]
Education (E)	0.3385	0.2521	0.5906
Income (Y)	0.1504	0.7027	0.8531

Notes: 1. The OLS regression coefficients of education and income in the moderate drinking equation.
2. Not the coefficient of the reduced-form parameter.

These results are consistent with the hypothesis that the well educated are aware of the evidence that moderate drinking generates health improvements [Kenkel (1991)].[59]

On the other hand, the reduced-form estimates presented in Table 5.18 suggest that smoking and excessive drinking have negative effects on health in terms of no recent illness and self-assessed good health. A 10 percentage point reduction in the proportion of

[59] The generality of this point can also be demonstrated by the use of algebra. Consider the following estimable linear structural equations;

$$H_3 = f_1 R_3 + f_2 E \quad (1)$$

$$R_3 = g_1 H_3 \quad (2)$$

where all other explanatory variables and intercepts are not shown. Taking the total differentials,

$$dH_3 = f_1 dR_3 + f_2 dE \quad (3)$$

$$dR_3 = g_1 dH_3 \quad (4)$$

Substituting (3) into (4) and collecting term on dE, the total derivative of R_3 with respect to E can be obtained as:

$$dR_3/dE = g_1 f_2/(1-g_1 f_1) \quad (5)$$

The expression (5) is itself the ceteris paribus total (indirect) effect of education on moderate drinking:

$$dR_3/dE = \partial R_3/\partial H_3 * \partial H_3/\partial E/(1-\partial R_3/\partial H_3 * \partial H_3/\partial R_3) \quad (6)$$

where $f_1 = \partial H_3/\partial R_3$, $f_2 = \partial H_3/\partial E$ and $g_1 = \partial R_3/\partial H_3$.

Provided that moderate drinking improves health, then, $\partial H_3/\partial R_3 > 0$.

Assume further that $\partial R_3/\partial H_3 * \partial H_3/\partial R_3 < 1$. Then, the empirical observation of the value of $\partial H_3/\partial E$ greater than zero implies that the marginal productivity of education on moderate drinking must be positive as well. It has been observed from the regression results in Table 5.16 that $\partial H_3/\partial R_3 = 0.2176$ and $\partial H_3/\partial E = 0.2618$ from the OLS estimates, and $\partial R_3/\partial H_3 = 0.5831$ from the TSLS estimates. On the basis of this, we conclude that the well educated are aware of the evidence that moderate drinking generates health improvements.

persons aged 18 years or over with either cigarettes consumption (R_1) or excessive alcohol consumption increases the proportion of persons with no recent illness by 0.942 percentage points and 5.004 percentage points, respectively. The higher the proportion of persons aged 18 years or over with either cigarettes consumption or excessive alcohol consumption in a community the poorer the health of children and adults. The reduced-form estimates also suggest that a 10 percentage point reduction in the proportion of persons aged 18 years or over with either cigarettes consumption or excessive alcohol consumption in a community increases the proportion of persons aged 18 years or over with self-assessed good health by 4.148 percentage points and 0.246 percentage points, respectively.

Table 5.18. Estimates of Direct, Indirect, and Total Effects of Smoking and Excessive Drinking on No Recent Illness and Self-assessed Good Health

Explanatory Variables	Dependent Variables					
	No recent illness (H_1)[1]			Self-assessed good health (H_3)[2]		
	Direct	Indirect	Total	Direct	Indirect	Total
Smoking (R_1)		-0.0942	-0.0942		-0.4148	-0.4148
Excessive drinking (R_4)		-0.5004	-0.5004	-0.0301	0.0055	-0.0246

Notes: **1.** No direct effects of smoking and excessive drinking on no recent illness. Therefore, the indirect effects equal the total effects; the estimates of the reduced-form parameters given in Appendix Table A-44.

2. No direct effect of smoking on self-assessed good health. Therefore, the indirect effect equals the total effect; an estimate of the reduced-form parameter given in Appendix Table A-44. The direct effect of excessive drinking on self-assessed good health is the OLS regression coefficient.

The results and analysis of all equations estimated are summarised as follows:

1. Other things being equal, a set of two indicators of health (no recent illness and self-assessed good health) and moderate drinking are positively inter-related. The higher the proportion of persons aged 18 years or over with moderate alcohol consumption, the higher the proportion of persons aged 15 years or over with post-school

qualifications, and the higher the gross nominal median income of persons aged 15 years or over, the better the health of the region in terms of no recent illness for children and adults.

2. The OLS estimates suggest that a higher proportion of persons aged 18 years or over with moderate alcohol consumption, a higher proportion of persons aged 15 years or over with post-school qualifications, and a higher level of nominal gross annual median income of persons aged 15 years or over in a community are associated with higher proportions of persons with either no recent illness or self-assessed good health (age 18 or over).

3. The standard error of the TSLS estimates on moderate drinking is smaller than that of the OLS estimates, implying that the estimated TSLS regression is preferred. Furthermore, the Durbin-Watson statistics for the moderate drinking equation estimated by TSLS suggests that the TSLS regression appears to be correctly specified. The TSLS estimates suggest that a higher proportion of persons with no recent illness, a higher proportion of persons aged 18 years or over, self-assessed good health, a higher proportion of persons aged 15 years or over with post-school qualifications, and a higher level of nominal gross annual median income of persons aged 15 years or over, are associated with a higher proportion of persons aged 18 years or over with moderate alcohol consumption.

4. The TSLS estimates suggest that a higher proportion of persons with no chronic condition in a community is associated with a lower proportion of persons aged 18 years or over with moderate alcohol consumption. Thus, the estimations suggest that a higher proportion of persons with no chronic condition is associated with a higher proportion of persons aged 18 years or over with excessive alcohol consumption.

5. A higher proportion of persons aged 15 years or over with post-school qualifications and a higher level of nominal gross annual median income of persons aged 15 years or

over in a community are associated with a higher proportion of persons aged 18 years or over with moderate alcohol consumption directly and indirectly, through a higher proportion of persons with either no recent illness or self-assessed good health (age 18 or over). Given the econometric results, it also suggests that the well-educated are aware of the evidence that moderate drinking generates health improvements.

6. A lower proportion of persons aged 18 years or over with either cigarettes consumption or excessive alcohol consumption in a community is associated with a higher proportion of persons with no recent illness and a higher proportion of persons aged 18 years or over with self-assessed good health. A higher proportion of persons aged 18 years or over with either cigarettes consumption or excessive alcohol consumption results in a greater likelihood of recent illness for children and adults.

5.2.2 Smoking and Excessive Drinking

It has been argued in the previous subsection that, ceteris paribus, a higher proportion of persons aged 18 years or over with moderate alcohol consumption is associated with a higher proportion of persons with no recent illness and with a higher proportion of persons aged 18 years or over with self-assessed good health. It has also been argued that the higher the proportion of persons aged 18 years or over with moderate alcohol consumption in a community, the better health it will exhibit on average in terms of no recent illness for children as well as for adults. This suggests that moderate drinking yields beneficial physical and/or psychological effects which have beneficial health effects.

On the contrary, the estimations in the previous subsection suggest that either smoking or excessive drinking has a detrimental effect on the health in a community; ceteris paribus, a higher proportion of persons aged 18 years or over with either cigarettes consumption or excessive alcohol consumption is associated with a lower proportion of persons with a set of two indicators of good health. Thus, a reduction in the demand for either cigarettes or excessive alcohol use is likely to have a substantial positive effect on health. This suggests that addicts with lower rates of discount respond more to a detrimental effect of smoking and excessive drinking on health and that a higher future cost (due perhaps to greater information about health risks) is likely to reduce the demand for cigarettes and excessive alcohol use, particularly among the rich [Becker, Grossman, and Murphy (1991)]. Therefore, communities with a higher proportion of persons aged 15 years or over with post-school qualifications and a higher level of average income respond more to knowledge about the detrimental effects of smoking and excessive drinking on the population's health. Campaigns against smoking and excessive drinking may be more beneficial to health in such communities [Collins and Lapsley (1993)].

In the present subsection, the analysis is extended to allow for a change in smoking in response to a change in excessive drinking and for a change in excessive drinking in response to a change in smoking, so that smoking and excessive drinking may be related for the purpose of policy formulation.

5.2.2.1 Diagnostic Testing of the Hypothesis

In the first stage of testing, the linear and natural logarithmic versions of the function are contrasted to each other using non-nested tests to determine which model is a better representation of behaviour. Then only the diagnostic tests of the linear functions which have survived stage one are reported.

Table 5.19. Non-Nested Test of Double Logarithmic versus Linear Models of Each of Smoking and Excessive Drinking[1]

	Test I[2]		Test II[3]	
Equation	Logarithmic model (H_1)	Linear model (H_0)	Logarithmic model (H_1)	Linear model (H_0)
Smoking	t = 2.261**	t = 0.092 Accept H_0	t = 1.958*	t = 0.162 Accept H_0
Excessive drinking	t = 3.404***	t = 0.081 Accept H_0	t = 3.602***	t = 0.725 Accept H_0

Notes: 1. The t-statistic denotes the estimated absolute value. ***, **, and * indicate significance at the 1%, 5%, and 10% levels on a two-tailed test, respectively. By "Accept H_0" we strictly mean "cannot reject H_0".
2. For the test procedure see MacKinnon, White, and Davidson (1983, pp.54-56).
3. For the test procedure see Beggs (1988, p.95).

In the MacKinnon-White-Davidson PE test of the smoking function in Table 5.19 under the column "Test I", the observed t-statistics on the added regressor of the double natural logarithmic and linear models are 2.261 and 0.092, respectively.[60] In the Beggs test of the same function in the table under the column "Test II", the observed t-statistics on the added regressor of the two models are 1.958 and 0.162, respectively. The critical t-values for forty-seven degrees of freedom at the 0.10 and 0.05 levels of significance on a two-tailed test are 1.680 and 2.015, respectively.

On the other hand, in the excessive drinking function, the MacKinnon-White-Davidson PE test statistics on the added regressor of the double natural logarithmic and

[60] The values of $\ln GNG_1$, $\ln GNG_4$, and $\ln GNG_8$ are less than zero. Therefore, the prediction of the double natural logarithmic model for the non-nested tests is obtained by relying upon the first-order Taylor series approximation $\ln(1+x) \approx x$. For example, $\ln GNG_1 = \ln(1+GNG_1-1) \approx GNG_1-1$. Expressing the linear structural equations (17) and (18) in natural logarithms results in the following functional forms:

$$\ln R_1 \approx R(\ln R_4, \ln E, \ln Y, \ln AGE_8, \ln R_3, \ln NES, \ln Ln, \ln OCC_1, GNG_1-1, \ln GNG_3, \ln GNG_9, \ln GNG_{11})$$

$$\ln R_4 \approx R(\ln R_1, \ln E, \ln Y, \ln AGE_8, \ln R_3, \ln NES, \ln Ln, GNG_1-1, \ln GNG_3, GNG_4-1, \ln GNG_5, \ln GNG_6, \ln GNG_7, GNG_8-1)$$

where the subscript i denoting the individual statistical region is omitted from the functions for the sake of convenience. Table 3.1 in Chapter 3 provides a description of the variables as well as their means and standard deviations.

linear models are 3.411 and 0.081, respectively. The Beggs test statistics on the added regressor of the two models of the function are 3.415 and 0.725, respectively. The critical t-value for forty-five degrees of freedom at the 0.01 level of significance on a two-tailed test is 2.700.

The absolute values of the t-statistics on the added regressors of the linear models are less than the critical t-values, thus the null hypothesis (H_0) of the linear form of a model against the double natural logarithmic model is not rejected in either function. Therefore, non-nested model tests of both the smoking and excessive drinking functions suggest that the linear models are preferred to the double natural logarithmic models. Hence, we choose the linear model for both functions.

In order to test the null hypothesis of exogeneity between smoking and excessive drinking, a variant of the Hausman test for exogeneity is conducted. The joint F-tests are reported in Table 5.20 under the column "Exogeneity Tests". In the smoking equation, the F-statistic for the joint significance is 4.920. The explanatory variables consist of the eleven control variables, the one suspected variable, and its residual. The 95 per cent critical value for F(1,46) is 4.05. In the excessive drinking equation, the F-statistic for the joint significance is 5.547. The 95 per cent critical value for F(1,44) is 4.06. The F-statistics are larger than the 95 per cent critical values, and thus the null hypothesis of exogeneity is rejected. Therefore, the regression results suggest that, ceteris paribus, smoking and excessive drinking are inter-related.[61]

---

61 It appears from the contingency table given in Appendix Table A-39 that the percentages of the top third of regions in terms of smoking are 18.0 per cent and 6.56 per cent for the top third of regions and the bottom third of regions in terms of excessive drinking, respectively. The table also implies that the percentages of the top third of regions in terms of excessive drinking are 18.0 per cent and 9.84 per cent for the top third of regions and the bottom third of regions in terms of smoking, respectively. Therefore, smoking and excessive drinking are positively dependent on one another.

Table 5.20. Diagnostic Evaluation of Each of the Smoking and Excessive Drinking Equations: The OLS Estimates

Equation	Summary and Diagnostic Statistics[1]								
	Exogeneity Tests[2]		Test of Functional Form Misspecification[3]						
	S.E. of estimates	Joint Test	RESET2	RESET3	R^2	D.W.	F-Stat	Hetero-skedasticity[4]	Homo-skedasticity[5]
Smoking (R_1)	3.080	F(1,46) = 4.920**	t = 0.145		0.569	1.676	F(12,48) = 5.286***	F(15,46) = 0.875	χ^2 (12) = 12.165
Excessive drinking (R_4)	2.210	F(1,44) = 5.547**	t = 2.288**	F(2,44) = 5.854***	0.641	2.127	F(14,46) = 5.866***	F(18,43) = 1.694*	χ^2 (14) = 19.455

Notes:

1. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. The null hypothesis of independence between R_1 and R_4 is rejected at the 0.10 level of significance. The calculated value of Chi-Square of Mantel-Haenszel is 3.584 with one degree of freedom. For the test each variable is classified as the top third, the middle third, and the bottom third of regions in terms of the sample proportions. For more details see Lewis, O'Brien, and Thampapillai (1990, pp.197-201). The contingency table between the two variables is given in Appendix Table A-38.
2. The exogeneity test is a variant of the Hausman test as described by Beggs (1988, p.96) applied to the suspected regressor of endogeneity.
3. For the test procedure see Dowrick (1993, p.2), and Lewis, O'Brien and Thampapillai (1990, p.296 and pp.302-307). The third power of the prediction of the smoking variable (R_1) for RESET3 test has too low tolerance (1.00E-04 limits reached) and therefore the test is not performed.
4. For the test procedure see Dowrick (1993, p.2).
5. For the test procedure see Breusch and Pagan (1979, p.1288).

The RESET tests for functional form misspecifications involve adding higher order powers of the prediction of the model to the regression. The null hypothesis of functional form misspecification is rejected if the absolute value of the t-statistic on the added regressor is smaller than the critical t-value for RESET2 test.

The regression results for the smoking equation in Table 5.20 under the column headed by RESET2 indicate that the t-statistic of the added regressor is smaller than the critical value for forty-seven degrees of freedom at the 0.05 level of significance. Thus, the model is correctly specified.

On the other hand, the regression results for the excessive drinking equation indicate that the t-statistic of the added regressor is larger than the critical value for forty-five degrees of freedom at the 0.05 level of significance on a two-tailed test. Thus, the RESET3 test is also conducted. The null hypothesis of functional form misspecification is rejected if the F-statistic for the joint significance is smaller than the critical F-value. The F-statistic is 5.854, which is larger than the 99 per cent critical value for F(2,44) of 5.12. Therefore, misspecification is implied in this equation. However, the observed value of R^2 is not low; that is, 64.1 per cent of the total variation in the dependent variable is explained by variation in the full set of independent variables. The estimated F-statistic is 5.866, which is larger than the 99 per cent critical value for F(14,46) of 2.50. Further, the observed D.W. value lies between the tabulated upper boundary value (d_U) of 2.177 and 4-d_U of 1.823, indicating that the null hypothesis of non-autoregressive residuals or specification is not rejected [Lewis, O'Brien and Thampapillai (1990, p.296), and Gerdtham and Jönsson (1992, p.19)].

The test for heteroskedasticity proposed by Dowrick (1993, p.2) is performed by regressing the squared residual of the model on the explanatory variables. In the smoking equation, a set of explanatory variables consists of eleven control variables, the one suspected regressor of endogeneity, its residual for the exogeneity tests, and the squared prediction of the model for RESET2. Including the intercept term, the numerator degrees of freedom are 15 and hence the denominator degrees of freedom are 46. In the excessive drinking equation, a set of explanatory variables consists of thirteen control variables, the one suspected regressor of endogeneity, its residual for the exogeneity tests, the squared prediction of the model for RESET2, and the third power of the prediction of the dependent variable for RESET3. Including the intercept term, the numerator degrees of freedom are 18 and the denominator degrees of freedom are 43. The null hypothesis of no heteroskedasticity is rejected if the F-statistic is larger than the critical value. The regression results in Table 5.20 under the column "Heteroskedasticity" suggest that the test rejects the null hypothesis at the 0.10 level of significance in the excessive drinking equation,[62] whereas the test cannot reject the null hypothesis in the smoking equation.

Therefore, the Breusch and Pagan test for homoskedasticity is also conducted; under the null hypothesis of homoskedasticity, the number of sample observations times R^2 from the secondary regression executed by regressing the squared residual of the model has a Chi-Squared distribution with degrees of freedom equal to the number of non-constant explanatory variables in the secondary regression. From the reported regressions provided in Table 5.20 under the column "Homoskedasticity", the calculated values for $\chi^2(12)$ and $\chi^2(14)$ are 12.165 and 19.455, respectively. The 95 per cent critical values for this statistic with twelve and fourteen degrees of freedom are 21.026 and 23.685 respectively. Therefore, whilst heteroskedasticity could not be detected in the first equation, there may be some complicated higher order heteroskedasticity in the second equation. Accordingly, the standard errors and t-statistics in the excessive drinking equation may be biased in unknown ways.

---

62 The null hypothesis of no heteroskedasticity is not rejected in this equation when the third power of the prediction of the dependent variable for RESET3 is excluded from the test. The estimated F-statistic is 1.490, which is smaller than the 90 per cent critical value for F(17,44) of 1.66.

In summary, the diagnostic statistical tests conducted can be concluded as follows:

1. The non-nested tests of the linear form of a model versus the double natural logarithmic model indicate that the linear model is preferred to the double natural logarithmic model in both the smoking and excessive drinking equations. Two tests are conducted; the MacKinnon-White-Davidson PE (1983) and the Beggs (1988) tests.

2. Results from a variant of the Hausman test for exogeneity as described by Beggs (1988) indicate that, ceteris paribus, excessive drinking is highly endogenous to smoking and vice versa. The independence test and the contingency table (Lewis, O'Brien and Thampapillai, 1990) also indicate that one is positively dependent on one another.

3. The RESET2 test for functional form misspecification, the observed values of R^2 and F (Lewis, O'Brien and Thampapillai, 1990), and the observed D.W. value (Gerdtham and Jönsson, 1992) of the smoking equation indicate that the model is correctly specified. On the contrary, RESET2 and RESET3 tests indicate that misspecification is implied in the excessive drinking equation. However, the observed R^2 value of 0.641 is not low and the F-statistic of 5.866 is greater than the 99 per cent critical value for F(14,46). Further, the observed D.W. value of 2.127 lies between the tabulated upper boundary value (d_U) and 4-d_U at the 5 per cent level of significance.

4. The Dowrick (1993) test indicates that heteroskedasticity could not be detected in the smoking equation estimated by OLS, whereas heteroskedasticity could be detected in the excessive drinking equation. However, the Breusch-Pagan (1979) test indicates that heteroskedasticity could not be detected in the two estimated regressions.

5.2.2.2 Estimates of the Inter-relationship Between Smoking and Excessive Drinking, and Impact of Education and Income

It has been suggested that the estimated ordinary least squares (OLS) models may adequately describe the behavioural relationships. Moreover, each of the estimated regression equations run by OLS in Table 5.20, taken as a whole, significantly explains the variation in the dependent variable, since the estimated F-statistics are greater than the 99 per cent critical values for F(12,48) of 2.58 and F(14,46) of 2.50. Therefore, it can be concluded that both regression equations demonstrate a statistically significant linear relationship between the dependent variable and the set of independent variables.

Table 5.21 presents the standard errors of the OLS and TSLS estimates (S.E. of estimates) for the smoking and excessive drinking equations.[63] The standard errors of the OLS and TSLS estimates of the two equations are small. In a comparison of the OLS and TSLS estimates, the former is smaller than the latter.

On the basis of the possible loss in efficiency in TSLS estimation, without consistency being assured for the excessive drinking equation, we concentrate on an analysis of the econometric results from the OLS estimates.

---

[63] The control variables of the OLS and TSLS regressions for smoking are E, Y, AGE_8, R_3, NES, Ln, OCC_1, GNG_1, GNG_3, GNG_9, GNG_{11}.
The control variables of the OLS and TSLS regressions for excessive drinking are E, Y, AGE_8, R_3, NES, Ln, GNG_1, GNG_3, GNG_4, GNG_5, GNG_6, GNG_7, GNG_8. Their estimated results are reported in Table 5.21 and Table 5.23.

Table 5.21. Estimates of the Smoking and Excessive Drinking Equations[1]

Explanatory Variables	Dependent Variables[2]			
	Smoking (R_1)		Excessive drinking (R_4)	
	OLS	TSLS	OLS	TSLS
Smoking (R_1)			0.2573 (2.917)***	0.0616 (0.267)
Excessive drinking (R_4)	0.6426 (3.593)***	1.1061 (1.853)*		
Education (E)	0.0140 (0.117)	0.0052 (0.040)	0.0623 (0.683)	0.0252 (0.236)
Income (Y)	0.6130 (1.925)*	0.3356 (0.694)	0.4332 (1.966)*	0.5810 (2.023)**
R^2	0.569	0.490	0.641	0.575
S.E. of estimates	3.073	3.345	2.211	2.405
D.W.	1.676	1.756	2.127	2.247

Notes: **1.** Estimates of all other explanatory variables are given in Table 5.23.
2. Values in parentheses are the estimated absolute t-values. *** , **, and * indicate significance at the 1%, 5%, and 10% levels on a two-tailed test, respectively.

An analysis of the estimated linear regression equations in Table 5.21 allows us to make the following comments.

A 10 percentage point reduction in the proportion of persons aged 18 years or over with excessive alcohol consumption in a community reduces the proportion of persons aged 18 years or over with cigarettes consumption by 6.426 percentage points, and a 10 percentage point reduction in the proportion of persons aged 18 years or over with cigarettes consumption in a community reduces the proportion of persons aged 18 years or over with excessive alcohol consumption by 2.573 percentage points. This suggests that smoking and excessive drinking are positively inter-related, and therefore that they complement each other.

The OLS estimates suggest that the higher the proportion of persons aged 15 years or over with post-school qualifications and the higher the level of nominal gross annual median income of persons aged 15 years or over in a community, the higher the proportion of persons aged 18 years or over with harmful lifestyle factors such as smoking and excessive drinking. Addicts with lower rates of discount respond more to changes in harmful future consequences of addictive goods, such as negative effects of smoking and excessive drinking on health [Becker, Grossman, and Murphy (1991)]. Therefore, communities with lower rates of discount respond more to knowledge of the detrimental effects of smoking and excessive drinking on health.[64] It has been observed in the previous subsection that higher proportion of persons aged 18 years or over with either cigarettes consumption or excessive alcohol consumption in a community are associated with lower proportion of persons with no recent illness and lower proportion of persons aged 18 years or over with self-assessed good health. Given the importance of the population's health, the community would be better off if it substantially reduced tobacco smoking and increased the proportion of the population who have never smoked, and if it minimised the harm associated with alcohol abuse and increased the proportion of persons aged 18 years or over with moderate alcohol consumption [Conway, Pinyopusarerk, Carter, Penm, and Stevenson (1993)]. Given the estimations, the campaign against smoking and excessive drinking appears to be potentially beneficial to health in the community [Collins and Lapsley (1993)].[65]

---

64 This approach could be justified under the assumption that community's preferences for health represent an aggregation of individual preferences. For example, in examining the production model of health measured by mortality rates in logarithmic form across 51 states of the United States as the unit of observation in 1960 (sample size of white population in the labour force), Auster, Leveson, and Sarachek (1969, p.418) note that relationships among aggregates may depend on characteristics and resources among individuals.

65 Becker, Grossman, and Murphy (1991, p.239) also argue that a higher future cost (due perhaps to greater information about health hazards) is likely to reduce the demand for each of cigarettes and excessive alcohol use, particularly among the rich.

5.2.3 Impact of the Control Variables

For each statement in this subsection the qualifications 'ceteris paribus' and 'on average' are to be understood. The regression results in Table 5.22 indicate that the moderate drinking equation has the standard error (S.E.) of the TSLS estimates smaller than that of the OLS estimates, implying that the estimated TSLS regression is preferred. Furthermore, the Durbin-Watson statistics (D.W.) for the moderate drinking equation estimated by TSLS suggests that the null hypothesis of no misspecification is not rejected at the 0.05 level of significance [Gerdtham and Jönsson (1992)]. On the basis of these considerations, we concentrate on the TSLS estimates for moderate drinking.

On the other hand, the regression results in Table 5.23 indicate that in each equation the standard error of the OLS estimates is smaller than that of the TSLS estimates. Given this and our previously noted doubts about TSLS obtaining efficiency and consistency in the excessive drinking equation, this subsection concentrates on an analysis of the OLS econometric results for lifestyle factors in terms of smoking and excessive drinking.

Table 5.22. Estimates of Other Explanatory Variables in No Recent Illness, Self-assessed Good Health, and moderate Drinking Equations[1]

Explanatory Variables	Dependent Variables[2]			
	No recent illness (H_1)	Self-assessed good health (H_3)	Moderate drinking (R_3)	
	OLS	OLS	OLS	TSLS
R_1			0.0556 (0.371)	0.1302 (0.897)
R_2	0.6666 (5.504)***			
R_4		-0.0301 (0.213)		
F	0.4956 (1.744)*	0.7312 (2.119)**		
OCC_1	-0.3454 (4.140)***	-0.0398 (0.466)		
AGE_3			-0.4776 (1.093)	-0.6223 (1.499)
AGE_4			-0.0464 (0.115)	-0.0992 (0.262)
AGE_5			-0.1278 (0.249)	-0.0734 (0.152)
$FAGE_3$			0.0228 (0.057)	0.1280 (0.336)
$FAGE_4$			-0.6785 (1.748)*	-0.5902 (1.583)
$FAGE_5$			0.0057 (0.012)	0.0103 (0.022)
G*VIC	-0.3134 (0.688)	0.6240 (1.216)	0.3968 (0.654)	0.1760 (0.303)
G*QLD	-3.9523 (3.443)***	1.0478 (0.739)	-4.1394 (3.012)***	-4.0397 (3.016)***
G*WA	-2.8824 (2.586)**	1.6033 (1.193)	1.1851 (0.777)	0.3629 (0.242)
G*TAS	1.4274 (0.767)	2.0195 (0.882)	3.7625 (1.500)	2.6644 (1.107)
G*(SA+NT)	0.1084 (0.155)	-1.0699 (1.269)	-1.3603 (1.407)	-0.7638 (0.810)
C	-0.7272 (4.403)***	-0.3710 (1.995)*		
B_1	1.6725 (2.834)***			
B	-0.0288 (0.176)			

Table 5.22. (Continued)[1]

Explanatory Variables	Dependent Variables[2] No recent illness (H_1) OLS	Self-assessed good health (H_3) OLS	Moderate drinking (R_3) OLS	Moderate drinking (R_3) TSLS
METD	-4.6082 (3.412)***	-1.0378 (0.631)		
GNG_8	0.5869 (0.977)	0.6225 (0.841)		
GNG_9	-0.3704 (1.834)*	-0.4009 (1.627)		
GNG_{10}	-0.2673 (1.900)*	-0.2504 (1.426)		
Constant	-27.6073 (1.532)	10.3138 (0.535)	22.3018 (1.335)	4.1426 (0.224)
R^2	0.759	0.630	0.685	0.722
S.E. of estimates	2.200	2.753	3.264	3.067
D.W.	2.429	1.796	2.139	2.292

Notes: **1.** The reduced-form estimates are given in Appendix Table A-44.
2. Values in parentheses are the estimated absolute t-values. ***, **, and * indicate significance at the 1%, 5%, and 10% levels on a two-tailed test, respectively.

1. In the moderate drinking equation estimated by TSLS, the age distribution variables (AGE_3, AGE_4, and AGE_5) are negatively related to moderate drinking. The estimated coefficient for the proportion of females aged between 25 years and 44 years to total females ($FAGE_4$) carries the negative sign. This suggests that females in this age category are non-drinkers, on average. It also suggests that a higher proportion of persons aged 18 years or over without alcohol consumption (R_2) in a community is associated with a higher proportion of persons with good health in terms of no recent illness.

2. Given the inclusion of moderate drinking in the health functions,[66] the regression results suggest that with two exceptions (Queensland and Australian Capital Territory) the role of state government intervention in the health sector in terms of moderate drinking is greater in New South Wales than in Victoria, South Australia, Western Australia, Tasmania, and Northern Territory.

3. In Table 5.23, our estimates suggest that the proportion of persons aged between 25 and 74 years (AGE_8) within a community has a positive effect on the proportion of persons aged 18 years or over with excessive alcohol consumption. On the other hand, Table 5.22 shows that each of the proportions of persons aged between 15 and 24 years (AGE_3), between 25 and 44 years (AGE_4), and between 45 and 64 years (AGE_5) have a negative impact on the proportion of persons aged 18 years or over with moderate alcohol consumption. Given the estimations, it can be suggested that, on average, a higher proportion of persons aged between 18 and 44 years in a community is associated with a higher proportion of persons aged 18 years or over with excessive alcohol consumption, while a higher proportion of persons aged between 45 and 74 years is associated with a higher proportion of persons aged 18 years or over without alcohol consumption.[67]

4. Moderate drinking carries a negative sign in each of the estimated equations for smoking and excessive drinking. A 10 percentage point increase in the proportion of persons aged 18 years or over with moderate alcohol consumption in a community

---

66 For example, see Berger and Leigh (1988, p.1346).

67 Hamilton and Hamilton (1993, p.27) utilise cross-section data for 1,954 males between the ages of 20 and 59 from the 1985 General Social Survey (GSS) Health and Social Support collected by Statistics Canada. They provide the results of the multinomial logit regression for non and heavy drinkers. The base category is moderate drinkers (p.14). Hamilton and Hamilton suggest that increasing age has a negative effect on the propensity to be a heavy drinker, while no such distinction is found for non-drinkers. However, given the estimated positive coefficient of each of the age variables (age 30s with the estimated t-statistic of 0.786, age 40s with that statistic of 1.419, and age 50s with that statistic of 1.521) on non-drinkers, our view is that increasing age is less likely to be associated with moderate drinking than non-drinking. Note that Hamilton and Hamilton do not make this inference.

Table 5.23. Estimates of Other Explanatory Variables in the Smoking and Excessive Drinking Equations[1]

Explanatory Variables	Dependent Variables[2]			
	Smoking (R_1)		Excessive drinking (R_4)	
	OLS	TSLS	OLS	TSLS
AGE_8	-0.1690 (0.914)	-0.3827 (1.163)	0.4567 (4.041)***	0.4493 (3.647)***
R_3	-0.0021 (0.018)	0.1462 (0.666)	-0.2568 (3.311)***	-0.3147 (3.006)***
NES	0.1745 (1.596)	0.3127 (1.517)	-0.2992 (4.331)***	-0.2973 (3.955)***
Ln	-0.3072 (2.812)***	-0.3029 (2.544)**	0.0565 (0.648)	-0.0055 (0.048)
OCC_1	-0.1908 (1.873)*	-0.1966 (1.770)*		
GNG_1	-0.0589 (0.368)	0.0240 (0.119)	-0.0598 (0.550)	-0.0949 (0.764)
GNG_3	-0.1866 (1.151)	-0.0725 (0.323)	-0.1467 (1.293)	-0.1855 (1.425)
GNG_4			-0.3903 (1.042)	-0.3823 (0.938)
GNG_5			0.1106 (0.527)	0.2586 (0.931)
GNG_6			-0.0975 (0.568)	-0.1270 (0.671)
GNG_7			-0.2928 (1.157)	-0.3556 (1.254)
GNG_8			0.3439 (0.646)	0.2669 (0.456)
GNG_9	-0.5182 (2.246)**	-0.4540 (1.727)*		
GNG_{11}	-0.1513 (0.744)	-0.1433 (0.646)		
Constant	40.7686 (3.503)***	42.1809 (3.300)***	-15.2450 (1.665)	-4.7701 (0.318)
R^2	0.569	0.490	0.641	0.575
S.E. of estimates	3.073	3.345	2.211	2.405
D.W.	1.676	1.756	2.127	2.247

Notes: 1. The reduced-form estimates are given in Appendix Table A-45.

2. Values in parentheses are the estimated absolute t-values. ***, **, and * indicate significance at the 1%, 5%, and 10% levels on a two-tailed test, respectively.

reduces the proportions of persons aged 18 years or over with cigarettes consumption and excessive alcohol consumption by 0.021 percentage points and 2.568 percentage points, respectively. The implication is that moderate drinking is a substitute for excessive drinking.

5. A 10 percentage point increase in the proportion of persons aged 15 years or over who are employed in the finance, property, and business services industry (GNG_9) in a community reduces the proportion of persons aged 18 years or over with cigarettes consumption by 5.182 percentage points.

6. In the smoking equation, a positive sign of the coefficient of the constant term suggests that smoking is addictive. In the excessive drinking equation estimated, a negative sign of the coefficient of the constant term is ambiguous in explaining that excessive drinking is addictive. The evidence also suggests that females are less likely to be excessive drinkers.

6.0 SUMMARY AND POLICY IMPLICATIONS

6.1 SUMMARY OF PRINCIPAL FINDINGS

This study first concentrates on testing for Lewis' proposition (1994) that a set of three indicators of good health (viz., no recent illness, no chronic condition, and self-assessed good health), education, and income are positively inter-related. Given the empirical results that moderate drinking reduces the levels of education and income, we reject the hypothesis that differences in education, income, and three health indicators are manifestations of moderate drinking. It also demonstrates that excessive drinking has detrimental effects on a set of three health indicators and that the negative indirect effects of excessive drinking on education and income via a set of three indicators of good health offset the positive direct effects of excessive drinking on education and income.

A set of two indicators of health (viz., no recent illness and self-assessed good health) and moderate drinking are found to be positively inter-related. The implication is that the positive correlations among education, income, and moderate drinking develop only after the hypothesis that moderate drinking yields beneficial physical and psychological effects which have beneficial health effects. Thus, the mechanism behind the education-income-moderate drinking relationships may also give rise to the education-income and two health indicators relationship. It also indicates that more education and higher income affect moderate drinking directly and indirectly via better health.

The regression estimates indicate that both smoking and excessive drinking have detrimental effects on a set of two indicators of good health directly and indirectly via moderate drinking, suggesting that the negative correlations between moderate drinking and excessive drinking develop only after the hypothesis that moderate drinking is negatively associated with smoking. Thus, the mechanism behind the excessive drinking-moderate drinking relationships may also give rise to the excessive drinking-smoking relationship.

Given the importance of health, we attempt to test that smoking and excessive drinking are positively inter-related. This then suggests that differences in both smoking and excessive drinking are causal to differentials in three indicators of health. This implies that smoking and excessive drinking may be considered as a package in policy formulation and health outcomes.

The first specification of the model estimated suggests that the negative indirect effects of excessive drinking on education and income via three indicators of good health offset the positive direct effects of excessive drinking on education and income. Thus, differences in both education and income are causal to both smoking and excessive drinking differentials.

6.1.1 The Multiple Inter-relationships Among Health Status, Education, and Income

Previous studies have suggested that there are inter-relationships between better health and more education, between better health and higher income, and between more education and higher income.

More recently, Lewis has formulated testable propositions that (a) good health makes it easier to study and proceed through the educational system, and thus facilitates learning, (b) higher education leads to higher income, since productivity and income increase with the skills and knowledge, (c) higher income is generally associated with better health as a result of better diet, and access to the better medical care, (d) good health affects the ability to earn income through higher productivity, (e) higher income facilitates additional years of schooling, and (f) good education contributes to better health by fostering better knowledge about health issues, increasing receptabilty to health information, and increasing the ability to deal with health bureaucracies.

The relevance of this framework is that there are multiple positive inter-relationships among health status, education, and income, suggesting the validity of developing a full simultaneous equations model; ceteris paribus, an increase in each of these variables causes increases in the other two. This implies that the cross effects among health status, education, and income are positive. It turns out that they complement each other and therefore health and education can be viewed as normal goods.

These inter-relationships are explored using data from the 1989-90 National Health Survey (NHS) for 61 statistical regions of Australia. We test the influence of other explanatory variables on health, education, and income. Moreover, we conduct diagnostic tests of the five models, one for a set of three indicators of good health (viz. no recent illness, no chronic condition, and self-assessed good or excellent health) as well as education and income. Note that both education and income are important determinants of three indicators of good health and that all three indicators of health status are important determinants of the levels of education and income.

The overall results and analysis of the estimated models allow for the following summary remarks to be made regarding the propositions tested:

1. The non-nested tests of the linear form of a model versus the double natural logarithmic model indicate that the linear model is preferred to the double natural logarithmic model with respect to all the equations for a set of three indicators of good health, education, and income. Only the linear models of the five dependent variables are used for the application of diagnostic tests. The MacKinnon-White-Davidson PE (1983, pp.54-56) test, the Beggs (1988, p.95) test, the Box-Cox procedure [Maddala (1977, p.317)], and the Theil Prediction Criterion [Maddala (1992, p.497)] are conducted.

2. The independence [Lewis, O'Brien and Thampapillai (1990, pp.197-201)] and endogeneity [Dowrick (1993, p.2)] tests, and the contingency table indicate that, ceteris paribus, each of the three indicators of health, education, and income are inter-related as hypothesised (Lewis, 1994); education and income affecting health, health and income affecting education, and health and education affecting income.

3. Results from the RESET2 test, the observed values of R^2 and F [Lewis, O'Brien and Thampapillai (1990, p.296)], and the observed value of D.W. [Gerdtham and Jönsson (1992, p.19)] indicate that correct specifications are implied in the five estimated ordinary least squares (OLS) regressions, indicating that the assumption of zero expected values of residuals is not violated.

4. The Breush-Pagan (1979, p.1288) and Dowrick (1993, p.2) tests indicate that heteroskedasticity could not be detected in any of the equations estimated by OLS, implying that the assumption that the variances of the regression disturbances are approximately constant for all of the sixty-one statistical regions cannot be rejected.

5. The standard errors of both the OLS and two-stage least squares (TSLS) estimates for the five production functions are small. In a comparison of the OLS and TSLS estimates, the former is smaller than the latter for no recent illness, education, and income. Therefore, the estimated OLS regressions are preferred. On the other hand, the standard errors of the TSLS estimates for either no chronic condition or self-assessed good health are smaller than those of the OLS estimates, implying that the estimated TSLS regressions are preferred. Furthermore, the D.W. statistics for these two regression equations estimated by TSLS indicate that the TSLS regressions appear to be correctly specified.

6. Other things being equal, all three indicators of health, education, and income are positively inter-related (Lewis, 1994). A higher proportion of persons aged 15 years

or over with post-school qualifications and a higher level of nominal gross annual median income of persons aged 15 years or over produce, on average, better health in terms of no recent illness and no chronic condition for children as well as for adults.

7. The regression estimates suggest that a 10 percentage point increase in the proportion of persons aged 15 years or over with post-school qualifications and a $1,000 increase in the nominal gross annual median income of persons aged 15 years or over in a community increase the proportion of persons with no recent illness by 1.065 percentage points and 0.777 percentage points, respectively. The estimates also suggest that a 10 percentage point increase in the proportion of persons aged 15 years or over with post-school qualifications and a $1,000 increase in the nominal gross annual median income of persons aged 15 years or over in a community increases the proportion of persons with no chronic condition by 10.078 percentage points and 0.059 percentage points, respectively. A 10 percentage point increase in the proportion of persons aged 15 years or over with post-school qualifications and a $1,000 increase in the nominal gross annual median income of persons aged 15 years or over in a community increase the proportion of persons aged 18 years or over with self-assessed good or excellent health by 5.382 percentage points and 0.593 percentage points, respectively.

8. The regression estimates suggest that a 10 percentage point increase in the proportion of persons with no recent illness, no chronic condition and self-assessed good or excellent health (age 18 or over), and a $1,000 increase in the nominal gross annual median income of persons aged 15 years or over increase the proportion of persons aged 15 years or over with post-school qualifications by 0.190 percentage points, 0.979 percentage points, 0.444 percentage points, and 1.455 percentage points, respectively. The estimations also suggest that the proportion of persons aged 15 years or over with post-school qualifications in a community is positively related to past family income.

9. The regression estimates suggest that a 10 percentage point increase in the proportion of persons with no recent illness, no chronic condition, self-assessed good or excellent health (age 18 or over), and post-school qualifications (age 15 or over) in a community increases the nominal gross annual median income of persons aged 15 years or over by $922, $100, $1, and $919, respectively.

10. The regression estimates indicate that three indicators of good health and education are normal necessities; a 10 per cent increase in nominal gross annual median income of persons aged 15 years or over increases the proportions of persons with no recent illness, no chronic condition, self-assessed good or excellent health (age 18 or over), and post-school qualifications (age 15 or over) by 4.31 per cent, 0.28 per cent, 1.21 per cent, and 5.01 per cent, respectively. Given the estimations, the benefits of government expenditures on the three indicators of health and education are progressive.

11. The observed high values of multiple determination (R^2) in the second stage of TSLS estimates imply that the TSLS method works well. The econometric results from the TSLS estimates are consistent with those from the OLS estimates with two exceptions: the estimated education and income regression equations. These differences may be due to the impact of university dummy variable on the three indicators of good health. Furthermore, it has been observed from the reduced-form equations that universities affect the three indicators of good health indirectly, through education and through income caused by education. The coefficient stability analysis [Beggs (1988, p.97) and Dowrick (1993, p.2 and pp.26-27)] indicates that the beneficial health effects of universities raise the levels of education and income in communities in which they are located; that is, in a community with a university a 10 percentage point increase in the proportion of persons with no recent illness is associated with a 1.640 percentage points increase in the proportion of persons aged 15 years or over with post-school qualifications, ceteris paribus. Other things being

equal, in a community with a university a 10 percentage point increase in the proportions of persons with no chronic condition and self-assessed good or excellent health (age 18 or over) are associated with \$159 and \$322 increases in nominal gross annual median income of persons aged 15 years or over, respectively. Given the regression results, increases in the number of university places may have positive effects on health, education, and income.

12. The estimates suggest that the higher the proportion of the population which is female in a community, ceteris paribus, the more likely it is that its population will be healthy. In addition, universities yield beneficial health effects to communities in which they are located, due to beneficial education effects. Given that communities with a university campus have many service industries and that the service sector is female-intensive, women may be attracted to communities with a significant service sector. Thus, the proportion of the population which is female is higher in communities with a university than in communities without a university. The combined effects of a university on health and gender suggest that a higher proportions of persons with a set of three indicators of good health are associated with a higher proportion of the population which is female.

13. The independence and exogeneity tests indicate that education enhances earnings via the production of marketable skills - a fundamental way of increasing productivity [Lewis (1991, p.1)]; that is, an increase in the proportion of persons aged 15 years or over with post-school qualifications in a community increases the proportion of skilled workers to all employees aged 15 years or over, which in turn enhances nominal gross annual median income of persons aged 15 years or over. The estimations also suggest that an increase in nominal gross annual median income of persons aged 15 years or over in a community increases the proportion of skilled workers to all employees aged 15 years or over, due to an increase in the proportion of persons aged 15 years or over with post-school qualifications. On the other hand, estimates of the reduced-form skill level parameter suggest that communities with better health due to more education and higher income have a higher proportion of persons with skill in the sense that skilled workers are more attracted to communities where better health, more education, and higher income are expected.

14. The estimates suggest that maderate drinking yields beneficial physical and psychological effects which have beneficial health effects. A 10 percentage point increase in the proportion of persons aged 18 years or over with moderate alcohol consumption in a community increases the proportion of persons with no recent illness, no chronic conditions, and self-assessed good or excellent health (age 18 or over) by 4.625 percentage points, 1.643 percentage points, and 2.478 percentage points, respectively. These results indicate that communities with relatively higher proportions of persons aged 18 years or over with moderate alcohol consumption have better health, on average, in terms of no recent illness and no chronic condition for children as well as for adults.

15. On the other hand, estimates of the reduced-form three health indicators parameters of no drinking and moderate drinking suggest that excessive drinking has total (direct and indirect) detrimental effects on a set of three indicators of good health. The estimates imply that communities with a relatively lower proportion of persons aged 18 years or over with excessive alcohol consumption produce better health for children as well as for adults.

16. Given the positive effect of public health expenditure on health [Leu (1986, p.59), Hitiris and Posnett (1992, p.179) and Anand and Ravallion (1993, p.141)], the estimates suggest that with two exceptions (Queensland and Australian Capital Territory) the role of state government intervention in the health sector in terms of no recent illness is greater in New South Wales than in all other states and territories.. Its role in the health sector in terms of no chronic condition is greater in New South

Wales than in Victoria, South Australia and Tasmania, while its role is smaller in New South Wales than in Queensland, Western Australia, Northern Territory and Australian Capital Territory. With one exception (Australian Capital Territory) the role of state government intervention in the health sector in terms of self-assessed good health is greater in New South Wales than in all other states and territories. Note that Commonwealth funding data by statistical regions or by states and territories are not available in the Australian Bureau of Statistics (ABS) Information Consultancy Service for in-depth data investigation. Thus, Commonwealth outlays on health are excluded from the measurements.

17. The estimates suggest that the effect of non-government schools on education is positive [Williams and Carpenter (1990, p.15)].

18. The estimates suggest that communities with a university have a higher proportion of persons aged 15 years or over with post-school qualifications than communities without a university.

19. Estimates of the reduced-form education and income parameters of moderate drinking suggest that a 10 percentage point increase in the proportion of persons aged 18 years or over with moderate alcohol consumption in a community increases the proportion of persons aged 15 years or over with post-school qualifications and nominal gross annual median income of persons aged 15 years or over by 2.797 percentage points and $400, respectively; this results from an increase in the proportion of persons with no recent illness, no chronic condition, and self-assessed good or excellent health (age 18 or over). These results are consistent with the health economics literature that the beneficial health effects of moderate drinking carry over to the labour market, raising productivity and income [Berger and Leigh (1988, p.1346), and Hamilton and Hamilton (1993, p.1)].

20. On the other hand, estimates of the reduced-form education and income parameters of no drinking and moderate drinking suggest that the negative indirect effects of excessive drinking on education and income via a set of three indicators of good health offset the positive direct effects of excessive drinking on education and income. It turns out that communities with a relatively lower proportion of persons aged 18 years or over with excessive alcohol consumption produce better health for children as well as for adults, which in turn generate higher levels of education and income.

21. The estimates suggest that a 10 percentage point increase in the proportion of persons aged 15 years or over without union membership in a community increases nominal gross annual median income of persons aged 15 years or over by $2,131, while a 10 percentage point increase in the proportion of persons aged 15 years or over with union membership increases the nominal gross annual median income of persons aged 15 years or over in a community by $1,273. This suggests that non-union members receive a wage earnings premium over union members in the community. Given the Brown-Medoff hypothesis (1978, p.364) that unionisation has a substantial positive effect on wage, these results indicate that non-union members are over represented in industries with higher earnings.

22. The estimates suggest that a lower proportion of part-time workers to full-time workers in a community is associated with a higher level of nominal gross annual median income of persons aged 15 years or over; a 10 percentage point decrease in the proportion of part-time workers to full-time workers in a community increases nominal gross annual median income by $99.4.

23. The estimates suggest that a 10 percentage point increase in the proportion of persons aged 15 years or over who are employed in the finance and business services increases nominal gross annual median income of persons aged 15 years or over by $193.6.

6.1.2 Lifestyle Factors - Smoking and Drinking Behaviours

This econometric study has provided empirical support to the proposition that each of three good health indicators, education, and income are positively inter-related; that is, a higher proportion of persons with a set of three indicators of good health, a higher proportion of persons aged 15 years or over with post-school qualifications, and a higher level of nominal gross annual median income of persons aged 15 years or over in a community are inter-related.

In addition, different patterns of behaviour contribute to differences in health. For instance, moderate drinking yields beneficial physical and psychological effects which have beneficial health effects [Hamilton and Hamilton (1993, p.1)]. The levels of education and income affect moderate drinking directly and indirectly via good health - an important determinant of the drinking behaviour, implying that two causal relationships potentially exist: moderate drinking affecting health, and health affecting moderate drinking.

However, these beneficial health effects deteriorate as alcohol use increases, together with a negative effect of smoking on health [National Health Strategy of Australia, Research Paper No.1 (1992, p.139), and Conway, Pinyopusarerk, Carter, Penm and Stevenson (1993, p.15)]. Addicts with lower rates of discount respond more to changes in harmful expected consequences of addictive goods, such as a negative effect on health of smoking and excessive drinking [Becker, Grossman, and Murphy (1991, p.239)]. It follows that a higher expected cost (due perhaps to greater information about health hazards) is likely to reduce the demand for both cigarettes and excessive alcohol use, particularly among the rich. Since individuals with more education and higher income have lower rates of discount, communities with a higher proportion of persons with post-school qualifications and a higher level of income have lower rates of discount. This suggests that a community would be better off if it substantially reduced tobacco smoking by reducing the proportion of the population who smoke on a regular basis and if it minimised the harm associated with alcohol abuse [Conway, Pinyopusarerk, Carter, Penm, and Stevenson (1993, pp.39-41)]. For example, the campaign against smoking and excessive drinking appears to be potentially beneficial to health in the community [Collins and Lapsley (1993, p.1)]. Such campaigns are more likely to be successful in communities made up of individuals with low discount rates. Smoking and excessive drinking complement each other; reducing one is likely to result in reductions in the other.

We explore the multiple positive inter-relationships between health status and moderate drinking, and between smoking and excessive drinking. We also test the influence of other explanatory variables on moderate drinking, excessive drinking, and smoking.

The overall results and analysis of the estimated models allow for the following summary remarks to be made regarding the propositions tested:

1. The linear model is preferred to the double natural logarithmic model with respect to all the equations for a set of two indicators of good health, moderate drinking, smoking, and excessive drinking. Only the linear models of the five dependent variables are used for the application of diagnostic tests.

2. Correct specifications are implied in the estimated OLS regressions for a set of two indicators of good health, moderate drinking, smoking, and excessive drinking.

3. Heteroskedasticity could not be detected in the estimated OLS regressions for a set of two indicators of good health, moderate drinking, smoking, and excessive drinking.

4. The standard errors of both the OLS and TSLS estimates for the five regressions are small. In a comparison of the OLS and TSLS estimates, the former is smaller than the latter with one exception (i.e., the moderate drinking regression equation).

5. Other things being equal, a set of the two indicators of good health and moderate drinking are positively inter-related. The results indicate that the higher proportion of persons aged 18 years or over with moderate alcohol consumption in a community, the better health will be in the community on average in terms of no recent illness for children as well as for adults.

6. The estimates suggest that a set of two indicators of good health, education, and income are positively associated with moderate drinking. A 10 percentage point increase in the proportion of persons with either no recent illness or self-assessed good or excellent health (age 18 or over), increases the proportion of persons aged 18 years or over with moderate alcohol consumption by 4.957 percentage points and 5.831 percentage points, respectively. A 10 percentage point increase in the proportion of persons aged 15 years or over with post-school qualifications and a $1,000 increase in the nominal gross annual median income of persons aged 15 years or over in a community increase the proportion of persons aged 18 years or over with moderate alcohol consumption by 2.925 percentage points and 0.065 percentage points, respectively.

7. The estimates suggest that the proportion of persons with no chronic condition in a community has a negative effect on the proportion of persons aged 18 years or over with moderate alcohol consumption.

8. It is suggested from the estimates that a 10 percentage point increase in the proportion of persons aged 15 years or over with post-school qualifications and a $1,000 increase in the nominal gross annual median income of persons aged 15 years or over in a community increase the proportion of persons aged 15 years or over with moderate alcohol consumption by 2.521 percentage points and 7.027 percentage points, respectively, due to an increase in the proportion of persons with either no recent illness or self-assessed good or excellent health (age 18 or over). The estimations also

suggest that the well-educated are aware of the evidence that moderate drinking generates health improvements.

9. Estimates of the reduced-form health parameters suggest that smoking and excessive drinking have negative effects on a set of two indicators of good health directly and indirectly, via moderate drinking; a 10 percentage point reduction in the proportion of persons aged 18 years or over with either cigarettes consumption or excessive alcohol consumption increases the proportion of persons with no recent illness by 0.942 percentage points and 5.004 percentage points, respectively. The lower the numbers of smokers and excessive drinkers in a community, the better health will be on average in terms of no recent illness for children as well as for adults. The reduced-form estimates also suggest that a 10 percentage point reduction in the proportion of persons aged 18 years or over with either cigarettes consumption or excessive alcohol consumption in a community increases the proportion of persons aged 18 years or over with self-assessed good or excellent health by 4.148 percentage points and 0.246 percentage points, respectively. Hence, we suggest that smoking and excessive drinking have negative effects on moderate drinking.

10. Other things being equal, excessive drinking and smoking are positively inter-related. The estimates suggest that a 10 percentage point reduction in the proportion of persons aged 18 years or over with excessive alcohol consumption in a community reduces the proportion of persons aged 18 years or over with cigarettes consumption by 6.426 percentage points. Conversely, it also suggests that a 10 percentage point reduction in the proportion of persons aged 18 years or over with cigarettes consumption in a community reduces the proportion of persons aged 18 years or over with excessive alcohol consumption by 2.573 percentage points.

11. The estimates suggest that smoking and excessive drinking rise with the levels of education and income; a 10 percentage point increase in the proportion of persons aged 15 years or over with post-school qualifications and a $1,000 increase in the nominal gross annual median income of persons aged 15 years or over in a community increase the proportion of persons aged 18 years or over with cigarettes consumption by 0.140 percentage points and 0.613 percentage points respectively. A 10 percentage point increase in the proportion of persons aged 15 years or over with post-school qualifications and a $1,000 increase in the nominal gross annual median income of persons aged 15 years or over in a community increase the proportion of persons aged 18 years or over with excessive alcohol consumption by 0.623 percentage points and 0.433 percentage points, respectively.

12. The estimates suggest that a higher proportion of females aged between 25 and 44 years to total females in a community is associated with a lower proportion of persons aged 18 years or over with moderate alcohol consumption. This suggests that females in this age category are likely to be non-drinkers. The OLS estimates suggest that a higher proportion of persons aged 18 years or over without alcohol consumption in a community is associated with a higher proportion of persons with good health in terms of no recent illness.

13. Given the inclusion of moderate drinking in the health functions [Berger and Leigh (1988, p.1346)], the estimates suggest that with two exceptions (Queensland and Australian Capital Territory) the impact of state government intervention in the health sector in terms of moderate drinking is greater in New South Wales than in Victoria, South Australia, Western Australia, Tasmania, and Northern Territory.

14. The estimates suggest that a higher proportion of persons aged between 25 and 74 years in a community is associated with a higher proportion of persons aged 18 years or over with excessive alcohol consumption. On the other hand, the estimates suggest that each of higher proportions of persons aged between 15 and 24 years, between 25 and 44 years, and between 45 and 64 years is associated with a lower proportion of persons aged 18 years or over with moderate alcohol consumption. Given the estimations, it is possible that, on average, a higher proportion of persons aged between 18 and 44 years in a community is associated with a higher proportion of persons aged 18 years or over with excessive alcohol consumption, while a higher proportion of persons aged between 45 and 64 years is associated with a higher proportion of persons aged 18 years or over without alcohol consumption.

15. The estimates suggest that moderate drinking is negatively associated with both smoking and excessive drinking. Given a negative effect of no chronic condition on moderate drinking the estimates imply that communities with a higher proportion of persons with no chronic condition have a higher proportion of persons aged 18 years or over with excessive alcohol consumption. The implication also is that moderate drinking is a substitute for smoking.

16. The estimates suggest that a 10 percentage point increase in the proportion of persons aged 15 years or over who are employed in the finance, property, and business services industry reduces the proportion of persons aged 18 years or over with cigarettes consumption by 5.182 percentage point.

17. In the smoking equation, a positive sign on the coefficient of the constant term suggests that smoking is addictive. In the excessive drinking equation, a negative sign on the coefficient of the constant term is ambiguous in explaining that excessive drinking is addictive. It is consistent with the notion that females are not frequently excessive drinkers and therefore excessive drinking is more likely to be found among male drinkers.

6.2 POLICY IMPLICATIONS

From the independence and endogeneity tests, a set of three indicators of good health, education, and income are found to be positively inter-related as proposed by Lewis. It is also observed that the positive correlations among education, income, and moderate drinking develop only after the hypothesis that moderate drinking yields beneficial health effects.

In an attempt to further examine the beneficial health effects of moderate drinking, this study concentrates on testing for the null hypothesis of no endogeneity between a set of two indicators of good health and moderate drinking, and rejects the hypothesis. However, it is also evident that these beneficial health effects deteriorate as alcohol use increases, together with a negative effect of smoking on health.

An implication of these results is that there are multiple inter-relationships among health status, education, income, and moderate drinking. Thus, an increase in each of these variables are associated with increases in the other three, ceteris paribus. This reveals that health status, education and moderate drinking complement each other and therefore can be viewed as normal goods. The empirical results indicate that differences in both education and income are causal to either smoking or excessive drinking differentials. This suggests that smoking and excessive drinking complement each other and therefore can also be viewed as normal goods.

The most important empirical results for policy purposes are briefly summarised in Tables 6.1 to 6.3.[68]

---

[68] No recent illness (H_1) is measured as the proportion of persons without medical conditions such as illness, injury or disability experienced in the two weeks prior to interview. No chronic condition (H_2) is measured as the proportion of persons without medical conditions such as illness, injury or disability which have lasted or are expected to last for a period of six months or more including long-term and permanent impairment or disabilities. Self-assessed good or excellent health (H_3) is measured as the proportion of persons aged 18 years and over with self-assessed good or excellent health status as opposed to poor or fair status at the survey interview. Education (E) is measured as the proportion of persons aged 15 years or over who obtained a qualification such as bachelor degree or higher, trade/apprenticeship, certificate/diploma and other since leaving school. Income (Y) is nominal gross annual median income of persons aged 15 years or over used as an alternative to the mean for a

Table 6.1. The Ceteris Paribus Direct, Indirect, and Total Proportionate Rates of Changes of Three Indicators of Health with Respect to Education and Income[1]

A 10 per cent increase in:	Percentage Change in: No recent illness			No chronic condition			Self-assessed good or excellent health		
	Direct	Indirect	Total	Direct	Indirect	Total	Direct	Indirect	Total
Education	1.719	2.421	4.140	13.920	0.155	14.075	3.206	1.027	4.233
Income[1]	4.315	1.051	5.366	0.280	8.937	9.217	1.215	1.832	3.047

Note: **1.** Obtained at mean value. For example, the ceteris paribus total proportionate rate of change of no recent illness with respect to income at mean is 0.5366.

Table 6.2. The Ceteris Paribus Direct, Indirect, and Total Proportionate Rates of Changes of Education and Income with Respect to Three Indicators of Health, Education, and Income[1]

A 10 per cent increase in:	Percentage Change in: Education			Income		
	Direct	Indirect	Total	Direct	Indirect	Total
No recent illness	0.118	1.015	1.133	1.661	0.067	1.728
No chronic condition	0.709	0.135	0.844	0.210	0.392	0.602
Self-assessed good or excellent health	0.745	0.003	0.748	0.005	0.630	0.635
Education				2.671	0.580	3.251
Income	5.006	0.163	5.169			

Notes: **1.** Obtained at mean value. For example, the ceteris paribus total proportionate rate of change of education with respect to income at mean is 0.5169.

2. Obtained from the ceteris paribus income elasticity of education at mean value. For example, the ceteris paribus (direct) mean income elasticity of education is 0.5010.

---

frequency distribution with open-ended class interval (hence, $60,001 or more). The measured median income is before taxes and government transfers.

Moderate drinking (R_3) is measured as the proportion of persons aged 18 years or over with the average daily consumption of alcohol greater than zero but less than 50 millilitres or 35 grams for males and 25 millilitres or 17.5 grams for females on a regular basis over the seven days prior to interview. Therefore, excessive drinking (R_4) is measured as the proportion of persons aged 18 years or over with the average daily consumption of alcohol greater than 50 millilitres or 35 grams for males and 25 millilitres or 17.5 grams for females. Smoking (R_1) is measured as the proportion of persons aged 18 years or over with the average daily consumption of one or more cigarettes (or pipes or cigars) at the time of interview of 1989-90.

Table 6.3. The Ceteris Paribus Total Proportionate Rates of Changes of Three Indicators of Health, Education, and Income with Respect to Lifestyle Factors[1]

	Percentage Change in:				
A 10 per cent increase in:	No recent illness	No chronic condition	Self-assessed good or excellent health	Education	Income
Moderate Drinking	8.301	8.884	1.198	3.045	1.266
Excessive Drinking	-1.932	-2.735	-0.035	-0.386	-0.125
Smoking	-0.923	-3.201	-1.500	-0.217	-0.254

Note: **1.** Obtained at mean value. For example, the ceteris paribus total proportionate rate of change of no chronic condition with respect to smoking at mean, $\eta_{R_1 H_2}$, is -0.3201; $\eta_{R_1 H_2}$ = $\xi_{R_1 R_4} * \eta_{R_4 H_2}$ where $\xi_{R_1 R_4}$ stands for the ceteris paribus direct proportionate rate of change of excessive drinking with respect to smoking at mean.

Tables 6.1 and 6.2 indicate that more education is proportionately associated with better health directly and indirectly, through higher income. For example, a 10 per cent increase in the proportion of persons aged 15 years or over with post-school qualifications increases the proportion of persons with no recent illness by 4.140 per cent; by 1.719 per cent directly and by 2.421 per cent indirectly, due to increases in nominal gross annual median income of persons aged 15 years or over directly and indirectly via the proportion of persons with the other two indicators of good health. Thus, increases in education and/or income can make a substantial contribution to health. Consequently when exploring ways of improving health policy makers should not limit themselves to thinking about increasing health care expenditure or making the health care system more efficient; improving general education levels may be as effective.

Improving the level of health is likely to increase income levels directly and indirectly, through increases in education. For example, a 10 per cent increase in the proportion of persons with no recent illness increases nominal gross annual median income of persons aged 15 years or over by 1.728 per cent; by 1.661 per cent directly and by 0.067 per cent indirectly, due to increases in the proportion of persons aged 15 years or over with post-school qualifications and the proportion of persons with the other two indicators of good health via the proportion of persons aged 15 years or over with post-school qualifications. In this sense expenditure on health care can be (at least partially) self-financing. (Increased government spending on health care results in improved health which results in higher income and more government revenue). There will also be additional benefits in the form of better education.

In Table 6.3, excessive drinking and smoking affect health (as is generally known and appreciated) but they also affect education and income. For example, a 10 per cent reduction in the proportion of persons aged 18 years or over with excessive alcohol consumption increases the proportion of persons aged 18 years or over with self-assessed good or excellent health, the proportion of persons aged 15 years or over with post-school qualifications, and nominal gross annual median income of persons aged 15 years or over by 0.035 per cent, 0.386 per cent, and 0.125 per cent, respectively. Therefore, campaigns to affect behaviour (e.g. anti binge drinking and anti smoking campaigns) will not only improve health (as primarily intended) but also income (and hence taxes and may be partially self-funding) and education.

On the basis of this study, the choice of policy instruments should not be based upon stability analysis for the single final target variable. Target variables can be influenced in a variety of ways.

1. The proportion of persons with a set of three indicators of good health in a community may be increased directly by an increased proportion of persons aged 18 years or over with moderate alcohol consumption and an increased expenditure on medical care. It may also be increased indirectly through an increased proportion of persons aged 15 years or over with post-school qualifications, and through healthy environment, better diet and access to the better medical care as a result of higher income.

2. The proportion of persons aged 15 years or over with post-school qualifications in a community may be increased directly by additional years of schooling such as university education. It may also be increased indirectly through an increased proportion of persons with a set of three indicators of good health and through higher past family income as well as current income. Furthermore, the proportion of persons aged 15 years or over with post-school qualifications may be increased indirectly by increased proportion of persons aged 18 years or over with moderate alcohol consumption, due to increased proportion of persons with a set of three indicators of good health.

3. Income in a community may be increased directly by an increased proportion of persons aged 15 years or over who are employed in the finance, property and business services industry, and a decreased proportion of part-time workers to full-time workers. It may also be increased indirectly through higher productivity in the community due to an increased proportion of persons with a set of three indicators of good health and through higher proportion of persons with the skills and knowledge as a result of more education. Furthermore, income may be increased indirectly by increased proportion of persons aged 18 years or over with moderate alcohol consumption, due to increased proportion of persons with a set of three indicators of good health.

4. A set of three indicators of good health and education are normal necessities. Given the empirical observations, the benefits from government expenditures on three indicators of health and education are progressive.

5. There is evidence of coefficient instability in education and income caused by the university dummies for the three indicators of health. Furthermore, it has been observed that universities affect a set of three indicators of good health indirectly, through education and income. The coefficient stability analysis indicates that universities have positive effects on education and income to communities. The implication is that the number of university places has a positive effect on health, education, and income.

6. The higher the proportion of the population which is female in a community, ceteris paribus, the more likely it is that its population will be healthy. In addition, universities yield beneficial health effects to communities in which they are located, due to beneficial education effects. Given that communities with a university campus have many service industries and that the service sector is female-intensive, women tend to move from a community without the service sector to a community with it. Thus, the proportion of the population which is female is higher in communities with a university than in communities without a university. The combined effects of a university on health and gender suggest that higher proportions of persons with a set of three indicators of good health in a community are associated with a higher proportion of the population which is female.

7. Community income may be increased by an increased proportion of skilled workers to all employees aged 15 years or over as a result of more education. On the other hand, communities with higher proportions of persons with a set of three indicators of good health due to a higher proportion of persons aged 15 years or over with post-school qualifications and a higher level of income tend to have a higher proportion of persons with skill in the sense that skilled workers are more attracted to communities where better health, more education, and higher income are expected.

8. The proportion of persons aged 18 years or over with moderate alcohol consumption in a community may be increased directly by an increased proportion of persons with no recent illness and self-assessed good or excellent health (age 18 or over), and indirectly by increased proportion of persons aged 15 years or over with post-school qualifications and increased level of income.

9. Given that education and income are positively associated with moderate drinking through an effect on better health, communities with a high propensity to invest in education engage in forward-looking habits with respect to health, which in turn mean

that those with a higher level of income and a higher proportion of persons aged 18 years or over with moderate alcohol consumption are future-oriented. Furthermore, the well-educated are more aware of the evidence that moderate drinking generates health improvements. From the inter-relationships among health status, education, and income, it also expected that the beneficial health effects of moderate drinking lead to productivity gains, which in turn raise the levels of education and income.

10. A higher proportion of persons with no chronic condition in a community is associated with a lower proportion of persons aged 18 years or over with moderate alcohol consumption. This suggests that communities with higher proportions of persons with no chronic condition have higher proportions of persons aged 18 years or over with excessive alcohol consumption.

11. The proportion of persons with positive health in a community may be increased by a reduction in the number of smokers or excessive drinkers. The proportion of persons aged 18 years or over with cigarettes consumption in a community may be reduced by decreased proportion of persons aged 18 years or over with excessive alcohol consumption and vice versa. Further, a higher proportion of persons aged 15 years or over with post-school qualifications and a higher level of income in a community are associated with a higher proportion of persons aged 18 years or over with either cigarettes consumption or excessive alcohol consumption. Therefore, given the importance of the population's health, communities will be better off if they substantially reduce tobacco smoking by reducing the proportion of the population who smoke on a regular basis and if they reduce the harm associated with alcohol abuse by increasing the proportion of persons aged 18 years or over with moderate alcohol consumption. Campaigns against smoking and excessive drinking are likely to yield substantial benefits to the community.

7.0 REFERENCES

Acs, G. and Danziger, S., "Educational Attainment, Industrial Structure, and Male Earnings Through the 1980s", Journal of Human Resources, 28(3), Summer, 1993, 618-648.

Anand, S. and Ravallion, M., "Human Development in Poor Countries: On the Role of Private Incomes and Public Services", Journal of Economic Perspectives, 7(1), Winter, 1993, 133-150.

Ashenfelter, O., "How Convincing is the Evidence Linking Education and Income?", prepared for the 22nd Conference of Economists, Economic Society of Australia, Murdoch University, Western Australia, September, 1993.

Auster, R., Leveson, I. and Sarachek, D., "The Production of Health, An Exploratory Study", Journal of Human Resources, 4(4), Fall, 1969, 411-436.

Australian Bureau of Statistics, 1989-90 National Health Survey: Summary of Results Australia, ABS Catalogue No.4364.0, Canberra, 1991.

Australian Bureau of Statistics, 1989-90 National Health Survey: User's Guide, ABS Catalogue No.4363.0, Canberra, 1991.

Australian Bureau of Statistics, 1991 Census Geographic Areas, ABS Catalogue No.2905.0, Canberra, 1992.

Australian Bureau of Statistics, 1991 Census of Population and Housing, ABS Catalogue No.2722.0, Canberra, 1992.

Australian Bureau of Statistics, Australian Standard Classification of Occupations (ASCO) Manual, ABS Catalogue No.1222.0, Canberra, 1991.

Australian Bureau of Statistics, Australian Standard Geographical Classification (ASGC) Manual, ABS Catalogue No.1216.0, Canberra, 1991.

Australian Bureau of Statistics, Australian Standard Industrial Classification (ASIC) Manual, ABS Catalogue No.1201.0, Canberra, 1991.

Australian Bureau of Statistics, Australian National Accounts, ABS Catalogue No.5221.0, Canberra, 1992.

Australian Bureau of Statistics, Building Activity Australia, ABS Catalogue No.8752.0, Canberra, 1992.

Australian Bureau of Statistics, Government Finance Statistics, ABS Catalogue No.5512.0, Canberra, 1992.

Australian Bureau of Statistics, Household Expenditure Survey States and Territories, ABS Catalogue No.6533.0, Canberra, 1990.

Australian Bureau of Statistics, Measuring Australia's Economy: A Student Guide, ABS Catalogue No.1360.0, Canberra, 1993.

Australian Bureau of Statistics, Monthly Summary of Statistics, ABS Catalogue No.1304.0, Canberra, 1993.

Australian Bureau of Statistics, Schools Australia, ABS Catalogue No.4221.0, Canberra, 1992.

Australian Bureau of Statistics, Trade Union Members Australia, ABS Catalogue No.6325.0, Canberra, 1992.

Australian Council for Educational Research (ACER), Australian Education Directory, edited by J. Badger and P. Mathews, ACER, Hawthorn, Victoria, 1992.

Australian Financial Review, Universities in Australia, edited by M. Dwyer and S. Lewis, The Financial Review Library, 1992.

Baumol, W.J., Economic Theory and Operations Analysis, Fourth Edition, Prentice-Hall, Inc., New Jersey, 1977.

Becker, G.S. and Murphy, K.M., "A Theory of Rational Addiction", Journal of Political Economy, 96(4), August, 1988, 675-700.

Becker, G.S., Grossman, M. and Murphy, K.M., "Rational Addiction and the Effect of Price on Consumption", American Economic Review, 81(2), May, 1991, 237-241.

Beggs, J.J., "Diagnostic Testing in Applied Econometrics", Economic Record, 64(185), June, 1988, 81-101.

Berger, M.C. and Leigh, J.P., "The Effect of Alcohol Use on Wages", Applied Economics, 20(10), October, 1988, 1343-1351.

Björklund, A., "Unemployment and Mental Health: Some Evidence from Panel Data", Journal of Human Resources, 20(4), Fall, 1985, 469-483.

Blandy, R. and Goldsworthy, T., "Private Returns to Education in South Australia" in Australian Labour Economics: Readings, edited by B.J. Chapman, J.E. Isaac and J.R. Niland, Third Edition, Macmillan Company of Australia Pty. Limited, Melbourne, 1984.

Blaug, M., "Analogies with Health Economics" in Economics of Education, Penguin Books, London, 1970.

Boissiere, M., Knight, J.B., and Sabot, R.H., "Earnings, Schooling, Ability, and Cognitive Skills", American Economic Review, 75(5), December, 1985, 1016-1030.

Breusch, T.S. and Pagan, A.R., "A Simple Test for Heteroscedasticity and Random Coefficient Variation", Econometrica, 47(5), September, 1979, 1287-1294.

Brown, C. and Medoff, J., "Trade Unions in the Production Process", Journal of Political Economy, 86(3), June, 1978, 355-378.

Carpenter, P.G. and Hayden, M., "Academic Achievement among Australian Youth", Australian Journal of Education, 29(3), November, 1985, 199-220.

Chirikos, T.N. and Nestel, G., "Further Evidence on the Economic Effects of Poor Health", Review of Economics and Statistics, 67(1), February, 1985, 61-69.

Collins, D. and Lapsley, H., "Some Research Directions for Policy Formulation in the Economics of Drug Abuse", presented at the International Symposium on the Economics of Drug Abuse, Canberra, April, 1993.

Conway, L.M., Pinyopusarerk, M., Carter, R.C., Penm, R., and Stevenson, C., "The Public Health Significance and Cost of Illness Resulting from Drug and Alcohol Abuse in Australia", presented at the International Symposium on the Economics of Drug Abuse, Canberra, April, 1993.

Cook, P.J. and Moore, M.J., "Drinking and Schooling", Journal of Health Economics, 12(4), December, 1993, 411-429.

Culyer, A.J., "The Demand for Health Care", in Need and the National Health Service, Martin Robinson, Oxford, 1976.

Dardanoni, V. and Wagstaff, A., "Uncertainty, Inequalities in Health and the Demand for Health", Journal of Health Economics, 6(4), December, 1987, 283-290.

Deeble, J., Medical Services Through Medicare, Background Paper No.2 National Health Strategy, Australia, February, 1991.

Dowrick, S., "Estimating the Impact of Government Consumption on Growth: Growth Accounting and Optimising Models", presented at the Combined Economics Department and Associated Research Programs Seminar/Workshop series, University of Wollongong, New South Wales, Australia, August, 1993.

Ehrenberg, R.G. and Smith, R.S., Modern Labor Economics: Theory and Public Policy, Second Edition, Scott, Foresman and Company, Illinois, 1985.

Folland, S., Goodman, A.C. and Stano, M., The Economics of Health and Health Care, Macmillan Publishing Company, New York, 1993.

Fuchs, V.R., Who Shall Live? Health, Economics, and Social Choice, Basic Books Inc., New York, 1974.

Fuchs, V.R., The Health Economy, Harvard University Press, Cambridge, Massachusetts, 1986.

Gerdtham, U-G. and Jönsson, B., "Conversion Factor Instability in International Comparisons of Health Care Expenditure", Journal of Health Economics, 10(2), July, 1991, 227-234.

Gerdtham, U-G. and Jönsson, B., "International Comparisons of Health Care Expenditure-Conversion Factor Instability, Heteroscedasticity, Outliers and Robust Estimators", Journal of Health Economics, 11(2), August, 1992, 189-197.

Giannaros, D.S., "Long-term Real Interest Rate Forecasting and the Role of Uncertainty and Deficits", Initial paper presented at Fifth International Symposium on Forecasting, Montreal, Canada, June, 1985.

Glover, J. and Woollacott, T., A Social Health Atlas of Australia Volume 1, Australian Bureau of Statistics, ABS Catalogue No.4385.0, Canberra, July, 1992a.

Glover, J. and Woollacott, T., A Social Health Atlas of Australia Volume 2, Australian Bureau of Statistics, ABS Catalogue No.4385.0, Canberra, September, 1992b.

Graetz, B., "Private Schools and Educational Attainment: Cohort and Generational Effects", Australian Journal of Education, 34(2), August, 1990, 174-191.

Gregory's Sydney Street Directory, A Division of Universal Press Pty. Limited, Macquarie Park, New South Wales, 1991.

Griliches, Z., "Production Functions in Manufacturing: Some Preliminary Results" in The Theory and Empirical Analysis of Production, edited by M. Brown, National Bureau of Economic Research, New York, 1967.

Grossman, M., "On the Concept of Health Capital and the Demand for Health", Journal of Political Economy, 80(2), March/April, 1972, 223-255.

Grossman, M. and Benham, L., "Health, Hours, and Wages", in The Economics of Health and Medical Care, edited by M. Perlman, Macmillan Press Limited, London, 1974.

Hamilton, V. and Hamilton, B.H., "Alcohol in the Workplace: Examining the Impact of Moderate and Heavy Drinking on Workers' Earnings", presented at the International Symposium on the Economics of Drug Abuse, Canberra, April, 1993.

Hausman, J.A., "Specification Tests in Econometrics", Econometrica, 46(6), November, 1978, 1251-1271.

Hirsch, B.T. and Addison, J.T., The Economic Analysis of Unions: New Approaches and Evidence, Allen & Unwin Inc., Massachusetts, 1986.

Hitiris, T. and Posnett, J., "The Determinants and Effects of Health Expenditure in Developed Countries", Journal of Health Economics, 11(2), August, 1992, 173-181.

Hogan, W.P., "The Illicit Drug Market: Demand and Participation", presented at the International Symposium on the Economics of Drug Abuse, Canberra, April, 1993.

Johnston, J., Econometric Methods, Third Edition, McGraw-Hill Book Company, New York, 1984.

Kemna, H.J.M.I., "Working Conditions and the Relationship Between Schooling and Health", Journal of Health Economics, 6(3), September, 1987, 189-210.

Kenkel, D.S., "Health Behavior, Health Knowledge, and Schooling", Journal of Political Economy, 99(2), April, 1991, 287-305.

Kroch, E.A. and Sjoblom, K., "Schooling as Human Capital or a Signal: Some Evidence", Journal of Human Resources, 29(1), Winter, 1993, 156-180.

Land Information Center, The Official Road Directory of New South Wales, N.S.W. Department of Lands, Bathurst, New South Wales, 1990.

Layard, P.R.G. and Walters, A.A., Microeconomic Theory, McGraw-Hill Book Company, New York, 1978.

Leamer, E.E., Specification Searches, John Wiley & Sons, Inc., New York, 1978.

Leu, R.E., "The Public-Private Mix and International Health Care Costs", in Public and Private Health Services, edited by A.J. Culyer and B. Jönsson, Basil Blackwell Limited, Oxford, 1986.

Lewis, D.E., "Comparative Quit Rates of Men and Women", Journal of Industrial Relations, September, 1979, 331-350.

Lewis, D.E., "Credit Transfers Between TAFE and Higher Education", Australia Department of Employment, Education and Training, Canberra, July, 1991.

Lewis, D.E., Lecture Notes of the Economics of Health Care, University of Wollongong, New South Wales, 1994.

Lewis, D.E., O'Brien, D.T. and Thampapillai, D., Statistics for Business and Economics, Harcourt Brace Jovanovich Publishers, Sydney, 1990.

MacKinnon, J.G. and White, H., "Some Heteroscedastic Consistent Covariance Matrix Estimators with Improved Finite Sample Properties", Journal of Econometrics, 29(3), September, 1985, 305-325.

MacKinnon, J.G., White, H. and Davidson, R., "Tests for Model Specification in the Presence of Alternative Hypotheses", Journal of Econometrics, 21(1), January, 1985, 53-70.

Maddala, G.S., Econometrics, McGraw-Hill, Inc., New York, 1977.

Maddala, G.S., Introduction to Econometrics, Second Edition, Macmillan Publishing Company, New York, 1992.

Maglen, L.R., "Challenging the Human Capital Orthodoxy: The Education-Productivity Link Re-examined", Economic Record, 66(195), December, 1990, 281-294.

Mansfield, E., Statistics for Business and Economics, Second Edition, W.W. Norton & Company, New York, 1983.

McClelland, A., In Fair Health? Equity and the Health System, Background Paper No.3 National Health Strategy, Australia, March, 1991a.

McClelland, A., Spending on Health, Background Paper No.7 National Health Strategy, Australia, July, 1991b.

Miller, P.W., "The Rate of Return to Education: Evidence from the 1976 Census", in Australian Labour Economics: Readings, edited by B.J. Chapman, J.E. Isaac and J.R.Niland, Third Edition, Macmillan Company of Australia Pty. Limited, Melbourne, 1984.

Mullahy, J. and Sindelar, J.L., "Gender Differences in Labor Market Effects of Alcoholism", American Economic Review, 81(2), May, 1991, 161-165.

Muurinen, J.M., "Demand for Health: A Generalised Grossman Model", Journal of Health Economics, 1(1), April, 1982, 5-28.

Narendranthan, W., Nickell, S., and Metcalf, D., "An Investigation into the Incidence and Dynamic Structure of Sickness and Unemployment in Britain, 1965-75", Journal of the Royal Statistical Society, 148 (Part 3), 1985, 254-267.

National Health Strategy of Australia, Enough to Make You Sick: How Income and Environment Affect Health, Research Paper No.1, Melbourne, September, 1992.

Neter, J., Wasserman, W., and Whitmore, G.A., Applied Statistics, Second Edition, Allyn and Bacon, Inc., Boston, 1982.

Newhouse, J.P. and Friedlander, L.J., "The Relationship Between Medical Resources and Measures of Health: Some Additional Evidence", Journal of Human Resources, 15(2), Spring, 1980, 200-218.

Norris, K., The Economics of Australian Labour Markets, Second Edition, Longman Cheshire Pty. Limited, Melbourne, 1989.

Olsen, J.A., "On What Basis Should Health Be Discounted?", Journal of Health Economics, 12(1), April, 1993a, 39-53.

Olsen, J.A., "But Health Can Still Be a Necessity...", Journal of Health Economics, 12(2), July, 1993b, 187-191.

Panik, M.J., "Classical Optimization: Foundations and Extensions", in Studies in Mathematical and Managerial Economics, edited by H. Theil, North-Holland Publishing Company American Elsevier, Inc., New York, 1976.

Parkin, D., McGuire, A., and Yule, B., "Aggregate Health Care Expenditures and National Income: Is Health Care a Luxury Good?", Journal of Health Economics, 6(2), June, 1987, 109-127.

Podgursky, M., "Unions and Family Income Inequality", Journal of Human Resources, 18(4), Fall, 1983, 574-591.

Richardson, J., "Health Economics: An Overview", in The Economics of Health Care, edited by J.W. Neville, CEDA, 1988.

Schulz, T.P., "Investments in the Schooling and Health of Women and Men: Quantities and Returns", Journal of Human Resources, 28(4), Fall, 1993, 694-734.

Sindelar, J.L., "Differential Use of Medical Care by Sex", Journal of Political Economy, 90(5), October, 1982, 1003-1019.

South Australian Health Commission, A Social Health Atlas of South Australia, SAHC Adelaide, South Australia, 1991.

UBD Street Directory, A Division of Universal Press Pty. Limited, Australia, 1993.

Umberson, D., "Gender, Marital Status and the Social Control of Health Behavior", Social Science Medicine, 34(8), April, 1992, 907-917.

Van deVen, W.P.M.M.and Van der Gaag, J., "Health as an Unobservable a MIMIC-Model of Demand for Health Care", Journal of Health Economics, 1(2), August, 1982, 157-183.

Van Vliet, R.C.J.A. and Van Praag, B.M.S., "Health Status Estimation on the Basis of MIMIC-Health Care Models", Journal of Health Economics, 6(1), March, 1987, 27-42.

Wagstaff, A., "The Demand for Health: Some New Empirical Evidence", Journal of Health Economics, 5(3), September, 1986, 195-233.

Wagstaff, A., "Econometric Studies in Health Economics: A Survey of the British Literature", Journal of Health Economics, 8(1), May, 1989, 1-51.

White, H., "Using Least Squares to Approximate Unknown Regression Functions", International Economic Review, 21(1), February, 1980, 149-170.

Williams, T. and Carpenter, P.G., "Private Schooling and Public Achievement", Australian Journal of Education, 34(1), April, 1990, 3-24.

Wise, D.A., "Academic Achievement and Job Performance", American Economic Review, 65(3), June, 1975, 350-366.

Wolpin, K.I., "Education and Screening", American Economic Review, 67(5), December, 1977, 949-958.

Wooden, M., "Factors Associated with Workplace Accidents", Journal of Occupational Health and Safety - Australia and New Zealand, 6(2), 1990, 97-102.

APPENDICES

Appendix Table A-1. Statistical Regional Structure in Australia

State, Major Statistical Region and Statistical Region[1]	State, Major Statistical Region and Statistical Region[1]
NEW SOUTH WALES	**QUEENSLAND**
Sydney Statistical Region	Brisbane Statistical Region
Inner Sydney	Brisbane City Inner Core
Eastern Suburbs	Brisbane City Outer Core
St. George - Sutherland	South and East Brisbane SD Balance
Canterbury - Bankstown	North and West Brisbane SD Balance[2]
Fairfield - Liverpool	Balance of Queensland
Inner Western Sydney	South and East Moreton
Outer South Western Sydney	North and West Moreton
Central Western Sydney	Wide Bay - Burnett
Outer Western Sydney	Darling Downs - South West
Blacktown - Baulkham Hills	Mackay - Fitzroy - Central West
Lower Northern Sydney	Northern - North - West
Hornsby - Ku-ring-gai	Far North
Manly - Warringah	
Gosford - Wyong	**SOUTH AUSTRALIA**
Balance of NSW	Adelaide Statistical Region
Hunter	Northern Adelaide
Illawarra	Western Adelaide
Richmond - Tweed	Eastern Adelaide
Mid-North Coast	Southern Adelaide
Northern	Balance of South Australia
Far West - North Western	Northern and Western SA
Central West	Southern and Eastern SA
South - Eastern	
Murray - Murrumbidgee	**WESTERN AUSTRALIA**
	Perth Statistical Region
VICTORIA	Central Metropolitan
Melbourne Statistical Region	East Metropolitan
Inner Melbourne Region	North Metropolitan
Southern Melbourne Region	South West Metropolitan
Inner Eastern Melbourne Region	South East Metropolitan
North Eastern Melbourne Region	Balance of Western Australia
North Western Melbourne Region	Lower Western WA[3]
Western Melbourne Region	
Mornington Peninsula Region	**TASMANIA**
Outer Eastern Melbourne Region	
Balance of Victoria	**NORTHERN TERRITORY**
South Western Victoria	
Western Victoria	**AUSTRALIAN CAPITAL TERRITORY**
Northern Victoria	
Eastern Victoria	**AUSTRALIA**

1. Statistical regions are ordered as they appear in the data package of the 1989-90 National Health Survey (NHS) provided by the Australian Bureau of Statistics (ABS).
2. Consists of South-West Brisbane and North Brisbane Subdivision Balances.
3. Includes Remainder - Balance of WA.

Appendix Table A-2. Rank Criterion Matrix of the Structural Equations for Three Indicators of Health (No Recent Illness, H_1, No Chronic Condition, H_2, Self-assessed Good Health, H_3), Education (E), and Income (Y)

$$A^*_{H_1} = \begin{array}{c} \begin{array}{ccccccccccccc} OCC_2 & NGSTD & UD & AE & K & Ln & I & PFL & GNG_1 & GNG_2 & GNG_3 & GNG_4 & GNG_6 \end{array} \\ \left[\begin{array}{ccccccccccccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ d_8 & d_{10} & d_{11} & d_{12} & 0 & 0 & 0 & 0 & d_{13} & d_{14} & d_{15} & d_{16} & d_{17} \\ 0 & 0 & 0 & 0 & e_{10} & e_{11} & e_{12} & e_{13} & 0 & e_{14} & 0 & 0 & 0 \end{array}\right] \end{array}$$

$$A^*_{H_2} = \begin{array}{c} \begin{array}{ccccccccccccc} OCC_2 & NGSTD & UD & AE & K & Ln & I & PFL & GNG_1 & GNG_2 & GNG_3 & GNG_4 & GNG_6 \end{array} \\ \left[\begin{array}{ccccccccccccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ d_8 & d_{10} & d_{11} & d_{12} & 0 & 0 & 0 & 0 & d_{13} & d_{14} & d_{15} & d_{16} & d_{17} \\ 0 & 0 & 0 & 0 & e_{10} & e_{11} & e_{12} & e_{13} & 0 & e_{14} & 0 & 0 & 0 \end{array}\right] \end{array}$$

$$A^*_{H_3} = \begin{array}{c} \begin{array}{ccccccccccccccc} OCC_2 & B_1 & B & NGSTD & UD & AE & K & Ln & I & PFL & GNG_1 & GNG_2 & GNG_3 & GNG_4 & GNG_6 \end{array} \\ \left[\begin{array}{ccccccccccccccc} 0 & a_8 & a_9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & b_8 & b_9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ d_8 & 0 & 0 & d_{10} & d_{11} & d_{12} & 0 & 0 & 0 & 0 & d_{13} & d_{14} & d_{15} & d_{16} & d_{17} \\ 0 & 0 & 0 & 0 & 0 & 0 & e_{10} & e_{11} & e_{12} & e_{13} & 0 & e_{14} & 0 & 0 & 0 \end{array}\right] \end{array}$$

Appendix Table A-2. (Continued)

$$
A_E^* = \begin{array}{c}
\begin{array}{ccccccccccccccc}
F & B_1 & B & G^*VIC & G^*QLD & G^*WA & G^*TAS & G^*(SA+NT) & METD & K & Ln & I & PFL & GNG_8 & GNG_9
\end{array} \\
\left[\begin{array}{ccccccccccccccc}
a_5 & a_8 & a_9 & a_{10} & a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & 0 & 0 & 0 & 0 & a_{16} & a_{17} \\
b_5 & b_8 & b_9 & b_{10} & b_{11} & b_{12} & b_{13} & b_{14} & b_{15} & 0 & 0 & 0 & 0 & b_{16} & b_{17} \\
c_5 & 0 & 0 & c_8 & c_9 & c_{10} & c_{11} & c_{12} & c_{13} & 0 & 0 & 0 & 0 & c_{14} & c_{15} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & e_9 & e_{10} & e_{11} & e_{12} & e_{13} & 0 & e_{15}
\end{array}\right]
\end{array}
$$

$$
A_Y^* = \begin{array}{c}
\begin{array}{ccccccccccccccccccc}
F & OCC_2 & B_1 & B & G^*VIC & G^*QLD & G^*WA & G^*TAS & G^*(SA+NT) & NGSTD & UD & AE & GNG_1 & GNG_3 & GNG_4 & GNG_6 & GNG_8
\end{array} \\
\left[\begin{array}{ccccccccccccccccc}
a_5 & 0 & a_8 & a_9 & a_{10} & a_{11} & a_{12} & a_{13} & a_{14} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_{16} \\
b_5 & 0 & b_8 & b_9 & b_{10} & b_{11} & b_{12} & b_{13} & b_{14} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & b_{16} \\
c_5 & 0 & 0 & 0 & c_8 & c_9 & c_{10} & c_{11} & c_{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & c_{14} \\
0 & d_8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & d_{10} & d_{11} & d_{12} & d_{13} & d_{15} & d_{16} & d_{17} & 0
\end{array}\right]
\end{array}
$$

Notes:
1. A_X^* denotes a rank criterion matrix for the structural equation X; X = H_1, H_2, H_3, E, Y.
2. 0 denotes a variable not in the structural equation.
3. All the equations are identified with rows and columns whose determinants are non-zero. The pair of the structural equations for H_1, H_2, and H_3 has no simultaneous relationships. See Baumol (1977, pp.237-252).

Appendix Table A-3. Rank Criterion Matrix of the Structural Equations for No Recent Illness (H_1), Self-assessed Good Health (H_3), and Moderate Drinking (R_3)

$$A^*_{H_1} = \begin{array}{c} \begin{array}{ccccccccc} H_2 & R_1 & R_4 & AGE_3 & AGE_4 & AGE_5 & FAGE_3 & FAGE_4 & FAGE_5 \end{array} \\ \left[\begin{array}{ccccccccc} 0 & 0 & g_4 & 0 & 0 & 0 & 0 & 0 & 0 \\ h_5 & h_6 & 0 & h_7 & h_8 & h_9 & h_{10} & h_{11} & h_{12} \end{array}\right] \end{array}$$

$$A^*_{H_3} = \begin{array}{c} \begin{array}{ccccccccccc} R_2 & B_1 & B & H_2 & R_1 & AGE_3 & AGE_4 & AGE_5 & FAGE_3 & FAGE_4 & FAGE_5 \end{array} \\ \left[\begin{array}{ccccccccccc} f_4 & f_{13} & f_{14} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & h_5 & h_6 & h_7 & h_8 & h_9 & h_{10} & h_{11} & h_{12} \end{array}\right] \end{array}$$

$$A^*_{R_3} = \begin{array}{c} \begin{array}{ccccccccccc} R_2 & R_4 & F & OCC_1 & C & B_1 & B & METD & GNG_8 & GNG_9 & GNG_{10} \end{array} \\ \left[\begin{array}{ccccccccccc} f_4 & 0 & f_5 & f_6 & f_{12} & f_{13} & f_{14} & f_{15} & f_{16} & f_{17} & f_{18} \\ 0 & g_4 & g_5 & g_6 & g_{12} & 0 & 0 & g_{13} & g_{14} & g_{15} & g_{16} \end{array}\right] \end{array}$$

Notes:
1. A^*_X denotes a rank criterion matrix for the structural equation X; X = H_1, H_3, R_3.
2. 0 denotes a variable not in the structural equation.
3. All the equations are identified with rows and columns whose determinants are non-zero. See Baumol (1977, pp.247-252).

Appendix Table A-4. Rank Criterion Matrix of the Structural Equations for Smoking (R_1) and Excessive drinking (R_4)

$$A^*_{R_1} = \begin{array}{c} \begin{array}{ccccc} GNG_4 & GNG_5 & GNG_6 & GNG_7 & GNG_8 \end{array} \\ \begin{bmatrix} \ell_{10} & \ell_{11} & \ell_{12} & \ell_{13} & \ell_{14} \end{bmatrix} \end{array}$$

$$A^*_{R_4} = \begin{array}{c} \begin{array}{ccc} OCC_1 & GNG_9 & GNG_{11} \end{array} \\ \begin{bmatrix} k_8 & k_{11} & k_{12} \end{bmatrix} \end{array}$$

Notes: 1. A^*_X denotes a rank criterion matrix for the structural equations X; $X = R_1, R_4$.
2. Two equations are identified with columns whose determinants are non-zero. See Baumol (1977, pp.247-252) and Johnston (1985, pp.460-461).

Appendix Table A-5. Simple Bivariate Correlation Coefficients Between Each Pair of the State Dummies[1]

	SD_1	SD_2	SD_3	SD_4	SD_5	SD_6	SD_7	SD_8
SD_1	1.000							
SD_2	-0.385	1.000						
SD_3	-0.365	-0.232	1.000					
SD_4	-0.257	-0.163	-0.155	1.000				
SD_5	-0.257	-0.163	-0.155	-0.109	1.000			
SD_6	-0.100	-0.064	-0.061	-0.043	-0.043	1.000		
SD_7	-0.100	-0.064	-0.061	-0.043	-0.043	-0.017	1.000	
SD_8	-0.100	-0.064	-0.061	-0.043	-0.043	-0.017	-0.017	1.000

Notes: 1. $SD_1 = 1$ for statistical regions in New South Wales (23), $SD_2 = 1$ for statistical regions in Victoria (12), $SD_3 = 1$ for statistical regions in Queensland (11), $SD_4 = 1$ for statistical regions in South Australia (6), $SD_5 = 1$ for statistical regions in Western Australia (6), $SD_6 = 1$ for Tasmania (1), $SD_7 = 1$ for Northern Territory (1) and $SD_8 = 1$ for Australian Capital Territory (1) with the number of statistical regions in parentheses, and zero elsewhere.

Appendix Table A-6. Simple Bivariate Correlation Coefficients Between Each Pair of the State Variables Interacted with Public Health Expenditure per Person (age 15 Years or Over)[1]

	G*VIC	G*QLD	G*WA	G*TAS	G*(SA+NT)	G*(1+ACT)
G*VIC	1.000					
G*QLD	-0.270	1.000				
G*WA	-0.229	0.681	1.000			
G*TAS	-0.230	0.754	0.967	1.000		
G*(SA+NT)	-0.221	0.735	0.959	0.995	1.000	
G*(1+ACT)	-0.217	0.745	0.970	0.997	0.997	1.000

Notes: 1. G is defined as nominal state and local governments outlays on health per person by States and Territories (age 15 years or over). New South Wales is treated as a base. Then other States variables are measured as: VIC = $1+SD_2$; QLD = $1+SD_3$; WA = $SD_5 + SD_4 + 2SD_3 + 4$; TAS = $SD_6 + SD_5 + 2SD_4 + 4SD_3 + 8$; SA + NT = $SD_7 + SD_6 + 2SD_5 + 5SD_4 + 9SD_3 + 18$; and ACT = $SD_8 + SD_7 + 2SD_6 + 4SD_5 + 8SD_4 + 16SD_3 + 32$ with VIC for Victoria, QLD for Queensland, WA for Western Australia, TAS for Tasmania, (SA+NT) for South Australia and Northern Territory, ACT for Australian Capital Territory, and (1+ACT) for New South Wales and Australian Capital Territory.

Appendix Table A-7. List of Australian Universities by Statistical Regions in Parentheses

NEW SOUTH WALES

Charles Sturt University

Albury campus*, Albury. N.S.W. 2640 (Murray-Murrumbidgee)

Bathurst campus**, Bathurst. N.S.W. 2795 (Central West)

Wagga Wagga campus*, Wagga Wagga. N.S.W. 2650 (Murray-Murrumbidgee)

Macquarie University

Main campus**, North Ryde. N.S.W. 2109 (Lower Northern Sydney)

Institute of Early Childhood Studies, Waverley. N.S.W. 2024 (Eastern Suburbs)

Center for Chiropractic and Osteopathy, Summer Hill. N.S.W. 2130 (Inner Western Sydney)

Southern Cross University, Lismore. N.S.W. 2480 (Richmond-Tweed)

University of Newcastle

Newcastle campus**, Callaghan. N.S.W. 2307 (Hunter)

Central Coast campus, Ourimbah. N.S.W. 2258 (Gosford-Wyong)

Newcastle city campus, Newcastle. N.S.W. 2300 (Hunter)

University of New England

Armidale campus**, Armidale. N.S.W. 2350 (Northern)

Mossman campus, Armidale. N.S.W. 2350 (Northern)

Coffs Harbour campus*, Coffs Harbour. N.S.W. 2450 (Mid-North Coast)

Orange campus*, Orange. N.S.W. 2800 (Central West)

University of New South Wales

Main campus**, Kensington. N.S.W. 2033 (Eastern Suburbs)

St. George campus*, Oatley. N.S.W. 2223 (St. George-Sutherland)

College of Fine Art*, Paddington. N.S.W. 2021 (Eastern Suburbs)

University of Sydney

Main campus**, Camperdown. N.S.W. 2050 ((Inner Sydney)

Faculty of Health Science, Lidcombe. N.S.W. 2141 (Central Western Sydney)

College of the Arts, Roseville. N.S.W. 2039 (Inner Sydney)

Conservatorium of Music, Sydney. N.S.W. 2000 (Inner Sydney)

University of Technology, Sydney

Broadway campus**, Sydney. N.S.W. 2007 (Inner Sydney)

Markets campus, Ultimo. N.S.W. 2007 (Inner Sydney)

Ku-ring-gai campus, Lindfield. N.S.W. 2070 (Lower Northern Sydney)

Design campus, Roseville. N.S.W. 2070 (Lower Northern Sydney)

Dunbar campus, St. Leonards. N.S.W. 2065 (Lower Northern Sydney)

Appendix Table A-7. (Continued)

University of Western Sydney

Hawkesbury campus**, Richmond. N.S.W. 2753 (Outer Western Sydney)

Macarthur campus*, Campbelltown. N.S.W. 2560 (Outer South Western Sydney)

Nepean campus*, Kingswood. N.S.W. 2747 (Outer Western Sydney)

University of Wollongong, Wollongong. N.S.W. 2500** (Illawarra)

VICTORIA

Ballarat University College

Ballarat, Victoria. 3353** (Western VIC. Central Highlands)

Deakin University

Geelong campus**, Geelong. VIC. 3217. (South Western VIC.; Barwon)

Burwood campus*, Burwood. VIC. 3125 (Inner Eastern Melbourne)

Rusden campus, Clayton. VIC. 3168 (Southern Melbourne)

Toorak campus, Malvern. VIC. 3144 (Southern Melbourne)

Warrnambool campus*, Warrnambool. VIC. 3280 (South Western VIC.; Barwon)

La Trobe University

Bundoora campus**, Bundoora. VIC. 3083 (North Eastern Melbourne)

Abbotsford campus, Abbotsford. VIC. 3067 (Inner Eastern Melbourne)

Albury-Wodonga campus, Wodonga. VIC. 3690 (Northern VIC.; Goulburn)

Bendigo campus, Flora Hill. VIC. 3550 (Northern VIC.; Loddon)

Carlton campus, Carlton. VIC. 3053 (Inner Melbourne)

Monash University

Clayton campus**, Clayton. VIC. 3168 (Southern Melbourne)

Caulfield campus*, Caulfield East. VIC. 3145 (Southern Melbourne)

Frankston campus, Frankston. VIC. 3191 (Mornington Peninsula)

College Gippsland campus, Churchill. VIC. 3842 (Eastern Victoria)

Royal Melbourne Institute of Technology (RMIT)

City campus**, Melbourne. VIC. 3001 (Inner Melbourne)

Bundoora campus*, Bundoora. VIC. 3083 (North Eastern Melbourne)

Coburg campus, Coburg. VIC. 3058 (North Western Melbourne)

Appendix Table A-7. (Continued)

Swinburne University of Technology

Main campus**, Hawthorn. VIC. 3122 (Inner Eastern Melbourne)

Prahran campus, Prahran. VIC. 3181 (Inner Melbourne)

Eastern campus, Lilydale. VIC. 3140 (Outer Eastern Melbourne)

University of Melbourne

Main campus**, Parkville. VIC. 3052 (Inner Melbourne)

Kew campus, Kew. VIC. 3101 (Inner Eastern Melbourne)

Hawthorn campus, Hawthorn. VIC. 3122 (Inner Eastern Melbourne)

Victoria College of the Arts, St. Kilda. VIC. 3004 (Inner Melbourne)

Victoria University of Technology

City campus**, Melbourne. VIC. 3000 (Inner Melbourne)

Footscray campus, Footscray. VIC. 3011 (Western Melbourne)

Melton campus, Melton South. VIC. 3338 (Western Melbourne)

St. Albans campus, St. Albans. VIC. 3021 (Western Melbourne)

Werribee campus, Werribee. VIC. 3030 (Western Melbourne)

QUEENSLAND

Bond University, Gold Coast. QLD. 4229** (South-East Moreton)

Griffith University

Nathan campus**, Nathan. QLD. 4111 (Brisbane City Outer Core)

Mount Gravatt campus, Mount Gravatt. QLD. 4122 (Brisbane City Outer Core)

Gold Coast University College*, Southport. QLD. 4205 (South-East Moreton)

James Cook University

Townsville campus**, Townsville. QLD. 4811 (Northern-North-West)

Cairns campus, Westcourt. QLD. 4870 (Far North)

Queensland University of Technology

Kelvin Grove campus**, Red Hill. QLD. 4059 (Brisbane City Inner Core)

Gardens Point, Brisbane. QLD. 4000 (Brisbane City Inner Core)

Carseldine campus, Zillmere. QLD. 4034 (Brisbane City Outer Core)

Kedron Park, Kedron. QLD. 4031 (Brisbane City Inner Core)

Sunshine Coast Centre, Nambour. QLD. 4560 (Northern-North-West)

Appendix Table A-7. (Continued)

University of Central Queensland

North Rockhampton. QLD. 4702** (Mackay-Fitzroy-Central West)

University of Queensland

St. Lucia campus**, St. Lucia. QLD. 4072 (Brisbane City Inner Core)

Gatton campus, Lawes. QLD. 4343 (North-West Moreton)

University of Southern Queensland

Toowoomba campus**, Toowoomba. QLD. 4350 (Darling Downs - South West)

Harvey Bay Study Centre, Pialba. QLD. 4655 (Wide Bay-Burnett)

SOUTH AUSTRALIA

University of Adelaide

North Terrace campus**, Adelaide. S.A. 5000 (Eastern Adelaide)

Roseworthy campus, Roseworthy. S.A. 5371 (Southern and Eastern S.A.)

Waite campus, Glen Osmond. S.A. 5064 (Eastern Adelaide)

Thebarton campus, Thebarton. S.A. 5031 (Western Adelaide)

The Flinders University, Bedford Park. S.A. 5042** (Southern Adelaide)

University of South Australia

City campus**, Adelaide. S.A. 5000 (Eastern Adelaide)

Underdale campus, Underdale. S.A. 5032 (Western Adelaide)

Whyalla campus, Whyalla. S.A. 5600 (Northern and Western S.A.)

WESTERN AUSTRALIA

Curtin University of Technology

Bentley campus**, Bentley. W.A. 6102 (Central Metropolitan)

Institute of Agriculture, Northam. W.A. 6401 (Balance of W.A.)

School of Mines, Kalgoorlie. W.A. 6430 (Balance of W.A.)

School of Occupational Therapy, Shenton Park. W.A. 6008 (Central Metropolitan)

Edith Cowan University

Churchlands campus**, Churchlands. W.A. 6018 (North Metropolitan)

Mount Lawley campus, Mount Lawley. W.A. 6050 (North Metropolitan)

Joondalup campus, Joondalup. W.A. 6027 (North Metropolitan)

Claremont campus, Claremont. W.A. 6010 (Central Metropolitan)

Bunbury campus, Bunbury. W.A. 6230 (Balance of W.A.)

Appendix Table A-7. (Continued)

Murdoch University, Murdoch. W.A. 6150** (South West Metropolitan)

University of Western Australia, Crawley. W.A. 6009** (Central Metropolitan)

TASMANIA

Australian Maritime College

Newnham campus*, Newnham. TAS. 7248 (TAS)

Beauty Point campus, Beauty Point. TAS. 7272 (TAS)

University of Tasmania

Hobart campus**, Sandy Bay. TAS. 7005 (TAS)

Launceston, Newnham. TAS. 7248 (TAS)

NORTHERN TERRITORY

Northern Territory University

Casuarina campus**, Casuarina. N.T. (N.T.)

Myilly campus, Larralceyah. N.T. (N.T.)

AUSTRALIAN CAPITAL TERRITORY

Australian Catholic University

MacKillop campus**, North Sydney. N.S.W. 2060 (Lower Northern Sydney)

Castle Hill campus*, Castle Hill. N.S.W. 2154 (Blacktown-Baulkham Hills)

Aguinas College*, Ballarat. VIC. 3350 (Western Vic.; Central Highlands)

Christ campus*, Oakleigh. VIC. 3600 (Southern Melbourne)

McAuley College*, Mitchelton. QLD. 4053 (Brisbane City Outer Core)

Signadou campus, Canberra. ACT. (ACT)

Australian Defence Force Academy, Campbell. ACT. 2600** (ACT)

Australian National University, Acton. ACT. 2601** (ACT)

University of Canberra, Bruce. ACT. 2616** (ACT)

Notes:
(i) * for campus with the admission office (19)
(ii) ** for campuses with both the chancellor and the admission offices (36)
(iii) Total number of campuses = 110; NSW (32), VIC (32), QLD (17), SA (8), WA (11), TAS (4), NT (2), and ACT (4).
(iv) Statistical region in parentheses.
(v) Bond University (QLD) as a private institution.

Source:
(i) Australian Standard Geographical Classification (ABS Cat.No.1216.0, 1991).
(ii) See Table 4.1 (Notes 4 and 5).
(iii) Gregory's, Sydney Street Directory, A Division of Universal Press Pty. Ltd., Macquarie Park. N.S.W. 1991.
(iv) Land Information Center, The Official Road Directory of New South Wales, N.S.W. Department of Lands, Bathurst, N.S.W. 1990.
(v) UBD, Street Directory, A Division of Universal Press Pty. Ltd., 1993.

Appendix Table A-8. Simple Bivariate Correlation Coefficients Among Selected Variables[1]

Variables[2]	H_1	H_2	H_3	E	Y	R_1
H_1	1.000					
H_2	0.721*	1.000				
H_3	-0.128	-0.242	1.000			
E	-0.288	-0.223	0.469	1.000		
Y	-0.144	-0.178	0.469	0.774*	1.000	
R_1	0.290	0.434	-0.340	-0.218	-0.096	1.000
R_2	0.370	0.390	-0.511	-0.525	-0.352	0.183
R_3	-0.182	-0.412	0.539	0.419	0.253	-0.413
R_4	-0.368	-0.053	0.069	0.276	0.228	0.310
F	-0.144	-0.282	0.131	0.244	0.114	-0.303
OCC_1	-0.142	-0.100	0.349	0.560	0.463	-0.314
OCC_2	0.360	0.352	-0.513	-0.794*	-0.678	0.344
G*VIC	0.211	0.099	0.091	-0.015	0.125	-0.061
G*QLD	-0.326	-0.266	0.056	-0.274	-0.191	0.026
G*WA	-0.240	-0.298	0.189	-0.287	-0.166	-0.009
G*TAS	-0.236	-0.252	0.116	-0.327	-0.201	0.017
G*(SA+NT)	-0.229	-0.215	0.100	-0.315	-0.188	0.037
AGE_3	0.134	0.070	0.233	0.175	0.426	0.244
AGE_4	-0.005	-0.098	0.187	0.314	0.578	0.307
AGE_5	-0.134	-0.171	-0.106	0.058	-0.213	-0.394
AGE_8	-0.020	-0.169	0.232	0.429	0.619	0.132
$FAGE_3$	0.095	0.049	0.221	0.219	0.447	0.224
$FAGE_4$	0.050	-0.021	0.178	0.288	0.559	0.254
$FAGE_5$	-0.136	-0.192	-0.154	-0.034	-0.295	-0.331

Appendix Table A-8. (Continued)[1]

Variables[2]	H_1	H_2	H_3	E	Y	R_1
C	-0.355	-0.336	-0.297	0.135	0.107	-0.142
B_1	-0.189	-0.267	0.013	0.302	0.342	-0.097
B	-0.052	-0.042	-0.015	0.105	0.152	0.119
METD	-0.268	-0.361	0.099	0.422	0.535	-0.178
NGSTD	-0.178	-0.343	0.149	0.580	0.633	-0.331
UD	0.002	-0.071	0.278	0.262	0.230	-0.081
AE	-0.300	-0.312	0.405	0.082	-0.227	-0.491
K	0.182	0.352	0.169	0.227	0.213	0.275
Ln	-0.390	-0.415	0.574	0.650	0.698	-0.356
I	0.059	-0.035	0.543	0.411	0.709*	-0.028
PFL	-0.169	-0.180	0.066	-0.168	-0.526	-0.326
NES	0.098	0.003	-0.238	0.286	0.443	0.062
GNG_1	0.313	0.292	-0.075	-0.503	-0.495	0.019
GNG_2	0.097	0.113	0.101	-0.188	-0.005	0.229
GNG_3	0.193	0.018	-0.237	-0.149	0.060	0.005
GNG_4	0.072	0.142	-0.058	0.058	0.018	-0.025
GNG_5	-0.032	0.109	0.043	0.016	-0.154	0.230
GNG_6	-0.017	-0.104	0.406	0.146	0.217	-0.138
GNG_7	-0.039	-0.065	-0.041	0.026	0.086	-0.037
GNG_8	-0.109	-0.019	-0.088	-0.021	0.213	0.031
GNG_9	-0.398	-0.478	0.330	0.737	0.727	-0.369
GNG_{10}	-0.120	-0.093	0.248	0.266	0.379	0.121
GNG_{11}	0.069	0.101	0.397	0.446	0.436	-0.131

Appendix Table A-8. (Continued)[1]

Variables[2]	R_2	R_3	R_4	F	OCC_1	OCC_2
R_2	1.000					
R_3	-0.821	1.000				
R_4	-0.493	-0.092	1.000			
F	-0.024	0.141	-0.173	1.000		
OCC_1	-0.536	0.406	0.316	0.127	1.000	
OCC_2	0.631	-0.517	-0.312	-0.343	-0.676	1.000
G*VIC	0.049	0.033	-0.135	0.083	0.144	-0.026
G*QLD	0.099	-0.205	0.140	-0.152	-0.235	0.153
G*WA	-0.107	0.114	0.012	-0.163	-0.220	0.069
G*TAS	-0.048	0.029	0.039	-0.166	-0.210	0.080
G*(SA+NT)	-0.046	0.008	0.068	-0.173	-0.193	0.066
AGE_3	0.027	-0.095	0.097	-0.102	0.053	-0.218
AGE_4	-0.081	-0.068	0.245	-0.089	0.058	-0.289
AGE_5	-0.106	0.205	-0.126	0.168	0.129	-0.085
AGE_8	-0.132	0.031	0.183	-0.018	0.180	-0.454
$FAGE_3$	-0.008	-0.057	0.102	-0.007	0.029	-0.210
$FAGE_4$	0.050	-0.149	0.140	-0.141	-0.006	-0.188
$FAGE_5$	-0.078	0.186	-0.147	0.117	0.000	-0.001
C	0.235	-0.204	-0.100	0.339	-0.272	-0.087
B_1	-0.372	0.241	0.281	0.140	0.490	-0.492
B	-0.298	0.095	0.376	0.006	0.523	-0.325
METD	0.002	0.137	-0.213	0.388	-0.073	-0.371

Appendix Table A-8. (Continued)[1]

Variables[2]	R_2	R_3	R_4	F	OCC_1	OCC_2
NGSTD	-0.187	0.226	-0.018	0.337	0.420	-0.636
UD	-0.341	0.271	0.182	0.103	0.393	-0.331
AE	-0.371	0.482	-0.088	0.114	0.028	-0.104
K	-0.199	0.073	0.235	-0.117	0.135	-0.123
Ln	-0.406	0.369	0.145	0.191	0.268	-0.577
I	-0.228	0.153	0.164	-0.160	0.244	-0.436
PFL	-0.346	0.401	-0.009	0.020	0.057	0.026
NES	0.260	-0.147	-0.230	0.165	-0.114	-0.153
GNG_1	0.129	-0.105	-0.064	-0.376	0.104	0.423
GNG_2	-0.185	0.070	0.216	-0.364	-0.036	0.292
GNG_3	0.487	-0.326	-0.353	0.144	-0.465	0.246
GNG_4	0.067	0.064	-0.215	-0.138	-0.029	0.146
GNG_5	-0.069	-0.004	0.125	-0.024	-0.301	0.109
GNG_6	-0.079	0.088	0.004	0.094	-0.111	-0.208
GNG_7	0.248	-0.229	-0.084	-0.161	-0.299	0.134
GNG_8	0.085	-0.143	0.069	-0.116	-0.191	-0.052
GNG_9	-0.430	0.381	0.168	0.353	0.455	-0.787*
GNG_{10}	-0.168	0.117	0.115	-0.031	0.122	-0.272
GNG_{11}	-0.392	0.291	0.240	-0.130	0.615	-0.539

Appendix Table A-8. (Continued)[1]

Variables[2]	G*VIC	G*QLD	G*WA	G*TAS	G*(SA+NT)	AGE_3
G*VIC	1.000					
G*QLD	-0.270	1.000				
G*WA	-0.229	0.681	1.000			
G*TAS	-0.230	0.754*	0.967*	1.000		
G*(SA+NT)	-0.221	0.735*	0.959*	0.995*	1.000	
AGE_3	0.166	0.186	0.195	0.209	0.217	1.000
AGE_4	0.061	0.015	0.136	0.101	0.105	0.324
AGE_5	-0.008	-0.171	-0.223	-0.212	-0.217	-0.322
AGE_8	0.152	0.004	0.078	0.063	0.068	0.675
$FAGE_3$	0.162	0.175	0.209	0.211	0.219	0.871*
$FAGE_4$	0.049	0.063	0.156	0.128	0.133	0.346
$FAGE_5$	-0.014	-0.160	-0.206	-0.197	-0.205	-0.351
C	-0.128	0.002	-0.143	-0.107	-0.105	-0.058
B_1	0.156	0.020	0.092	0.084	0.082	0.259
B	0.008	0.054	0.114	0.102	0.098	0.187
METD	0.084	-0.228	-0.074	-0.108	-0.122	0.207
NGSTD	0.146	-0.169	-0.179	-0.202	-0.205	0.301
UD	0.063	0.129	0.118	0.136	0.127	0.276
AE	-0.233	0.020	0.130	0.085	0.067	-0.410
K	0.084	-0.256	0.156	0.123	0.161	0.175
Ln	0.017	-0.047	-0.056	-0.114	-0.115	0.236
I	0.167	0.041	0.075	0.059	0.065	0.535
PFL	-0.102	0.064	0.229	0.225	0.213	-0.422
NES	0.264	-0.346	-0.256	-0.296	-0.285	0.262
GNG_1	-0.048	0.167	0.078	0.116	0.111	-0.165
GNG_2	-0.164	0.125	0.205	0.139	0.127	0.050
GNG_3	0.275	-0.177	-0.227	-0.210	-0.209	-0.019
GNG_4	0.181	-0.334	-0.216	-0.238	-0.237	-0.029
GNG_5	0.039	0.065	-0.083	-0.092	-0.086	-0.012
GNG_6	0.024	0.138	0.128	0.092	0.081	0.350
GNG_7	-0.232	0.131	0.112	0.075	0.063	-0.126
GNG_8	-0.020	0.097	-0.115	-0.076	-0.082	0.122
GNG_9	-0.002	-0.155	-0.134	-0.180	-0.182	0.221
GNG_{10}	0.185	0.126	0.148	0.174	0.177	0.343
GNG_{11}	0.135	-0.073	0.109	0.118	0.142	0.293

Appendix Table A-8. (Continued)[1]

Variables[2]	AGE_4	AGE_5	AGE_8	$FAGE_3$	$FAGE_4$	$FAGE_5$
AGE_4	1.000					
AGE_5	-0.595	1.000				
AGE_8	0.644	0.012	1.000			
$FAGE_3$	0.414	-0.411	0.608	1.000		
$FAGE_4$	0.913*	-0.596	0.581	0.395	1.000	
$FAGE_5$	-0.623	0.939*	-0.075	-0.466	-0.634	1.000
C	0.121	0.186	0.207	-0.013	0.121	0.172
B_1	0.285	0.011	0.422	0.299	0.151	-0.053
B	0.286	-0.136	0.269	0.216	0.094	-0.191
METD	0.315	0.010	0.415	0.209	0.316	-0.003
NGSTD	0.288	0.134	0.542	0.254	0.214	0.079
UD	0.179	-0.160	0.213	0.213	0.141	-0.183
AE	-0.483	0.297	-0.463	-0.414	-0.431	0.318
K	0.198	-0.249	0.101	0.239	0.205	-0.304
Ln	0.270	-0.003	0.383	0.198	0.274	-0.068
I	0.643	-0.391	0.613	0.514	0.625	-0.426
PFL	-0.508	0.395	-0.420	-0.440	-0.520	0.432
NES	0.366	0.105	0.564	0.283	0.375	0.048
GNG_1	-0.347	-0.048	-0.444	-0.185	-0.309	-0.056
GNG_2	0.002	-0.257	-0.158	0.077	0.007	-0.237
GNG_3	0.195	-0.007	0.154	-0.025	0.230	0.041
GNG_4	-0.086	0.120	-0.005	-0.063	0.016	0.092
GNG_5	-0.183	0.104	-0.092	-0.001	-0.181	0.204
GNG_6	0.145	-0.078	0.290	0.325	0.153	-0.066
GNG_7	0.213	-0.077	0.052	-0.167	0.256	-0.026
GNG_8	0.337	-0.098	0.301	0.046	0.328	-0.112
GNG_9	0.363	0.077	0.515	0.183	0.313	0.009
GNG_{10}	0.298	-0.254	0.290	0.370	0.307	-0.254
GNG_{11}	0.307	-0.160	0.336	0.272	0.312	-0.270

Appendix Table A-8. (Continued)[1]

Variables[2]	C	B_1	B	METD	NGSTD	UD
C	1.000					
B_1	0.200	1.000				
B	-0.117	0.790*	1.000			
METD	0.557	0.285	-0.076	1.000		
NGSTD	0.351	0.573	0.307	0.566	1.000	
UD	-0.116	0.357	0.358	0.001	0.330	1.000
AE	-0.138	-0.299	-0.374	-0.150	-0.212	-0.196
K	-0.263	0.005	0.007	-0.005	-0.077	0.117
Ln	0.189	0.114	-0.131	0.510	0.419	0.135
I	-0.147	0.114	0.088	0.209	0.307	0.230
PFL	-0.203	-0.063	-0.068	-0.337	-0.235	-0.044
NES	0.471	0.198	-0.050	0.622	0.525	-0.043
GNG_1	-0.615	-0.295	0.095	-0.791*	-0.525	-0.108
GNG_2	-0.403	-0.106	0.129	-0.394	-0.284	0.058
GNG_3	0.508	-0.148	-0.343	0.494	0.084	-0.247
GNG_4	0.043	-0.190	-0.218	0.069	-0.142	-0.272
GNG_5	0.006	-0.268	-0.281	-0.187	-0.253	-0.120
GNG_6	0.118	0.038	-0.167	0.205	0.127	-0.049
GNG_7	0.178	-0.144	-0.230	0.190	0.063	-0.178
GNG_8	0.368	0.044	-0.055	0.269	0.083	-0.051
GNG_9	0.298	0.508	0.221	0.624	0.759*	0.199
GNG_{10}	-0.173	-0.011	-0.029	0.119	0.121	0.250
GNG_{11}	-0.229	0.416	0.367	0.048	0.334	0.461

Appendix Table A-8. (Continued)[1]

Variables[2]	AE	K	Ln	I	PFL	NES
AE	1.000					
K	-0.090	1.000				
Ln	0.196	0.041	1.000			
I	-0.280	0.149	0.516	1.000		
PFL	0.610	-0.061	-0.144	-0.349	1.000	
NES	-0.529	0.073	0.200	0.178	-0.520	1.000
GNG_1	0.103	-0.106	-0.448	-0.096	0.272	-0.579
GNG_2	0.009	0.174	-0.164	0.076	-0.003	-0.284
GNG_3	-0.290	-0.167	0.051	0.065	-0.397	0.558
GNG_4	0.006	0.026	-0.036	0.121	0.091	0.088
GNG_5	0.297	0.009	0.040	-0.174	0.197	-0.117
GNG_6	0.176	-0.048	0.463	0.344	-0.020	-0.005
GNG_7	-0.034	-0.109	-0.003	0.075	-0.187	0.168
GNG_8	-0.203	-0.284	0.142	0.271	-0.344	0.231
GNG_9	-0.028	-0.088	0.647	0.406	-0.197	0.395
GNG_{10}	-0.103	0.243	0.232	0.354	-0.234	0.093
GNG_{11}	-0.149	0.441	0.207	0.457	0.017	-0.036

Appendix Table A-8. (Continued)[1]

Variables[2]	GNG_1	GNG_2	GNG_3	GNG_4	GNG_5	GNG_6
GNG_1	1.000					
GNG_2	0.309	1.000				
GNG_3	-0.385	-0.300	1.000			
GNG_4	0.066	0.052	0.101	1.000		
GNG_5	-0.144	-0.062	-0.068	-0.026	1.000	
GNG_6	-0.229	-0.118	0.086	-0.182	0.295	1.000
GNG_7	-0.177	0.049	0.207	-0.042	-0.125	0.072
GNG_8	-0.316	-0.143	0.308	0.031	0.036	0.147
GNG_9	-0.562	-0.293	0.047	-0.084	-0.268	0.251
GNG_{10}	-0.175	-0.071	-0.163	-0.009	0.111	-0.174
GNG_{11}	-0.053	-0.036	-0.361	0.026	-0.298	-0.014

Appendix Table A-8. (Continued)[1]

Variables[2]	GNG_7	GNG_8	GNG_9	GNG_{10}	GNG_{11}
GNG_7	1.000				
GNG_8	0.226	1.000			
GNG_9	0.092	0.206	1.000		
GNG_{10}	-0.173	0.010	0.025	1.000	
GNG_{11}	-0.108	-0.055	0.262	0.250	1.000

Notes: **1.** * indicates a high correlation between the two variables ($r > 0.7$). See, for example, Lewis, O'Brien and Thampapillai (1990, p.309).

2. $G_j * STATES_j$ reflects the data limitation in the absence of surveys of public health expenditures by the statistical regions and allows us to seperately compare the six different states and territories with New South Wales, the most populous of the six states and two territories, treated as a base and combining South Australia and Northern Territory as one group. The STATES variables are measured as: VIC = $1+SD_2$, QLD = $1+SD_3$, WA = $SD_5+SD_4+2SD_3+4$, TAS = $SD_6+SD_5+2SD_4+4SD_3+8$, ACT = $SD_8+SD_7+2SD_6+4SD_5+8SD_4+16SD_3+32$, SA+NT = $SD_7+SD_6+2SD_5+5SD_4+9SD_3+18$ where $SD_1 = 1$ for statistical regions in New South Wales (NSW), $SD_2 = 1$ for those in Victoria (VIC), $SD_3 = 1$ for those in Queensland (QLD), $SD_4 = 1$ for those in South Australia (SA), $SD_5 = 1$ for those in Western Australia (WA), $SD_6 = 1$ for Tasmania (TAS), $SD_7 = 1$ for Northern Territory (NT) and $SD_8 = 1$ for Australian Capital Territory (ACT), and 0 elsewhere.

Appendix Table A-9 Contingency Table: Cross-Classification Between No Recent Illness (H_1) and Education (E)

Level of Education												
Top Third				Middle Third				Bottom Third				
TH_1	MH_1	BH_1	Sub-Total	TH_1	MH_1	BH_1	Sub-Total	TH_1	MH_1	BH_1	Sub-Total	Total
0.150	0.350	0.500	1.000	0.286	0.333	0.381	1.000	0.450	0.350	0.200	1.000	3.000
0.049 (3)	0.115 (7)	0.164 (10)	0.328 (20)	0.098 (6)	0.115 (7)	0.131 (8)	0.344 (21)	0.148 (9)	0.115 (7)	0.066 (4)	0.328 (20)	1.000 (61)

Notes:
1. TH_1, MH_1, and BH_1 imply, respectively, the top third of regions, the middle third of regions, and the bottom third of regions in terms of sample proportions.
2. Values in parentheses denote the number of statistical regions.
3. For more details see Lewis, O'Brien, and Thampapillai (1990, pp.197-201).

Appendix Table A-10 Contingency Table: Cross-Classification Between No Recent Illness (H_1) and Income (Y)

Level of Income												
Top Third				Middle Third				Bottom Third				
TH_1	MH_1	BH_1	Sub-Total	TH_1	MH_1	BH_1	Sub-Total	TH_1	MH_1	BH_1	Sub-Total	Total
0.190	0.381	0.429	1.000	0.286	0.148	0.286	1.000	0.421	0.211	0.368	1.000	3.000
0.066 (4)	0.131 (8)	0.147 (9)	0.344 (21)	0.098 (6)	0.148 (9)	0.098 (6)	0.344 (21)	0.131 (8)	0.066 (4)	0.115 (7)	0.311 (19)	1.000 (61)

Notes: **1.** TH_1, MH_1,and BH_1 imply, respectively, the top third of regions, the middle third of regions, and the bottom third of regions in terms of sample proportions.
2. Values in parentheses denote the number of statistical regions.
3. For more details see Lewis, O'Brien, and Thampapillai (1990, pp.197-201).

Appendix Table A-11 Contingency Table: Cross-Classification Between No Chronic Condition (H_2) and Education (E)

Level of Education												
Top Third				Middle Third				Bottom Third				
TH_2	MH_2	BH_2	Sub-Total	TH_2	MH_2	BH_2	Sub-Total	TH_2	MH_2	BH_2	Sub-Total	Total
0.250	0.250	0.500	1.000	0.381	0.381	0.238	1.000	0.400	0.350	0.250	1.000	3.000
0.082 (5)	0.082 (5)	0.164 (10)	0.328 (20)	0.131 (8)	0.131 (8)	0.082 (5)	0.344 (21)	0.131 (8)	0.115 (7)	0.082 (5)	0.328 (20)	1.000 (61)

Notes: **1.** TH_2, MH_2, and BH_2 imply, respectively, the top third of regions, the middle third of regions, and the bottom third of regions in terms of sample proportions.
2. Values in parentheses denote the number of statistical regions.
3. For more details see Lewis, O'Brien, and Thampapillai (1990, pp.197-201).

Appendix Table A-12. Contingency Table: Cross-Classification Between No Chronic Condition (H_2) and Income (Y)

Level of Income												
Top Third				Middle Third				Bottom Third				
TH_2	MH_2	BH_2	Sub-Total	TH_2	MH_2	BH_2	Sub-Total	TH_2	MH_2	BH_2	Sub-Total	Total
0.238	0.381	0.381	1.000	0.333	0.286	0.381	1.000	0.474	0.316	0.210	1.000	3.000
0.082 (5)	0.131 (8)	0.131 (8)	0.344 (21)	0.115 (7)	0.098 (6)	0.131 (8)	0.344 (21)	0.148 (9)	0.098 (6)	0.066 (4)	0.311 (19)	1.000 (61)

Notes: **1.** TH_2, MH_2, and BH_2 imply, respectively, the top third of regions, the middle third of regions, and the bottom third of regions in terms of sample proportions.
2. Values in parentheses denote the number of statistical regions.
3. For more details see Lewis, O'Brien, and Thampapillai (1990, pp.197-201).

Appendix Table A-13. Contingency Table: Cross-Classification Between Self-assessed Good Health (H_3) and Education (E)

Level of Education												
Top Third				Middle Third				Bottom Third				
TH_3	MH_3	BH_3	Sub-Total	TH_3	MH_3	BH_3	Sub-Total	TH_3	MH_3	BH_3	Sub-Total	Total
0.600	0.200	0.200	1.000	0.286	0.333	0.381	1.000	0.100	0.400	0.500	1.000	3.000
0.196 (12)	0.066 (4)	0.066 (4)	0.328 (20)	0.098 (6)	0.115 (7)	0.131 (8)	0.344 (21)	0.033 (2)	0.131 (8)	0.164 (10)	0.328 (20)	1.000 (61)

Notes: **1.** TH_3, MH_3, and BH_3 imply, respectively, the top third of regions, the middle third of regions, and the bottom third of regions in terms of sample proportions.
2. Values in parentheses denote the number of statistical regions.
3. For more details see Lewis, O'Brien, and Thampapillai (1990, pp.197-201).

Appendix Table A-14. Contingency Table: Cross-Classification Between Self-assessed Good Health (H_3) and Income (Y)

Level of Income												
Top Third				Middle Third				Bottom Third				
TH_3	MH_3	BH_3	Sub-Total	TH_3	MH_3	BH_3	Sub-Total	TH_3	MH_3	BH_3	Sub-Total	Total
0.571	0.238	0.191	1.000	0.333	0.333	0.333	1.000	0.053	0.368	0.579	1.000	3.000
0.197 (12)	0.082 (5)	0.066 (4)	0.345 (21)	0.115 (7)	0.115 (7)	0.115 (7)	0.345 (21)	0.016 (1)	0.114 (7)	0.180 (11)	0.310 (19)	1.000 (61)

Notes: **1.** TH_3, MH_3, and BH_3 imply, respectively, the top third of regions, the middle third of regions, and the bottom third of regions in terms of sample proportions.
2. Values in parentheses denote the number of statistical regions.
3. For more details see Lewis, O'Brien, and Thampapillai (1990, pp.197-201).

Appendix Table A-15. Contingency Table: Cross-Classification Between Education (E) and Income (Y)

Level of Income												
Top Third				Middle Third				Bottom Third				
TE	ME	BE	Sub-Total	TE	ME	BE	Sub-Total	TE	ME	BE	Sub-Total	Total
0.850	0.050	0.100	1.000	0.143	0.571	0.286	1.000	0.050	0.400	0.550	1.000	3.000
0.279 (17)	0.016 (1)	0.033 (2)	0.328 (20)	0.049 (3)	0.197 (12)	0.098 (6)	0.344 (21)	0.016 (1)	0.031 (8)	0.181 (11)	0.328 (20)	1.000 (61)

Notes:
1. TE, ME, and BE imply, respectively, the top third of regions, the middle third of regions, and the bottom third of regions in terms of sample proportions.
2. Values in parentheses denote the number of statistical regions.
3. For more details see Lewis, O'Brien, and Thampapillai (1990, pp.197-201).

Appendix Table A-16. Reduced-form Estimates of the Pairs of Structural Equations for Three Indicators of Health, Education, and Income

Explanatory Variable	No recent illness (H_1)	No chronic condition (H_2)	Self-assessed good health (H_3)	Education (E)	Income (Y)
	Dependent Variables[1]				
R_2	0.5450 (3.174)***	1.0639 (3.595)***	0.1187 (0.731)	0.0427 (0.198)	-0.0041 (0.059)
R_3	0.4725 (2.812)***	0.5908 (2.040)**	0.1847 (1.162)	0.2797 (1.324)	0.0400 (0.586)
F	0.3829 (1.209)	-0.4134 (0.757)	0.1624 (0.543)	-0.6814 (1.712)*	-0.0680 (0.528)
OCC_1	-0.1782 (1.726)*	-0.3169 (1.780)*	0.0847 (0.868)	0.1760 (1.356)	0.1165 (2.775)***
OCC_2	-0.0320 (0.174)	-0.1410 (0.446)	-0.1077 (0.621)	-0.1520 (0.659)	0.0389 (0.522)
C	-0.8108 (3.881)***	-1.2819 (3.560)***	-0.0055 (0.028)	0.3001 (1.142)	-0.0100 (0.118)
B_1	1.5655 (2.240)**	0.0514 (0.043)	-0.8149 (1.233)	-1.0648 (1.212)	0.0618 (0.218)
B	-0.1239 (0.572)	0.6999 (1.875)*	0.2348 (1.146)	0.2604 (0.956)	-0.0374 (0.425)
G*VIC	-1.3176 (2.399)**	-1.2722 (1.344)	1.0691 (2.059)**	-0.4909 (0.711)	-0.0562 (0.252)
G*QLD	-2.8120 (1.959)*	1.5416 (0.623)	2.0521 (1.512)	6.0914 (3.376)***	0.0678 (0.116)
G*WA	-2.0310 (1.516)	-6.3697 (2.758)***	1.1503 (0.908)	-0.5586 (0.332)	-0.7232 (1.328)
G*TAS	0.8155 (0.386)	-8.9189 (2.451)**	0.9724 (0.487)	-2.7652 (1.042)	-0.0444 (0.052)
G*(SA+NT)	0.0351 (0.043)	4.5396 (3.257)***	-0.7745 (1.013)	0.5156 (0.507)	0.1534 (0.467)
METD	-4.7661 (2.725)**	-1.7323 (0.575)	0.9758 (0.590)	1.0395 (0.473)	0.5324 (0.749)
NGSTD	0.0627 (0.935)	-0.0739 (0.640)	-0.0821 (1.296)	-0.0539 (0.639)	0.0376 (1.379)
UD	0.7594 (0.907)	2.4478 (1.695)*	0.9074 (1.146)	0.4037 (0.383)	-0.7204 (2.116)**
AE	-0.1207 (0.960)	0.0623 (0.287)	0.4611 (3.878)***	0.2813 (1.780)*	0.0284 (0.556)

Appendix Table A-16. (Continued)

Explanatory Variable	Dependent Variables[1] No recent illness (H_1)	No chronic condition (H_2)	Self-assessed good health (H_3)	Education (E)	Income (Y)
K	0.6911 (1.176)	3.4440 (3.399)***	0.4763 (0.857)	2.0906 (2.829)***	0.1770 (0.741)
Ln	0.0226 (0.202)	-0.0616 (0.319)	0.0163 (0.154)	-0.0820 (0.582)	0.0794 (1.744)*
I	0.1079 (0.759)	-0.0444 (0.181)	0.5022 (3.738)***	0.3762 (2.106)**	0.2149 (3.720)***
PFL	-0.0513 (0.580)	0.0354 (0.233)	-0.1147 (1.373)	-0.1303 (1.173)	-0.1484 (4.133)***
GNG_1	-0.0948 (0.467)	-0.3412 (0.976)	-0.2506 (1.307)	-0.6879 (2.699)**	-0.1835 (2.227)**
GNG_2	-0.0862 (0.290)	0.1249 (0.244)	-0.0742 (0.264)	-0.4856 (1.300)	0.2901 (2.402)**
GNG_3	0.3249 (1.927)*	0.0167 (0.057)	-0.1483 (0.930)	-0.2943 (1.388)	-0.0537 (0.783)
GNG_4	0.2219 (0.422)	1.5274 (1.686)	-0.3777 (0.760)	0.7946 (1.202)	-0.0382 (0.179)
GNG_6	0.2650 (1.112)	0.6812 (1.659)	-0.0128 (0.057)	-0.1726 (0.576)	-0.0518 (0.535)
GNG_8	0.3190 (0.479)	1.6908 (1.473)	-0.9256 (1.470)	-2.9592 (3.536)***	-0.2334 (0.862)
GNG_9	-0.3190 (1.128)	-0.2302 (0.472)	-0.2197 (0.822)	0.5795 (1.630)	0.1235 (1.074)
GNG_{10}	-0.0913 (0.438)	-0.1811 (0.504)	-0.2744 (1.393)	-0.0302 (0.115)	0.0641 (0.757)
Constant	-14.5931 (0.536)	-1.6851 (0.036)	4.9120 (0.191)	24.4438 (0.714)	-0.4942 (0.045)
R^2	0.820	0.849	0.850	0.917	0.955
F-Stat	4.881***	6.007***	6.040***	11.789***	22.700***

Notes: 1. Values in parentheses are the estimated absolute t-values. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Appendix Table A-17. Estimates of the No Recent Illness Equation[1]

Explanatory Variables	Specification I[2]		Specification II[3]	
	OLS	TSLS	OLS	TSLS
E	0.2499 (2.334)**	0.2707 (1.376)		
Y			0.8606 (4.432)***	0.8848 (3.738)***
R_2	0.6747 (4.895)***	0.6825 (4.386)***	0.6315 (5.420)***	0.6330 (5.172)***
R_3	0.4053 (2.573)**	0.4068 (2.480)**	0.4622 (3.316)***	0.4643 (3.168)***
F	0.2634 (0.835)	0.2727 (0.812)	0.4799 (1.689)*	0.4891 (1.623)
OCC_1	-0.2879 (3.085)***	-0.2939 (2.731)***	-0.3253 (4.002)***	-0.3283 (3.792)***
C	-0.8231 (4.433)***	-0.8289 (4.187)***	-0.6919 (4.275)***	-0.6901 (4.061)***
B_1	1.5424 (2.300)**	1.5545 (2.211)**	1.6333 (2.770)***	1.6400 (2.649)**
B	-0.0001 (0.005)	0.0018 (0.009)	-0.0437 (0.268)	-0.0440 (0.257)
G*VIC	-0.1253 (0.243)	-0.1058 (0.190)	-0.4186 (0.941)	-0.4203 (0.901)
G*QLD	-3.7457 (2.869)***	-3.8057 (2.653)**	-3.7119 (3.296)***	-3.7312 (3.149)***
G*WA	-2.3235 (1.848)*	-2.3503 (1.777)*	-2.8253 (2.535)**	-2.8485 (2.426)**
G*TAS	0.3580 (0.171)	0.4687 (0.200)	1.0580 (0.579)	1.1150 (0.575)
G*(SA+NT)	0.4381 (0.553)	0.4072 (0.475)	0.2080 (0.299)	0.1911 (0.260)

Appendix Table A-17. (Continued)[1]

Explanatory Variables	Specification I[2]		Specification II[3]	
	OLS	TSLS	OLS	TSLS
METD	-3.1314 (2.131)**	-3.2056 (1.964)*	-4.4424 (3.308)***	-4.5043 (3.124)***
GNG_8	1.1240 (1.693)*	1.1658 (1.528)	0.3460 (0.622)	0.3383 (0.579)
GNG_9	-0.1841 (0.827)	-0.2005 (0.759)	-0.3190 (1.626)	-0.3283 (1.557)
GNG_{10}	-0.1116 (0.730)	-0.1233 (0.673)	-0.2328 (1.700)*	-0.2402 (1.620)
Constant	-10.6449 (0.537)	-12.4063 (0.501)	-21.7260 (1.267)	-22.6313 (1.221)
R^2	0.680	0.655	0.753	0.728
S.E. of estimates	2.504	2.602	2.202	2.309
Exogeneity (joint) Test[4]	$F(1,41)$ = 4.963**		$F(1,41)$ = 7.107**	
S.E. of estimates	2.533		2.227	
ξ[5]			0.478	0.491

Notes: **1.** Values in parentheses are the estimated absolute t-values. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

2. Education is the suspected regressor of endogeneity. The reduced-form estimates are given in Appendix Table A-24.

3. Income is the suspected regressor of endogeneity. The reduced-form estimates are given in Appendix Table A-25.

4. The exogeneity test is a variant of the Hausman test as described by Beggs (1988, p.96) applied to each of education and income, the suspected regressors of endogeneity.

5. ξ is the income elasticity at mean. Mean values of the predicted values of H_1 and Y are 29.16 and 16.19, respectively.

Appendix Table A-18. Estimates of the No Chronic Condition Equation[1]

Explanatory Variables	Specification I[2]		Specification II[3]	
	OLS	TSLS	OLS	TSLS
E	0.6410 (3.759)***	1.1248 (3.720)***		
Y			0.8269 (2.147)**	0.8190 (1.803)*
R_2	0.7255 (3.305)***	0.9070 (3.793)***	0.5336 (2.308)**	0.5332 (2.269)**
R_3	0.1363 (0.543)	0.1710 (0.678)	0.1622 (0.587)	0.1615 (0.574)
F	0.1204 (0.240)	0.3363 (0.0652)	0.1495 (0.265)	0.1465 (0.253)
OCC_1	-0.3942 (2.652)**	-0.5337 (3.226)***	-0.3145 (1.950)*	0.3135 (1.886)*
C	-1.5316 (5.719)***	-1.6645 (5.470)***	-1.2953 (4.034)***	-1.2959 (3.972)***
B_1	1.4575 (1.365)	1.7395 (1.610)	1.3112 (1.121)	1.3090 (1.101)
B	0.2408 (0.812)	0.3036 (1.016)	0.1476 (0.456)	0.1477 (0.450)
G*VIC	-0.6235 (0.760)	-0.1720 (0.201)	-1.2795 (1.450)	-1.2789 (1.428)
G*QLD	-4.5136 (2.171)**	-5.9079 (2.680)**	-3.3256 (1.489)	-3.3193 (1.460)
G*WA	-5.2038 (2.599)**	-5.8262 (2.866)***	-5.1702 (2.338)**	-5.1626 (2.290)**
G*TAS	-6.2231 (1.868)*	-3.6515 (1.016)	-7.6816 (2.118)**	-7.7002 (2.070)**
G*(SA+NT)	3.8632 (3.064)***	3.1462 (2.389)**	4.2364 (3.071)***	4.2419 (3.011)***
METD	-1.6143 (0.690)	-3.3379 (1.331)	-1.4457 (0.543)	-1.4254 (0.515)

Appendix Table A-18. (Continued)[1]

Explanatory Variables	Specification I[2]		Specification II[3]	
	OLS	TSLS	OLS	TSLS
GNG_8	2.9632 (2.802)***	3.9351 (3.356)***	1.4102 (1.279)	1.4128 (1.260)
GNG_9	-1.0629 (2.996)***	-1.4421 (3.554)***	-0.8783 (2.257)**	-0.8752 (2.162)**
GNG_{10}	-0.4582 (1.882)*	-0.7299 (2.594)**	-0.3496 (1.287)	-0.3472 (1.220)
Constant	23.2655 (0.736)	-17.6455 (0.464)	46.5211 (1.368)	46.8169 (1.316)
R^2	0.770	0.769	0.724	0.716
S.E. of estimates	3.988	3.999	4.369	4.433
Exogeneity (joint) Test[4]	$F(1,41)$ $= 8.807$***		$F(1,41)$ $= 6.134$**	
S.E. of estimates	3.854		4.421	
ξ[5]			0.393	0.389

Notes: **1.** Values in parentheses are the estimated absolute t-values. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

2. Education is the suspected regressor of endogeneity. The reduced-form estimates are given in Appendix Table A-23.

3. Income is the suspected regressor of endogeneity. The reduced-form estimates are given in Appendix Table A-24.

4. The exogeneity test is a variant of the Hausman test as described by Beggs (1988, p.96) applied to each of education and income, the suspected regressors of endogeneity.

5. ξ is the income elasticity at mean. Mean values of the predicted values of H_2 and Y are 34.07 and 16.19, respectively.

Appendix Table A-19. Estimates of the Self-assessed Good Health Equation[1]

Explanatory Variables	Specification I[2]		Specification II[3]	
	OLS	TSLS	OLS	TSLS
E	0.3839 (3.201)***	0.6772 (3.327)***		
Y			0.8723 (3.494)***	0.9935 (3.392)***
R_2	0.0395 (0.265)	0.1193 (0.770)	-0.3312 (0.230)	-0.0286 (0.198)
R_3	0.1967 (1.197)	0.1771 (1.083)	0.2797 (1.725)*	0.2877 (1.760)*
F	0.5483 (1.540)	0.665 (1.853)*	0.7039 (1.967)*	0.7470 (2.052)**
OCC_1	0.0071 (0.081)	-0.0428 (0.467)	-0.0181 (0.205)	-0.0306 (0.341)
C	-0.4647 (2.416)**	-0.5112 (2.652)**	-0.3045 (1.601)	-0.2907 (1.512)
G*VIC	0.7456 (1.382)	1.1039 (1.929)*	0.3134 (0.614)	0.3185 (0.620)
G*QLD	1.1909 (0.796)	0.2702 (0.172)	1.6920 (1.177)	1.5942 (1.099)
G*WA	2.0902 (1.489)	1.8859 (1.349)	1.5928 (1.142)	1.4865 (1.055)
G*TAS	1.1366 (0.476)	3.0433 (1.169)	0.8830 (0.382)	1.1945 (0.507)
G*(SA+NT)	-0.8040 (0.910)	-1.4003 (1.491)	-0.7111 (0.829)	-0.8066 (0.926)
METD	0.1491 (0.090)	-0.8303 (0.477)	-0.7012 (0.413)	-0.9975 (0.571)
GNG_8	1.0951 (1.450)	1.6971 (2.063)**	0.0104 (0.015)	-0.0308 (0.043)
GNG_9	-0.2523 (0.999)	-0.4828 (1.710)*	-0.2776 (1.117)	-0.3231 (1.260)

Appendix Table A-19. (Continued)[1]

Explanatory Variables	Specification I[2]		Specification II[3]	
	OLS	TSLS	OLS	TSLS
GNG_{10}	-0.1173 (0.664)	-0.3022 (1.482)	-0.1562 (0.886)	-0.1952 (1.062)
Constant	21.4052 (0.950)	-2.0650 (0.080)	20.1997 (0.918)	15.7636 (0.691)
R^2	0.577	0.584	0.592	0.587
S.E. of estimates	2.908	2.887	2.858	2.875
Exogeneity (joint) Test[4]	F(1,43) = 4.241**		F(1,43) = 4.088**	
S.E. of estimates	2.837		2.869	
ξ[5]			0.179	0.204

Notes: **1.** Values in parentheses are the estimated absolute t-values. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

2. Education is the suspected regressor of endogeneity. The reduced-form estimates are given in Appendix Table A-24.

3. Income is the suspected regressor of endogeneity. The reduced-form estimates are given in Appendix Table A-25.

The exogeneity test is a variant of the Hausman test as described by Beggs (1988, p.96) applied to each of education and income, the suspected regressors of endogeneity.

5. ξ is the income elasticity at mean. Mean values of the predicted values of H_3 and Y are 79.00 and 16.19, respectively.

Appendix Table A-20. The Ceteris Paribus Estimates of the Education Equation[1,2]

Suspected Regressors of Endogeneity	Specification I		Specification II		Specification III		Specification IV	
	OLS	TSLS	OLS	TSLS	OLS	TSLS	OLS	TSLS
H_1	0.1841 (0.894)	0.1122 (0.424)	0.1467 (0.853)	0.2266 (0.998)				
H_2					0.1523 (1.467)	0.2537 (2.082)**	0.1047 (1.162)	0.1910 (1.674)*
H_3					0.4006 (1.662)	0.9612 (2.903)***	0.0444 (0.197)	0.3651 (1.002)
Y			1.4941 (4.546)***	1.6516 (4.256)***			1.4550 (4.055)***	1.3597 (2.984)***
R^2	0.741	0.738	0.824	0.814	0.760	0.785	0.826	0.822
S.E. of estimates	4.070	4.098	3.395	3.488	3.963	3.747	3.410	3.450

Notes: **1.** The control variables are R_2, R_3, OCC_1, OCC_2, C, NGSTD, UD, GNG_1, GNG_2, GNG_3, GNG_4, GNG_6, GNG_{10}, AE. Their estimated results are available on request.

2. Values in parentheses are the estimated absolute t-values. ***, **, and * indicate significance at the 1%, 5%, and 10% levels on a two-tailed test, respectively.

Appendix Table A-21. The Ceteris Paribus Estimates of the Income Equation[1,2]

Suspected Regressors of Endogeneity	Specification I		Specification II		Specification III		Specification IV		Specification V		Specification VI	
	OLS	TSLS	OLS	TSLS	OLS	TSLS	OLS	TSLS	OLS	TSLS	OLS	TSLS
H_1									0.1249	0.0806	0.1015	0.0431
									(2.604)**	(1.115)	(2.274)**	(0.614)
H_2	0.0592	0.0567	0.0374	0.0298								
	(2.340)**	(1.745)*	(1.505)	(0.855)								
H_3					0.0055	-0.0296	-0.0195	-0.0763	0.0241	0.0000	-0.0017	-0.0580
					(0.109)	(0.462)	(0.425)	(1.128)	(0.502)	(0.004)	(0.038)	(0.831)
E			0.0899	0.0803			0.1070	0.1096			0.0953	0.1040
			(2.810)***	(1.846)*			(3.400)***	(2.686)**			(3.117)***	(2.473)**
R^2	0.934	0.930	0.943	0.935	0.926	0.926	0.941	0.936	0.935	0.928	0.947	0.937
S.E. of estimates	0.886	0.908	0.828	0.885	0.936	0.934	0.846	0.878	0.884	0.932	0.810	0.884

Notes:

1. The control variables are R_2, R_3, OCC_1, C. K, Ln, I, PFL, GNG_2, GNG_9, GNG_{10}, METD. Their estimated results are available on request.
2. Values in parentheses are the estimated absolute t-values. ***, **, and * indicate significance at the 1%, 5%, and 10% levels on a two-tailed test, respectively.

Appendix Table A-22. Estimation of the Education Equation[1]

Explanatory Variables	Specification I[2]			Specification II[3]	
	OLS	TSLS	REG (TSLS)[4]	OLS	TSLS
H_1	0.0213 (0.076)	-0.4001 (0.769)	0.0194 (0.041)		
H_2	0.1447 (1.000)	0.2629 (1.005)	0.0788 (0.360)		
H_3	0.4006 (1.643)	0.3448 (0.373)	0.4716 (0.615)		
H_1*UD			0.3518 (1.460)		
Y			1.5691 (4.514)***	1.5075 (4.606)***	1.4586 (3.541)***
R_2	-0.2925 (1.156)	-0.1668 (0.447)	-0.3120 (1.001)	-0.1320 (0.670)	-0.1324 (0.626)
R_3	-0.2415 (0.927)	-0.0617 (0.170)	-0.2582 (0.848)	-0.1610 (0.833)	-0.1626 (0.783)
OCC_1	0.2027 (1.397)	0.1908 (1.243)	-0.0011 (0.008)	0.0245 (0.189)	0.0310 (0.219)
OCC_2	-0.3530 (1.399)	-0.3499 (0.789)	-0.0098 (0.026)	-0.2210 (1.072)	-0.2310 (1.024)
C	0.1811 (0.553)	0.0161 (0.027)	0.2514 (0.495)	-0.0496 (0.212)	-0.0525 (0.209)
NGSTD	0.1290 (1.311)	0.1481 (1.166)	0.0250 (0.230)	-0.0286 (0.354)	-0.0252 (0.287)
UD	-0.8404 (0.593)	-0.5684 (0.245)	-10.1375 (1.403)	1.2760 (1.103)	1.2367 (0.987)
AE	0.0929 (0.568)	0.0484 (0.196)	0.2121 (1.027)	0.2849 (2.277)**	0.2804 (2.066)**
GNG_1	-0.5191 (2.235)**	-0.4565 (1.756)*	-0.4227 (1.898)*	-0.3545 (1.840)*	-0.3613 (1.730)*
GNG_2	-0.0362 (0.082)	0.0037 (0.005)	-1.1007 (1.725)*	-0.7960 (1.934)*	-0.7641 (1.650)

Appendix Table A-22. (Continued)[1]

Explanatory Variables	Specification I[2] OLS	TSLS	REG (TSLS)[4]	Specification II[3] OLS	TSLS
GNG_3	-0.0947 (0.415)	0.0247 (0.068)	-0.5269 (1.630)	-0.3449 (1.823)*	-0.3340 (1.602)
GNG_4	1.2398 (1.645)	1.1579 (1.234)	0.9399 (1.177)	1.0764 (1.734)*	1.0935 (1.632)
GNG_6	-0.1731 (0.522)	-0.0724 (0.103)	-0.7323 (1.214)	-0.3704 (1.405)	-0.3552 (1.221)
GNG_{10}	0.1045 (0.427)	0.0924 (0.304)	-0.4463 (1.610)	-0.3612 (1.624)	-0.3453 (1.388)
Constant	26.5306 (0.970)	28.0059 (0.425)	6.1676 (0.111)	36.8766 (1.968)*	37.6453 (1.846)*
R^2	0.760	0.743	0.834	0.821	0.794
S.E. of estimates	4.009	4.148	3.413	3.385	3.631
Exogeneity (joint) Test[5]	$F(3,37) = 6.620$***			$F(1,43) = 12.626$***	
S.E. of estimates	4.087			3.421	
ξ[6]				0.519	0.502

Notes: 1. Values in parentheses are the estimated absolute t-values. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

2. Three indicators of health (H_1, H_2 and H_3) are the suspected regressors of endogeneity. The reduced-form estimates are given in Appendix Table A-24.

3. Income (Y) is the suspected regressor of endogeneity. The reduced-form estimates are given in Appendix Table A-26.

4. REG (TSLS) performs the coefficient stability tests due to the university location dummy in the two-stage least squares estimates (TSLS). The null hypothesis of no coefficient instability is rejected at the 1% level: The Beggs (1988, p.97) and Dowrick (1993, p.2) tests $F(2,41) = 10.852$ and $F(22,39) = 8.774$, respectively. H_1*UD and Y are added to the equation. The income variable (Y) is also added to this equation since improved specification often comes at the price of a radical increase in multicollinearity. The increase in multicollinearity from the addition of the income variable is not destructive. The inclusion of the income variable increases the coefficient of multiple determination (R_2) at the same time. For more details see Auster, Leveson, and Sarachek (1969, p.426). Furthermore, the null hypothesis of functional form misspecification is rejected: RESET2 test T = 0.500.

5. The exogeneity test is a variant of the Hausman test as described by Beggs (1988, p.96) applied to the suspected regressor(s) of endogeneity.

6. ξ is the income elasticity at mean. Mean values of the predicted values of E and Y are 47.06 and 16.19, respectively.

Appendix Table A-23. Estimates of the Income Equation[1]

Explanatory Variables	Specification I[2] OLS	TSLS	REG (TSLS)[4]	Specification II[3] OLS	TSLS
H_1	0.0887 (1.563)	0.0346 (0.335)	0.0863 (0.895)		
H_2	0.0357 (1.183)	0.0309 (0.522)	-0.0194 (0.329)		
H_3	0.0295 (0.616)	-0.1433 (1.150)	-0.1437 (1.222)		
H_2*UD			0.0209 (0.582)		
H_3*UD			0.0492 (0.871)		
UD			-5.0047 (1.028)		
E			0.0955 (3.010)***	0.1049 (3.406)***	0.1454 (2.352)**
R_2	-0.0464 (0.879)	-0.0124 (0.200)	-0.0219 (0.379)	0.0132 (0.287)	0.0174 (0.356)
R_3	-0.0394 (0.690)	0.0368 (0.496)	-0.0029 (0.041)	-0.0001 (0.017)	-0.0040 (0.084)
OCC_1	0.1010 (4.383)***	0.1038 (4.231)***	0.0973 (4.079)***	0.0811 (3.564)***	0.0724 (2.722)***
C	0.1367 (2.249)**	0.0710 (0.919)	0.0406 (0.538)	0.0402 (0.740)	0.0263 (0.438)
METD	1.1097 (2.421)**	0.9774 (2.035)**	1.1449 (2.573)**	1.0992 (2.524)**	1.1540 (2.479)**
K	0.1677 (0.937)	0.2667 (1.198)	0.1112 (0.511)	0.1782 (1.092)	0.1035 (0.524)
Ln	0.1051 (3.227)***	0.1453 (3.104)***	0.1109 (2.498)**	0.0666 (2.303)**	0.0559 (1.667)
I	0.1128 (3.331)***	0.1636 (3.558)***	0.1639 (3.798)***	0.1492 (5.207)***	0.1509 (4.974)***

Appendix Table A-23. (Continued)[1]

Explanatory Variables	Specification I[2] OLS	TSLS	REG (TSLS)[4]	Specification II[3] OLS	TSLS
PFL	-0.1017 (3.686)***	-0.1104 (3.623)***	-0.1062 (3.699)***	-0.1187 (4.766)***	-0.1151 (4.309)***
GNG_2	0.3967 (4.251)***	0.3707 (3.691)***	0.3972 (4.109)***	0.3456 (4.008)***	0.3517 (3.847)***
GNG_9	0.2971 (3.875)***	0.2605 (3.032)***	0.1661 (1.879)*	0.1520 (1.993)*	0.1149 (1.224)
GNG_{10}	0.2048 (3.663)***	0.1737 (2.705)***	0.1429 (2.313)**	0.1093 (2.178)**	0.0963 (1.731)*
Constant	-4.8420 (0.772)	1.7649 (0.205)	4.4619 (0.529)	-4.7017 (0.842)	-4.6680 (0.792)
R^2	0.937	0.934	0.950	0.941	0.934
S.E. of estimates	0.880	0.906	0.825	0.839	0.886
Exogeneity (joint) Test[5]	$F(3,39)$ = 38.359***			$F(1,45)$ = 52.814***	
S.E. of estimates	0.872			0.842	

Notes: 1. Values in parentheses are the estimated absolute t-values. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

2. Three indicators of health (H_1, H_2, and H_3) are the suspected regressors of endogeneity. The reduced-form estimates are given in Appendix Table A-25.

3. Education (E) is the suspected regressor of endogeneity. The reduced-form estimates are given in Appendix Table A-26.

4. REG (TSLS) performs the coefficient stability tests due to the university location dummy in the two-stage least squares estimates (TSLS). The null hypothesis of no coefficient instability is rejected at the 1% level: The Beggs (1988, p.97) and Dowrick (1993, p.2) tests $F(4,41) = 40.906$ and $F(22,39) = 38.239$, respectively. H_2*UD, H_3*UD, UD, and E are added to the equation. The education variable (E) is also added to this equation since improved specification often comes at the price of a radical increase in multicollinearity. The increase in multicollinearity from the addition of the education variable increases the coefficient of multiple determination (R^2) at the same time. For more details see Auster, Leveson, and Sarachek (1969, p.426). The null hypothesis of functional form misspecification is rejected: RESET2 test $t = 0.312$.

5. The exogeneity test is a variant of the Hausman test as described by Beggs (1988, p.96) applied to the suspected regressor(s) of endogeneity.

Appendix Table A-24. Reduced-form Estimates of the Pairs of Structural Equations for Three Indicators of Health and Education

	Dependent Variables[1]				
	Specification I[2]			Specification II[3]	
Explanatory Variables	No recent illness (H_1)	No chronic condition (H_2)	Education (E)	Self-assessed good health (H_3)	Education (E)
R_2	0.5079 (3.260)***	0.8750 (2.910)***	-0.0428 (0.184)	0.1086 (0.662)	0.0365 (0.173)
R_3	0.3993 (2.516)**	0.3239 (1.057)	0.0603 (0.255)	0.0473 (0.281)	0.0748 (0.345)
F	0.3224 (1.051)	-0.4364 (0.737)	-0.8023 (1.753)*	0.0589 (0.173)	-0.7553 (1.726)*
OCC_1	-0.1555 (1.621)	-0.2612 (1.411)	0.1992 (1.392)	0.0437 (0.434)	0.1350 (1.042)
OCC_2	-0.1145 (0.706)	-0.2409 (0.769)	-0.4430 (1.830)*	-0.4571 (2.624)**	-0.4765 (2.124)**
C	-0.8647 (4.632)***	-1.4792 (4.105)***	0.0434 (0.156)	-0.2711 (1.343)	-0.0552 (0.212)
B_1	1.5741 (2.361)**	0.5557 (0.432)	-0.8312 (0.836)		
B	-0.1976 (0.959)	0.3833 (0.963)	0.0268 (0.087)		
G*VIC	-1.4341 (2.696)**	-1.3936 (1.357)	-0.8947 (1.128)	0.4020 (0.721)	-1.1785 (1.642)
G*QLD	-3.3760 (2.928)***	-2.3929 (1.075)	3.7708 (2.192)**	1.7715 (1.344)	3.7750 (2.225)**
G*WA	-1.5306 (1.209)	-5.9097 (2.419)**	0.5388 (0.285)	2.4644 (1.867)*	0.3512 (0.207)
G*TAS	0.0685 (0.036)	-8.8124 (2.411)**	-4.0214 (1.424)	-1.1801 (0.571)	-4.7762 (1.796)*
G*(SA+NT)	0.2915 (0.382)	4.7477 (3.225)***	1.0216 (0.898)	-0.1415 (0.179)	1.3326 (1.310)
METD	-4.2093 (2.804)***	-1.7806 (0.614)	1.5776 (0.704)	1.9041 (1.106)	1.4137 (0.638)
NGSTD	0.0574 (0.915)	-0.0875 (0.723)	-0.0504 (0.539)	-0.0824 (1.180)	-0.0756 (0.840)
UD	0.9937 (1.236)	2.4864 (1.603)	1.0006 (0.834)	1.8831 (2.076)**	1.2536 (1.074)

Appendix Table A-24. (Continued)

	Dependent Variables[1]				
	Specification I[2]			Specification II[3]	
Explanatory Variables	No recent illness (H_1)	No chronic condition (H_2)	Education (E)	Self-assessed good health (H_3)	Education (E)
AE	-0.2037 (1.920)*	-0.0712 (0.348)	0.0015 (0.009)	0.2743 (2.815)***	0.0792 (0.631)
GNG_1	-0.0343 (0.202)	-0.5677 (1.733)*	-0.5286 (2.087)**	0.1707 (0.936)	-0.4846 (2.062)**
GNG_2	0.1072 (0.427)	0.2455 (0.507)	0.0677 (0.181)	0.5379 (1.895)*	0.1445 (0.395)
GNG_3	0.3941 (2.827)***	-0.1009 (0.375)	-0.0775 (0.372)	0.2311 (1.488)	-0.0191 (0.095)
GNG_4	0.2999 (0.676)	1.4008 (1.635)	1.1920 (1.800)*	0.3911 (0.785)	1.3444 (2.096)**
GNG_6	0.4191 (2.364)**	0.6030 (1.762)*	0.1802 (0.681)	0.5211 (2.691)**	0.1960 (0.786)
GNG_8	0.4064 (0.708)	0.8977 (0.810)	-2.5612 (2.992)***	0.0082 (0.013)	-2.5333 (3.058)***
GNG_9	-0.3150 (1.232)	-0.6131 (1.242)	0.4998 (1.311)	-0.1018 (0.375)	0.5257 (1.502)
GNG_{10}	0.0377 (0.234)	-0.2580 (0.831)	0.2788 (1.162)	0.1993 (1.164)	0.3682 (1.671)
Constant	11.6627 (0.595)	77.5364 (2.048)**	104.6634 (3.576)***	61.4150 (2.861)***	100.9848 (3.654)***
R^2	0.802	0.791	0.871	0.741	0.866
F-Stat	5.658***	5.293***	9.431***	4.600***	10.364***

Notes: **1.** Values in parentheses are the estimated absolute t-values. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

2. The reduced-form estimates of the pairs of structural equations for no recent illness (H_1), no chronic condition (H_2), and education (E) are contained in Specification I.

3. The reduced-form estimates of the pairs of structural equations for self-assessed good health (H_3) and education (E) are contained in Specification II.

Appendix Table A-25. Reduced-form Estimates of the Pairs of Structural Equations for Three Indicators of Health and Income

Explanatory Variables	Dependent Variables[1]				
	Specification I[2]			Specification II[3]	
	No recent illness (H_1)	No chronic condition (H_2)	Income (Y)	Self-assessed good health (H_3)	Income (Y)
R_2	0.5740 (3.872)***	0.8319 (3.195)***	0.0001 (0.002)	-0.0258 (0.174)	0.0203 (0.379)
R_3	0.4850 (3.119)***	0.4993 (1.828)*	0.0164 (0.257)	0.2175 (1.389)	0.0464 (0.823)
F	0.4410 (1.489)	-0.0120 (0.023)	-0.0606 (0.499)	0.7202 (2.245)**	-0.0908 (0.786)
OCC_1	-0.2366 (2.801)***	-0.2806 (1.891)*	0.0980 (2.833)***	0.0630 (0.832)	0.0841 (3.081)***
C	-0.6160 (3.576)***	-1.1121 (3.676)***	0.0390 (0.553)	-0.2227 (1.266)	0.0567 (0.894)
B_1	1.8238 (2.858)***	0.6885 (0.614)	0.2312 (0.885)		
B	-0.1230 (0.671)	0.4072 (1.265)	-0.0876 (1.168)		
G*VIC	-0.7063 (1.482)	-1.2238 (1.461)	-0.1156 (0.592)	0.0554 (0.118)	-0.0388 (0.230)
G*QLD	-2.5410 (1.776)*	1.5528 (0.618)	0.1939 (0.331)	1.5391 (0.983)	0.3377 (0.599)
G*WA	-2.3478 (1.879)*	-5.3090 (2.419)**	-0.2573 (0.503)	1.6654 (1.253)	-0.4102 (0.857)
G*TAS	0.6545 (0.338)	-8.0886 (2.378)**	-0.8443 (1.065)	1.4214 (0.661)	-0.9838 (1.269)
G*(SA+NT)	0.1237 (0.167)	3.9508 (3.039)***	0.3551 (1.172)	-1.0163 (1.267)	0.4370 (1.512)
METD	-3.3079 (2.284)**	-0.6053 (0.238)	1.2629 (2.129)**	-0.4084 (0.259)	1.3195 (2.325)**
K	0.8139 (1.457)	3.8861 (3.960)***	0.2918 (1.275)	0.6316 (1.041)	0.3636 (1.663)

Appendix Table A-25. (Continued)

Explanatory Variables	Dependent Variables[1] Specification I[2] No recent illness (H_1)	No chronic condition (H_2)	Income (Y)	Specification II[3] Self-assessed good health (H_3)	Income (Y)
Ln	-0.0038 (0.039)	-0.0064 (0.037)	0.0696 (1.747)*	0.2451 (2.526)**	0.0785 (2.246)**
I	0.2659 (2.947)***	0.1490 (0.940)	0.1593 (4.308)***	0.2992 (3.000)***	0.1506 (4.192)***
PFL	-0.1101 (1.424)	0.0676 (0.498)	-0.1144 (3.615)***	-0.0095 (0.113)	-0.1063 (3.512)***
GNG_2	-0.1548 (0.599)	-0.0363 (0.080)	0.3897 (3.681)***	-0.1866 (0.642)	0.3942 (3.767)***
GNG_8	0.0860 (0.139)	2.1565 (1.990)*	-0.0825 (0.326)	-0.0793 (0.115)	-0.0884 (0.354)
GNG_9	-0.2516 (1.170)	-0.4318 (1.142)	0.2415 (2.740)***	-0.3016 (1.304)	0.2709 (3.251)***
GNG_{10}	-0.2017 (1.406)	-0.2758 (1.095)	0.1633 (2.778)***	-0.1352 (0.852)	0.1675 (2.929)***
Constant	-29.4877 (1.333)	-25.7708 (0.663)	2.0767 (0.229)	2.0681 (0.086)	0.0676 (0.008)
R^2	0.763	0.793	0.940	0.704	0.938
F-Stat	5.993***	7.133***	29.032***	5.123***	32.521***

Notes: **1.** Values in parentheses are the estimated absolute t-values. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

2. The reduced-form estimates of the pairs of structural equations for no recent illness (H_1), no chronic condition (H_2), and income (Y) are contained in Specification I.

3. The reduced-form estimates of the pairs of structural equations for self-assessed good health (H_3) and income (Y) are contained in Specification II.

Appendix Table A-26. Reduced-form Estimates of the Pairs of Structural Equations for Education and Income

	Dependent Variables[1]			
	Education (E)		Income (Y)	
Explanatory Variables	Coefficient	t-Statistic	Coefficient	t-Statistic
R_2	-0.0451	0.187	-0.0108	0.192
R_3	-0.0219	0.096	0.0128	0.240
OCC_1	0.1256	0.864	0.1252	3.685***
OCC_2	-0.2649	0.999	0.0313	0.505
C	-0.0565	0.202	-0.0030	0.046
NGSTD	0.0657	0.670	0.0425	1.853*
UD	0.2569	0.201	-0.7463	2.503**
AE	0.2407	1.431	0.0319	0.811
METD	-3.1103	1.370	0.2824	0.533
K	0.1280	1.569	0.2324	1.383
Ln	0.2806	1.838*	0.0932	2.613**
I	0.0798	0.395	0.1994	4.227***
PFL	-0.2477	1.980*	-0.1471	5.034***
GNG_1	-0.5822	2.038**	-0.1724	2.583**
GNG_2	-0.3840	0.851	0.2497	2.368***
GNG_3	-0.2407	1.042	-0.0284	0.526
GNG_4	1.5513	2.016*	0.0273	0.152
GNG_6	-0.3778	1.085	-0.0611	0.750
GNG_9	0.3712	0.899	0.1039	1.076
GNG_{10}	-0.1679	0.570	0.0717	1.042
Constant	25.2668	0.802	-4.3627	0.593
R^2	0.814		0.947	
F-Stat	8.742***		36.039***	

Notes: 1. t-Statistic denotes the estimated absolute t-values. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Appendix Table A-27. The Ceteris Paribus Estimates When Omitting H_1 from the Education Equation, and H_2 and H_3 from the Income Equation[1]

Suspected Regressors of Endogeneity	Dependent Variables[2]							
	Education (E)				Income (Y)			
	Specification I		Specification II		Specification I		Specification II	
	OLS	TSLS	OLS	TSLS	OLS	TSLS	OLS	TSLS
H_1					0.1214	0.0805	0.1018	0.0681
					(2.579)**	(1.220)	(2.342)**	(1.076)
H_2	0.1523	0.2537	0.1047	0.1910				
	(1.467)	(2.082)**	(1.162)	(1.674)*				
H_3	0.4006	0.9612	0.0444	0.3651				
	(1.662)	(2.903)***	(0.197)	(1.002)				
E							0.0951	0.0922
							(3.199)***	(2.337)**
Y			1.4550	1.3597				
			(4.055)***	(2.984)***				
R^2	0.760	0.785	0.826	0.822	0.935	0.928	0.947	0.936
S.E. of estimates	3.963	3.747	3.410	3.450	0.877	0.922	0.801	0.881

Notes: 1. The estimated results of all other control variables are given in Appendix Table A-28.

2. Values in parentheses are the estimated absolute t-values. ***, **, and * indicate significance at the 1%, 5%, and 10% levels on a two-tailed test, respectively.

Appendix Table A.28. Estimates of the Control Variables on Education and Income

Control Variables	Dependent Variables[1]							
	Education (E)				Income (Y)			
	Specification I		Specification II		Specification I		Specification II	
	OLS	TSLS	OLS	TSLS	OLS	TSLS	OLS	TSLS
R_2	-0.2875	-0.4304	-0.1934	-0.2905	-0.0499	-0.0323	-0.0316	-0.0174
	(1.192)	(1.813)*	(0.926)	(1.300)	(0.959)	(0.558)	(0.660)	(0.313)
R_3	-0.2325	-0.2878	-0.1470	-0.1758	-0.0489	-0.0298	-0.0474	-0.0315
	(1.015)	(1.318)	(0.741)	(0.859)	(0.943)	(0.512)	(1.001)	(0.566)
OCC_1	0.2019	0.1811	0.0201	0.0181	0.1048	0.1045	0.0841	0.0844
	(1.411)	(1.335)	(0.154)	(0.132)	(4.607)***	(4.364)***	(3.861)***	(3.457)***
OCC_2	-0.3515	-0.1173	-0.2016	-0.0837				
	(1.414)	(0.459)	(0.929)	(0.356)				
C	0.1756	0.4737	0.0940	0.2949	0.1195	0.1049	0.0800	0.0689
	(0.556)	(1.430)	(0.345)	(0.949)	(2.052)**	(1.659)	(1.464)	(1.105)
NGSTD	0.1305	0.1818	0.0024	0.0451				
	(1.369)	(1.942)*	(0.027)	(0.462)				
UD	-0.8337	-1.9539	1.0266	0.2745				
	(0.596)	(1.390)	(0.798)	(0.184)				
K					0.2480	0.2896	0.0927	0.1323
					(1.486)	(1.599)	(0.579)	(0.712)
Ln					0.1133	0.1069	0.0851	0.0806
					(3.785)***	(3.317)***	(2.962)***	(2.460)**

Appendix Table A.28. (Continued)

Control Variables	Dependent Variables[1]							
	Education (E)				Income (Y)			
	Specification I		Specification II		Specification I		Specification II	
	OLS	TSLS	OLS	TSLS	OLS	TSLS	OLS	TSLS
I					0.1193	0.1279	0.1275	0.1344
					(3.790)***	(3.723)***	(4.411)***	(4.079)***
PFL					-0.1047	-0.1125	-0.1000	-1.1067
					(3.820)***	(3.749)***	(3.985)***	(3.707)***
GNG_1	-0.5153	-0.4966	-0.3150	-0.3044				
	(2.297)**	(2.320)**	(1.581)	(1.468)				
GNG_2	-0.0357	-0.3660	-0.7692	-0.8906	0.3817	0.3642	0.3877	0.3728
	(0.081)	(0.834)	(1.839)*	(2.022)**	(4.139)***	(3.687)***	(4.598)***	(3.947)***
GNG_3	-0.0891	-0.2171	-0.3292	-0.3770				
	(0.418)	(1.036)	(1.707)*	(1.883)*				
GNG_4	1.2323	0.8753	0.9235	0.7109				
	(1.669)	(1.218)	(1.443)	(1.071)				
GNG_6	-0.1700	-0.5402	-0.3898	-0.5738				
	(0.522)	(1.562)	(1.367)	(1.801)*				
GNG_9					0.2785	0.2682	0.1865	0.1807
					(3.714)***	(3.368)***	(2.509)**	(2.130)**
GNG_{10}	0.1039	0.0091	-0.3000	-0.3082	0.1927	0.1759	0.1541	0.1412
	(0.430)	(0.039)	(1.301)	(1.278)	(3.510)***	(2.912)***	(2.986)***	(2.368)**

Appendix Table A.28. (Continued)

Control Variables	Dependent Variables[1]							
	Education (E)				Income (Y)			
	Specification I		Specification II		Specification I		Specification II	
	OLS	TSLS	OLS	TSLS	OLS	TSLS	OLS	TSLS
METD					1.0823	1.0401	1.1909	1.1522
					(2.375)**	(2.161)**	(2.849)***	(2.491)**
AE	0.0904	-0.0259	0.3043	0.2388				
	(0.571)	(0.162)	(2.083)**	(1.389)				
Constant	26.3227	-13.1024	27.2451	4.0418	-2.8582	-3.5084	-3.0910	-3.6292
	(0.978)	(0.428)	(1.177)	(0.140)	(0.486)	(0.564)	(0.575)	(0.610)
R^2	0.760	0.785	0.826	0.823	0.935	0.928	0.947	0.936
S.E. of estimates	3.963	3.747	3.410	3.450	0.877	0.922	0.801	0.881

Notes: **1.** Values in parentheses are the estimated absolute t-values. ***, **, and * indicate significance at the 1%, 5%, and 10% levels on a two-tailed test, respectively.

Appendix Table A-29. Contingency Table: Cross-Classification Between No Recent Illness (H_1) and Gender (F)

Level of Gender												
Top Third				Middle Third				Bottom Third				
TH_1	MH_1	BH_1	Sub-Total	TH_1	MH_1	BH_1	Sub-Total	TH_1	MH_1	BH_1	Sub-Total	Total
0.200	0.450	0.350	1.000	0.333	0.190	0.476	1.000	0.350	0.400	0.250	1.000	3.000
0.066 (4)	0.147 (9)	0.115 (7)	0.328 (20)	0.115 (7)	0.066 (4)	0.164 (10)	0.344 (21)	0.115 (7)	0.131 (8)	0.082 (5)	0.328 (20)	1.000 (61)

Notes:
1. TH_1, MH_1 and BH_1 imply, respectively, the top third of regions, the middle third of regions, and the bottom third of regions in terms of sample proportions.
2. Values in parentheses denote the number of statistical regions.
3. For more details see Lewis, O'Brien, and Thampapillai (1990, pp.197-201).

Appendix Table A-30. Contingency Table: Cross-Classification Between No Chronic Condition (H_2) and Gender (F)

Level of Gender												
Top Third				Middle Third				Bottom Third				
TH_2	MH_2	BH_2	Sub-Total	TH_2	MH_2	BH_2	Sub-Total	TH_2	MH_2	BH_2	Sub-Total	Total
0.250	0.100	0.650	1.000	0.333	0.476	0.191	1.000	0.450	0.400	0.150	1.000	3.000
0.082 (5)	0.033 (2)	0.213 (13)	0.328 (20)	0.115 (7)	0.164 (10)	0.066 (4)	0.344 (21)	0.148 (9)	0.131 (8)	0.049 (3)	0.328 (20)	1.000 (61)

Notes: **1.** TH_2 MH_2 and BH_2 imply, respectively, the top third of regions, the middle third of regions, and the bottom third of regions in terms of sample proportions.
2. Values in parentheses denote the number of statistical regions.
3. For more details see Lewis, O'Brien, and Thampapillai (1990, pp.197-201).

Appendix Table A-31. Contingency Table: Cross-Classification Between Self-assessed Good Health (H_3) and Gender (F)

Level of Gender												
Top Third				Middle Third				Bottom Third				
TH_3	MH_3	BH_3	Sub-Total	TH_3	MH_3	BH_3	Sub-Total	TH_3	MH_3	BH_3	Sub-Total	Total
0.500	0.400	0.100	1.000	0.238	0.191	0.571	1.000	0.250	0.350	0.400	1.000	3.000
0.164 (10)	0.131 (8)	0.033 (2)	0.328 (20)	0.082 (5)	0.065 (4)	0.197 (12)	0.344 (21)	0.082 (5)	0.115 (7)	0.131 (8)	0.328 (20)	1.000 (61)

Notes: **1.** TH_3, MH_3, and BH_3 imply, respectively, the top third of regions, the middle third of regions, and the bottom third of regions in terms of sample proportions.
2. Values in parentheses denote the number of statistical regions.
3. For more details see Lewis, O'Brien, and Thampapillai (1990, pp.197-201).

Appendix Table A-32. Reduced-form Estimates on Three Indicators of Health and Gender

Explanatory Variable	Dependent Variables[1] No recent illness (H_1)	No chronic condition (H_2)	Self-assessed good health (H_3)	Gender (F) Specification I[2]	Gender (F) Specification II[3]
E	0.0972 (0.935)	0.5667 (3.106)***	0.2520 (1.964)*	-0.0189 (0.346)	-0.0134 (0.250)
Y	0.6964 (3.326)***	0.3410 (0.928)	0.5342 (2.010)*	-0.1628 (1.483)	-0.1616 (1.461)
R_2	0.6807 (5.504)***	0.7282 (3.355)***	0.0614 (0.420)	0.0284 (0.438)	0.0427 (0.704)
R_3	0.4966 (3.520)***	0.1803 (0.728)	0.2912 (1.803)*	0.0687 (0.928)	0.0594 (0.883)
OCC_1	-0.3271 (3.862)***	-0.4136 (2.782)***	-0.0166 (0.189)	0.0368 (0.829)	0.0317 (0.866)
C	-0.6473 (3.986)***	-1.4471 (5.077)***	-0.2918 (1.543)	0.1612 (1.894)*	0.1084 (1.378)
B_1	1.4235 (2.429)**	1.4035 (1.365)		-0.5024 (1.635)	
B	0.0246 (0.149)	0.2523 (0.874)		0.1078 (1.249)	
G*VIC	-0.1934 (0.420)	-0.6586 (0.814)	0.6809 (1.280)	0.2421 (1.002)	0.0779 (0.352)
G*QLD	-3.6856 (3.165)***	-4.4886 (2.196)**	1.3555 (0.926)	0.5382 (0.881)	0.4208 (0.691)
G*WA	-2.9756 (2.611)**	-5.5203 (2.760)***	1.5142 (1.086)	-0.1880 (0.315)	-0.1218 (0.210)
G*TAS	1.2336 (0.649)	-5.7930 (1.737)*	1.7800 (0.750)	-0.3911 (0.392)	-0.3276 (0.331)
G*(SA+NT)	0.1846 (0.258)	3.7384 (2.974)***	-0.9925 (1.135)	0.1537 (0.409)	0.1059 (0.291)
METD	-4.0930 (3.034)	-2.0917 (0.883)	-0.4347 (0.259)	1.0396 (1.470)	0.8248 (1.179)
GNG_8	0.2785 (0.474)	2.5558 (2.477)**	0.2108 (0.284)	-0.6223 (2.019)**	-0.5631 (1.826)*

Appendix Table A-32. (Continued)

Explanatory Variable	Dependent Variables[1]				
	No recent illness (H_1)	No chronic condition (H_2)	Self-assessed good health (H_3)	Gender (F)	
				Specification I[2]	Specification II[3]
GNG_9	-0.2862 (1.426)	-1.1141 (3.162)***	-0.2988 (1.191)	0.1699 (1.614)	0.1397 (1.338)
GNG_{10}	-0.2440 (1.702)*	-0.5232 (2.079)**	-0.2072 (1.145)	0.0470 (0.626)	0.0591 (0.784)
Constant	-5.7773 (0.435)	25.2844 (1.086)	40.8433 (2.561)**	44.0472 (6.332)***	45.8753 (6.912)***
R^2	0.741	0.775	0.592	0.408	0.371
F-Stat	7.253***	8.690***	4.349***	1.744*	1.772*

Notes: **1.** Values in parentheses are the estimated absolute t-values. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

2. The reduced-form estimates from the exogeneity test of the gender variable in the no recent illness (H_1) and no chronic condition (H_2) equations.

3. The reduced-form estimates from the exogeneity test of the gender variable in the self-assessed good health (H_3) equation.

Appendix Table A-33 Contingency Table: Cross-Classification Between Education (E) and Skill Level (OCC_1)

Level of Skill												
Top Third				Middle Third				Bottom Third				
TE	ME	BE	Sub-Total	TE	ME	BE	Sub-Total	TE	ME	BE	Sub-Total	Total
0.571	0.191	0.238	1.000	0.286	0.381	0.333	1.000	0.105	0.474	0.421	1.000	3.000
0.197 (12)	0.065 (4)	0.082 (5)	0.344 (21)	0.098 (6)	0.131 (8)	0.115 (7)	0.344 (21)	0.033 (2)	0.148 (9)	0.131 (8)	0.312 (19)	1.000 (61)

Notes: **1.** TE, ME,and BE imply, respectively, the top third of regions, the middle third of regions, and the bottom third of regions in terms of sample proportions.
2. Values in parentheses denote the number of statistical regions.
3. For more details see Lewis, O'Brien, and Thampapillai (1990, pp.197-201).

Appendix Table A-34. Contingency Table: Cross-Classification Between Income (Y) and Skill Level (OCC_1)

Level of Skill												
Top Third				Middle Third				Bottom Third				
TY	MY	BY	Sub-Total	TY	MY	BY	Sub-Total	TY	MY	BY	Sub-Total	Total
0.619	0.143	0.238	1.000	0.333	0.238	0.429	1.000	0.053	0.648	0.263	1.000	3.000
0.213 (13)	0.049 (3)	0.082 (5)	0.344 (21)	0.115 (7)	0.082 (5)	0.147 (9)	0.344 (21)	0.016 (1)	0.213 (13)	0.082 (5)	0.311 (19)	1.000 (61)

Notes: **1.** TY, MY and BY imply, respectively, the top third of regions, the middle third of regions, and the bottom third of regions in terms of sample proportions.
2. Values in parentheses denote the number of statistical regions.
3. For more details see Lewis, O'Brien, and Thampapillai (1990, pp.197-201).

Appendix Table A-35. Reduced-form Estimates on Education, Income, and Skill Level

Explanatory Variable	Dependent Variables[1] Education (E)	Income (Y)	Skill level (OCC_1)
H_1	0.1265 (0.465)	0.0884 (1.252)	-0.1060 (0.354)
H_2	0.1194 (0.829)	0.0289 (0.773)	0.0380 (0.240)
H_3	0.1808 (0.687)	0.0463 (0.678)	0.1812 (0.626)
R_2	-0.2093 (0.802)	-0.1004 (1.483)	-0.2726 (0.950)
R_3	-0.1428 (0.534)	-0.0673 (0.971)	-0.2125 (0.723)
OCC_2	-0.3066 (1.300)	-0.0734 (1.200)	-0.9725 (3.748)***
C	0.1347 (0.416)	0.0287 (0.341)	-0.5096 (1.430)
NGSTD	0.0873 (0.896)	0.0629 (0.491)**	0.2323 (2.168)**
UD	0.1687 (0.128)	-0.5807 (1.701)*	1.8679 (1.290)
AE	0.1998 (1.009)	0.0164 (0.320)	-0.2292 (1.053)
METD	-2.3793 (0.989)	0.5214 (0.835)	-1.5995 (0.604)
K	0.8052 (1.047)	0.1432 (0.718)	0.3622 (0.428)
Ln	0.3081 (2.031)	0.1309 (3.328)***	0.2593 (1.554)
I	-0.0483 (0.229)	0.1207 (2.203)**	-0.5137 (2.211)**
PFL	-0.2010 (1.520)	-0.1420 (4.139)***	-0.1302 (0.895)

Appendix Table A-35. (Continued)

Explanatory Variable	Dependent Variables[1]		
	Education (E)	Income (Y)	Skill level (OCC_1)
GNG_1	-0.3709 (1.330)	-0.0382 (0.527)	0.7701 (2.510)**
GNG_2	-0.2834 (0.608)	0.2890 (2.389)**	-0.0779 (0.152)
GNG_3	-0.2551 (1.037)	-0.0426 (0.668)	0.0870 (0.321)
GNG_4	1.5386 (2.015)*	0.1813 (0.916)	1.8526 (2.206)**
GNG_6	-0.4802 (1.360)	-0.1340 (1.463)	-0.3748 (0.965)
GNG_9	0.5694 (1.287)	0.1607 (1.400)	-0.1310 (0.269)
GNG_{10}	-0.0271 (0.085)	0.0938 (1.131)	-0.3314 (0.942)
Constant	26.1349 (0856)	5.8844 (0.743)	98.2684 (2.927)***
R^2	0.820	0.937	0.804
F-Stat	7.889***	25.906***	7.069***

Notes: 1. Values in parentheses are the estimated absolute t-values. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Appendix Table A-36. Contingency Table: Cross-Classification Between Moderate Drinking (R_3) and No Recent Illness (H_1)

Level of Moderate Drinking												
Top Third				Middle Third				Bottom Third				
TH_1	MH_1	BH_1	Sub-Total	TH_1	MH_1	BH_1	Sub-Total	TH_1	MH_1	BH_1	Sub-Total	Total
0.158	0.368	0.474	1.000	0.136	0.455	0.409	1.000	0.600	0.200	0.200	1.000	3.000
0.049 (3)	0.115 (7)	0.148 (9)	0.311 (19)	0.049 (3)	0.164 (10)	0.148 (9)	0.361 (22)	0197 (12)	0.066 (4)	0.066 (4)	0.328 (20)	1.000 (61)

Notes: **1.** TH_1, MH_1 and BH_1 imply, respectively, the top third of regions, the middle third of regions, and the bottom third of regions in terms of sample proportions.
2. Values in parentheses denote the number of statistical regions.
3. For more details see Lewis, O'Brien, and Thampapillai (1990, pp.197-201).

Appendix Table A-37. Contingency Table: Cross-Classification Between Moderate Drinking (R_3) and Self-assessed Good Health (H_3)

Level of Moderate Drinking												
Top Third				Middle Third				Bottom Third				
TH_3	MH_3	BH_3	Sub-Total	TH_3	MH_3	BH_3	Sub-Total	TH_3	MH_3	BH_3	Sub-Total	Total
0.631	0.211	0.158	1.000	0.227	0.455	0.318	1.000	0.150	0.250	0.600	1.000	3.000
0.197 (12)	0.065 (4)	0.049 (3)	0.311 (19)	0.082 (5)	0.164 (10)	0.115 (7)	0.361 (22)	0.049 (3)	0.082 (5)	0.197 (12)	0.328 (20)	1.000 (61)

Notes: **1.** TH_3, MH_3,and BH_3 imply, respectively, the top third of regions, the middle third of regions, and the bottom third of regions in terms of sample proportions.
2. Values in parentheses denote the number of statistical regions.
3. For more details see Lewis, O'Brien, and Thampapillai (1990, pp.197-201).

Appendix Table A-38. Contingency Table: Cross-Classification Between Moderate Drinking (R_3) and No Chronic Condition (H_2)

Level of Moderate Drinking												
Top Third				Middle Third				Bottom Third				
TH_2	MH_2	BH_2	Sub-Total	TH_2	MH_2	BH_2	Sub-Total	TH_2	MH_2	BH_2	Sub-Total	Total
0.053	0.421	0.526	1.000	0.273	0.409	0.318	1.000	0.700	0.150	0.150	1.000	3.000
0.016 (1)	0.131 (8)	0.164 (10)	0.311 (19)	0.098 (6)	0.148 (9)	0.115 (7)	0.361 (22)	0.230 (14)	0.049 (3)	0.049 (3)	0.328 (20)	1.000 (61)

Notes: **1.** The null hypothesis of independence between the two variables is rejected at the 0.01 level of significance. The calculated values of Chi-Square of Pearson (4), Likelihood ratio (4), and Mantel-Haenszel (1) with the number of degrees of freedom in parentheses are 19.467, 20.968, and 14.997, respectively. The contingency table implies that two variables are negatively dependent on one another.

2. TH_2, MH_2, and BH_2 imply, respectively, the top third of regions, the middle third of regions, and the bottom third of regions in terms of sample proportions.

3. Values in parentheses denote the number of statistical regions.

4. For more details see Lewis, O'Brien, and Thampapillai (1990, pp.197-201).

Appendix Table A-39. Contingency Table: Cross-Classification Between Smoking (R_1) and Excessive Drinking (R_4)

Level of Smoking												
Top Third				Middle Third				Bottom Third				
TR_4	MR_4	BR_4	Sub-Total	TR_4	MR_4	BR_4	Sub-Total	TR_4	MR_4	BR_4	Sub-Total	Total
0.550	0.250	0.200	1.000	0.200	0.450	0.350	1.000	0.286	0.286	0.428	1.000	3.000
0.180 (11)	0.082 (5)	0.066 (4)	0.328 (20)	0.066 (4)	0.148 (9)	0.115 (7)	0.328 (20)	0.098 (6)	0.098 (6)	0.148 (9)	0.344 (21)	1.000 (61)

Notes: **1.** TR_4, MR_4 and BR_4 imply, respectively, the top third of regions, the middle third of regions, and the bottom third of regions in terms of sample proportions.
2. Values in parentheses denote the number of statistical regions.
3. For more details see Lewis, O'Brien, and Thampapillai (1990, pp.197-201).

Appendix Table T-40. Contingency Table: Cross-Classification Between Moderate Drinking (R_3) and Smoking (R_1)

Level of Moderate Drinking												
Top Third				Middle Third				Bottom Third				
TR_1	MR_1	BR_1	Sub-Total	TR_1	MR_1	BR_1	Sub-Total	TR_1	MR_1	BR_1	Sub-Total	Total
0.105	0.368	0.527	1.000	0.364	0.272	0.364	1.000	0.500	0.350	0.150	1.000	3.000
0.032 (2)	0.115 (7)	0.164 (10)	0.311 (19)	0.131 (8)	0.099 (6)	0.131 (8)	0.361 (22)	0.164 (10)	0.115 (7)	0.049 (3)	0.328 (20)	1.000 (61)

Notes:

1. The null hypothesis of independence between the two variables is rejected at the 0.10 level of significance. The calculated values of Chi-Square of Pearson (4), Likelihood ratio (4), and Mantel-Haenszel (1) with the number of degrees of freedom in parentheses are 9.140, 10.271, and 8.468, respectively. The contingency table implies that two variables are negatively dependent on one another.
2. TR_1, MR_1, and BR_1 imply, respectively, the top third of regions, the middle third of regions, and the bottom third of regions in terms of sample proportions.
3. Values in parentheses denote the number of statistical regions.
4. For more details see Lewis, O'Brien, and Thampapillai (1990, pp.197-201).

Appendix Table A-41. Contingency Table: Cross-Classification Between No Recent Illness (H_1) and Smoking (R_1)

Level of Smoking												
Top Third				Middle Third				Bottom Third				
TH_1	MH_1	BH_1	Sub-Total	TH_1	MH_1	BH_1	Sub-Total	TH_1	MH_1	BH_1	Sub-Total	Total
0.450	0.350	0.200	1.000	0.300	0.400	0.300	1.000	0.143	0.286	0.571	1.000	3.000
0.148 (9)	0.115 (7)	0.065 (4)	0.328 (20)	0.098 (6)	0.132 (8)	0.098 (6)	0.328 (20)	0.049 (3)	0.098 (6)	0.197 (12)	0.344 (21)	1.000 (61)

Notes:
1. The null hypothesis of independence between the two variables is rejected at the 0.10 level of significance. The calculated values of Chi-Square of Person (4), Likelihood ratio (4), and Mantel-Haenszel (1) with the number of degrees of freedom in parentheses are 7.891, 7.956, and 7.153, respectively. The contingency table implies that two variables are positively dependent on one another.
2. TH_1, MH_1, and BH_1 imply, respectively, the top third of regions, the middle third of regions, and the bottom third of regions in terms of sample proportions.
3. Values in parentheses denote the number of statistical regions.
4. For more details see Lewis, O'Brien, and Thampapillai (1990, pp.197-201).

Appendix Table A-42. Contingency Table: Cross-Classification Between No Chronic Condition (H_2) and Smoking (R_1)

Level of Smoking												
Top Third				Middle Third				Bottom Third				
TH_2	MH_2	BH_2	Sub-Total	TH_2	MH_2	BH_2	Sub-Total	TH_2	MH_2	BH_2	Sub-Total	Total
0.600	0.200	0.200	1.000	0.300	0.350	0.350	1.000	0.142	0.429	0.429	1.000	3.000
0.196 (12)	0.066 (4)	0.066 (4)	0.328 (20)	0.098 (6)	0.115 (7)	0.115 (7)	0.328 (20)	0.048 (3)	0.148 (9)	0.148 (9)	0.344 (21)	1.000 (61)

Notes:

1. The null hypothesis of independence between the two variables is rejected at the 0.05 level of significance. The calculated values of Chi-Square of Pearson (4), Likelihood ratio (4), and Mantel-Haenszel (1) with the number of degrees of freedom in parentheses are 9.741, 9.966, and 7.018, respectively. The contingency table implies that two variables are positively dependent on one another.
2. TH_2, MH_2, and BH_2 imply, respectively, the top third of regions, the middle third of regions, and the bottom third of regions in terms of sample proportions.
3. Values in parentheses denote the number of statistical regions.
4. For more details see Lewis, O'Brien, and Thampapillai (1990, pp.197-201).

Appendix Table A-43. Contingency Table: Cross-Classification Between Self-assessed Good Health (H_3) and Smoking (R_1)

Level of Smoking												
Top Third				Middle Third				Bottom Third				
TH_3	MH_3	BH_3	Sub-Total	TH_3	MH_3	BH_3	Sub-Total	TH_3	MH_3	BH_3	Sub-Total	Total
0.150	0.350	0.500	1.000	0.350	0.300	0.350	1.000	0.476	0.286	0.238	1.000	3.000
0.049 (3)	0.115 (7)	0.164 (10)	0.328 (20)	0.115 (7)	0.098 (6)	0.115 (7)	0.328 (20)	0.164 (10)	0.098 (6)	0.082 (5)	0.344 (21)	1.000 (61)

Notes: **1.** The null hypothesis of independence between the two variables is rejected at the 0.25 level of significance. The calculated values of Chi-Square of Pearson (4), Likelihood ratio (4), and Mantel-Haenszel (1) with the number of degrees of freedom in parentheses are 5.475, 5.794, and 5.055, respectively. The contingency table implies that two variables are negatively dependent on one another.

2. TH_3, MH_3, and BH_3 imply, respectively, the top third of regions, the middle third of regions, and the bottom third of regions in terms of sample proportions.

3. Values in parentheses denote the number of statistical regions.

4. For more details see Lewis, O'Brien, and Thampapillai (1990, pp.197-201).

Appendix Table A-44. Reduced-form Estimates of the Pairs of Structural Equations for No Recent Illness, Self-assessed Good Health, and Moderate Drinking[1]

Explanatory Variables	Dependent Variables[2] No recent illness (H_1) Coefficient	t-Statistic	Self-assessed good health (H_3) Coefficient	t-Statistic
E	-0.0960	0.945	0.1059	0.780
Y	0.4892	2.227**	0.3753	1.279
H_2	0.2407	3.033***	0.1258	1.186
R_1	-0.0942	0.878	-0.4148	2.894***
R_2	-0.0973	0.924	-0.3971	2.820***
R_4	-0.5004	4.023***	-0.0246	0.148
AGE_3	0.8374	2.918***	0.9967	2.599**
AGE_4	0.4099	1.364	0.6883	1.714*
AGE_5	0.3717	1.167	-0.1264	0.297
$FAGE_3$	-0.5412	2.045**	-0.4985	1.410
$FAGE_4$	0.1967	0.722	-0.2353	0.646
$FAGE_5$	-0.1142	0.373	0.1274	0.311
F	0.7023	2.766***	0.7286	2.147**
OCC_1	-0.2471	2.373**	0.0911	0.655
G*VIC	-0.5668	1.385	0.4561	0.834
G*QLD	-1.3642	1.116	2.6712	1.708*
G*WA	-2.0917	1.913*	2.1373	1.462
G*TAS	1.0906	0.639	1.4056	0.616
G*(SA+NT)	-0.1523	0.220	-1.1074	1.195
C	-0.4508	2.461**	-0.1292	0.528
B_1	1.4646	2.750***	-1.0013	1.407
B	-0.1441	0.848	-0.1055	0.464
METD	-2.9342	2.316**	1.3156	0.777
GNG_8	-0.8821	1.448	-0.6090	0.748
GNG_9	-0.3856	1.836*	-0.6042	2.152**
GNG_{10}	-0.2093	1.662	-0.2862	1.701*
Constant	-4.0081	0.233	31.3307	1.362
R^2	0.864		0.773	
F-Stat	8.318***		4.450***	

Notes: 1. The sum of R_2 (no drinking), R_3 (moderate drinking), and R_4 (excessive drinking) is 100 per cent and the observed coefficient of multiple determination (R^2) in the reduced-form moderate drinking equation is 1.000. Therefore, the estimates of the reduced-form moderate drinking parameters are not reported here.

2. t-Statistic denotes the estimated absolute t-values. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Appendix Table A-45. Reduced-form Estimates of the Pairs of Structural Equations for Smoking and Excessive Drinking

Explanatory Variables	Dependent Variables[1]			
	Smoking (R_1)		Excessive drinking (R_4)	
	Coefficient	t-Statistic	Coefficient	t-Statistic
E	0.0482	0.297	0.0168	0.145
Y	1.1402	3.216***	0.6353	2.515**
AGE_8	0.1903	0.990	0.4665	3.407***
R_3	-0.2507	2.158**	-0.3357	4.056***
NES	-0.0471	0.425	-0.3026	3.834***
Ln	-0.3636	2.726***	-0.0241	0.253
OCC_1	-0.2931	2.269**	-0.0444	0.482
GNG_1	-0.0924	0.485	-0.0729	0.537
GNG_3	-0.3025	1.647	-0.1854	1.417
GNG_4	-0.2503	0.423	-0.4061	0.963
GNG_5	0.1428	0.346	0.3409	1.158
GNG_6	-0.0318	0.111	-0.1614	0.792
GNG_7	-0.8455	1.935*	0.4152	1.334
GNG_8	-0.1359	0.159	0.1965	0.322
GNG_9	-0.4567	1.407	0.0388	0.168
GNG_{11}	-0.0714	0.281	0.1043	0.576
Constant	39.1926	2.860***	-2.5878	0.265
R^2	0.506		0.579	
F-Stat	2.818***		3.778***	

Notes: 1. t-Statistic denotes the estimated absolute t-values. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.