Health Status and Labor Supply: Interrelationship and Determinants*

Siu Fai Leung
and
Chi Tat Wong

Hong Kong University of Science and Technology

This Version: May 28, 2002

Abstract

This paper investigates empirically the interrelationship between health status and labor supply as well as their respective determinants. By exploiting the logical consistency conditions, we provide a unifying framework to develop three different simultaneous equations models to capture various possible forms of interdependence between health and labor supply. In contrast to the ad hoc models in the literature, our models are derived in a coherent and systematic way. We estimate the models using a large cross-section data set obtained from a recent survey of the Hong Kong population. Our results indicate that separate estimation of the health and labor supply equations will produce misleading results, and that the two equations should be jointly estimated. We find strong evidence that health status is a significant determinant of employment, but not vice versa.

Keywords: health status, self-reported health, labor supply, employment, simultaneous equations model, limited dependent variable, logical consistency

Journal of Economic Literature Classification System: C350, I120, J200

*We wish to thank Teh-wei Hu for kindly granting us access to the data set and for giving us many useful comments on this study. Special thanks go to Chunrong Ai, Takeshi Amemiya, Songnian Chen, Randall P. Ellis, Cheng Hsiao, Francis T. Lui, Albert C. Ma, Walter Oi, and Adrian Pagan for very helpful comments and suggestions. We also thank for Shihti Yu for his help on the computer programming of the models used in this study.

Correspondence to: Siu Fai Leung, Department of Economics, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong. Email: `sfleung@ust.hk`

1 Introduction

The objective of this paper is to investigate empirically the interrelationship between health and labor supply as well as their respective determinants. While there is a large literature on the effect of health on labor supply (see Currie and Madrian (1999) for a survey), there are very few studies on the effect of labor supply on health in a fully simultaneous equations framework. In this paper, we develop three different simultaneous equations models to capture various possible forms of interdependence between health and labor supply. We also estimate the models using a large cross-sectional data set obtained from a recent survey of the Hong Kong population.

Most of the studies on health and labor supply focus on the impact of health on retirement. Using a "bivariate" logit model, Anderson and Burkhauser (1984, 1985) find that there is a significant negative interaction between health and retirement. Based on the estimates of their hazard models, Diamond and Hausman (1984) and Hausman and Wise (1985) find that health and retirement are negatively related. Sickles and Taubman (1986) formulate a random effect model to estimate a structural model of health and retirement. Their results indicate that improvements in health reduce the probability of retirement. In a study of disability and labor force participation (LFP), Stern (1989) finds that disability lowers the probability of LFP, but LFP increases the probability of disability. Although the finding on the negative effect of health on retirement is quite well established, the explanation is still unclear. For example, a number of authors argue that some healthy people choose to retire early and report their health conditions to be poor in order to take advantage of the social security system (see, e.g., Parsons (1982), Anderson and Burkhauser (1984, 1985), and Sickles and Taubman (1986)). This disincentive effect has been labeled the self-justification hypothesis in the literature.

While most studies find that health has a significant impact on retirement, the finding on the effect of employment on health is not as unambiguous. Ekerdt et. al. (1983) discuss the ambiguity concerning whether work improves or deteriorates health. Self-esteem, identity, and personal fulfillment from supplying labor efforts improve health. However, work pressure or poor working environment worsens health. Using multiple regression analysis, they find that retirement does not affect health. In contrast, using separate estimation, Ross and Mirowsky (1995) find that health is protected by employment and improvements in health increase the probability of employment.

A naive way to study the relationship between health and labor supply is to put employment as an explanatory variable in the health equation and vice versa,[1] and estimate the two equations separately without taking into account any possible interdependence between the two equations. This naive method ignores the possible endogeneity of the employment variable in the health equation, and the possible endogeneity of the health variable in the employment equation. Instead of assuming away the endogeneity problem, we provide three different models to test for endogeneity.

Our paper is different from the literature in three distinct ways. First, in contrast to the literature in which models are chosen in an ad hoc way,[2] we formulate a general structural model and exploit the logical consistency conditions to derive systematically

[1] In this paper, we use the term labor supply and employment interchangeably.

[2] For example, the model in Sickles and Taubman (1986) is partly structural and partly reduced form, whereas Stern's model is reduced form. They do not offer any explanation for their choice of models.

various interdependent models of health and labor supply. The unifying framework reveals transparently and coherently the linkages between the general model and the ad hoc models as well as their variants. It also demonstrates that the endogeneity problem can be tackled by means of a simultaneous equations model without imposing any exclusion restrictions on the explanatory variables. Second, in contrast to the literature in which typically only one model is estimated, we estimate a full range of models and compare the results. We examine three simultaneous equations models and one of them has never been considered in the health and labor supply literature. Third, we consider the entire adult population, instead of only focusing on the retired elderly. In a recent survey of the literature, Currie and Madrian (1999, p.3353) conclude that "a glaring limitation of the existing literature is the intense focus on elderly white men, to the virtual exclusion of most other groups." We believe that it is worthwhile to study the interrelationship between health and labor supply for the broader adult population.

The plan of the rest of the paper is as follows. Section 2 discusses how heath and labor supply are measured in this study. We present three simultaneous equations models and provide a unifying framework to develop the econometric models. We outline the estimation strategy for the models and discuss the computational problems. Section 3 describes the variables and discusses the estimation results. Section 4 concludes the paper.

2 Econometric Models

2.1 *Measurement of Health Status and Labor Supply*

Our measure of a person's health status is obtained from his/her response to the following question in the survey: "Compared with people of your age, do you consider your health condition better, worse, or more or less the same?" In reply, the respondent can choose one of the following five choices: much worse, worse, more or less the same, better, much better. Clearly, this is a self-reported subjective measure of health status. This type of question is very common in health surveys. There are many concerns about using self-reported health as a measure of health condition. First, different people may have different perceptions about what it means by "better" in answering the above question. A "much worse" for one may be perceived as "much better" by another. Second, the answer may be sensitive to short term events. For example, someone who was sick at the time of the survey might report his/her health status to be much worse than others. Despite these and other concerns, self-reported health is still the most popular measure of health available. A large number of studies use self-assessed health as a measure of the "true" health status.[3] In addition, many studies show that it is a good measure of health. Graur, West, and Gregory (1988) find that self-reported health helps to forecast a person's future health outcomes (such as morbidity and mortality). Taubman and Rosen (1982) find that self-reported health is close to "objective" health. From their unique data set, they find that, between 1969 and 1973, 23% of those in worse health in 1969 died, while only 7% of those who were in better health died. They claim that

[3] See, e.g., Lee (1982), Anderson and Burkhauser (1984, 1985), Sickles and Taubman (1986), Stern (1989), and Kenkel (1995).

self-reported health closely resembles the "true" health status. The survey conducted by Idler and Benyamini (1997) shows that in 23 out of 27 studies across 10 countries from 1982 to 1996, self-reported health ratings "reliably predict survival in populations even when known health risk factors have been accounted for". These studies lend credence to the usefulness of self-reported health.

Let y_1 be the self-reported health status, measured in the form of a polychotomous variable with five possible values $\{0, 1, 2, 3, 4\}$. The five values of y_1 correspond to the five answers to the health status question: much worse, worse, more or less the same, better, much better. Let y_1^* denote the unobserved health. We assume that the relationship between the observed and unobserved variables is given by

$$y_1 = \begin{cases} 4 & \text{if} \quad \alpha_3 < y_1^* < \alpha_4 \\ 3 & \text{if} \quad \alpha_2 < y_1^* \leq \alpha_3 \\ 2 & \text{if} \quad \alpha_1 < y_1^* \leq \alpha_2 \\ 1 & \text{if} \quad \alpha_0 < y_1^* \leq \alpha_1 \\ 0 & \text{if} \quad \alpha_{-1} < y_1^* \leq \alpha_0 \end{cases}, \tag{1}$$

where α_m $(m = -1, 0, 1, 2, 3, 4)$ is a threshold parameter, and $\alpha_m < \alpha_{m+1}$. Assume $\alpha_{-1} = -\infty$, $\alpha_4 = \infty$, and, without loss of generality, set $\alpha_0 = 0$.

Let y_2 denote the employment status, measured in the form of a dichotomous variable with two possible values $\{0, 1\}$. The two values of y_2 correspond to the respondent's economic activity status: 0 if unemployed and 1 if employed. Let y_2^* be the propensity to work and assume that

$$y_2 = \begin{cases} 1 & \text{if} \quad y_2^* > 0 \\ 0 & \text{if} \quad y_2^* \leq 0 \end{cases}. \tag{2}$$

2.2 *Model Specification*

We consider three simultaneous equations models.

Model I:

$$y_1^* = x\pi_1 + \varepsilon_1 \tag{3}$$

$$y_2^* = x_2\beta_2 + y_1\gamma_2 + \varepsilon_2 \tag{4}$$

where x is an $n \times k$ matrix for all the explanatory variables, π_1 is a $k \times 1$ vector of parameters, x_2 is an $n \times k_2$ matrix of explanatory variables for the employment equation, β_2 is a $k_2 \times 1$ vector of parameters, and γ_2 is a scalar parameter. (3) is a reduced form equation and (4) is a structural equation.

Model II:

$$y_1^* = x_1\beta_1 + y_2\gamma_1 + \varepsilon_1 \tag{5}$$

$$y_2^* = x\pi_2 + \varepsilon_2 \tag{6}$$

where x_1 is an $n \times k_1$ matrix of explanatory variables for the health equation, β_1 is a $k_1 \times 1$ vector of parameters, and γ_1 is a scalar parameter. Model II is a mirror image of Model I. Here (5) is a structural equation and (6) is a reduced form equation.

Model III:

$$y_1^* = x_1\beta_1 + y_2^*\theta_1 + \varepsilon_1 \tag{7}$$

$$y_2^* = x_2\beta_2 + y_1^*\theta_2 + \varepsilon_2 \tag{8}$$

where θ_1 and θ_2 are scalar parameters. Both (7) and (8) are structural equations.

Model I is similar to the one in Sickles and Taubman (1986), but their model has a more complicated error structure because it is formulated for a panel data set. Model II, where employment is explicitly treated as an explanatory variable in the health equation, has never been considered in the literature. Model III is similar to the one in Stern (1989), but his model is simpler than ours because his model is a standard bivariate probit model, whereas our model is composed of an ordered probit and a standard probit.

Models I and II are triangular systems, whereas Model III is a simultaneous system. We choose these three models because Models I and III have been used in the literature, while Model II is a natural extension of Model I. These three models appear to be ad hoc and no justification on the specification of Models I and III has been offered in the literature. In the following, we offer a unifying framework to obtain the models in a coherent and systematic way. We formulate a general model and exploit the logical consistency conditions to derive Models I through III as well as other variants.

Consider the general model:

$$y_1^* = x_1\beta_1 + y_1\delta_1 + y_2\gamma_1 + y_2^*\theta_1 + u_1 \tag{9}$$

$$y_2^* = x_2\beta_2 + y_1\gamma_2 + y_2\delta_2 + y_1^*\theta_2 + u_2 \tag{10}$$

where x_1 and x_2 are explanatory variables (x_1 and x_2 can contain some common variables), β_k, δ_k, γ_k, and θ_k are parameters ($k = 1, 2$), u_1 and u_2 are error terms. Substitute (10) into (9),

$$\begin{aligned} y_1^* &= x_1\beta_1 + y_1\delta_1 + y_2\gamma_1 + (x_2\beta_2 + y_1\gamma_2 + y_2\delta_2 + y_1^*\theta_2 + u_2)\theta_1 + u_1 \\ &= x_1\beta_1 + y_1\delta_1 + y_2\gamma_1 + x_2\beta_2\theta_1 + y_1\gamma_2\theta_1 + y_2\delta_2\theta_1 + y_1^*\theta_2\theta_1 + u_2\theta_1 + u_1 \end{aligned} \tag{11}$$

Assuming $\theta_1\theta_2 \neq 1$, it follows that

$$y_1^* = \frac{1}{1-\theta_1\theta_2}\left[y_1(\delta_1 + \gamma_2\theta_1) + y_2(\gamma_1 + \delta_2\theta_1) + x_1\beta_1 + x_2\beta_2\theta_1 + u_1 + u_2\theta_1\right], \tag{12}$$

Similarly, substituting (9) into (10), and solving for y_2^*, we obtain

$$y_2^* = \frac{1}{1-\theta_1\theta_2}\left[y_1(\gamma_2 + \delta_1\theta_2) + y_2(\delta_2 + \gamma_1\theta_2) + x_1\beta_1\theta_2 + x_2\beta_2 + u_1\theta_2 + u_2\right]. \tag{13}$$

The following proposition states the conditions for the model (9) and (10) to be logically consistent.

Proposition 1 For (9) and (10) to be logically consistent, either (C1) or (C2), or both, must hold:

$$\delta_1 + \gamma_2\theta_1 = 0,\ \ \delta_2 = 0,\ \text{ and } \gamma_1 = 0 \tag{C1}$$

$$\delta_2 + \gamma_1\theta_2 = 0, \ \delta_1 = 0, \ \text{and } \gamma_2 = 0 \tag{C2}$$

Proof: From Proposition 2 in Appendix I, the logical consistency of (9) and (10) requires that either (C3) or (C4), or both, must hold:

$$\delta_1 + \gamma_2\theta_1 = 0, \ \delta_2 + \gamma_1\theta_2 = 0, \ \text{and } \gamma_1 + \delta_2\theta_1 = 0 \tag{C3}$$

$$\delta_1 + \gamma_2\theta_1 = 0, \ \delta_2 + \gamma_1\theta_2 = 0, \ \text{and } \gamma_2 + \delta_1\theta_2 = 0 \tag{C4}$$

First, consider (C3). The third equality in (C3) implies that $\gamma_1 = -\delta_2\theta_1$. Substituting this into the second equality in (C3), $\delta_2 - \delta_2\theta_1\theta_2 = 0$, or $\delta_2(1 - \theta_1\theta_2) = 0$. Since $\theta_1\theta_2 \neq 1$, this implies that $\delta_2 = 0$. It follows from the third equality in (C3) that $\gamma_1 = 0$. Thus, the second and third equalities in (C3) imply that $\delta_2 = 0$ and $\gamma_1 = 0$, hence (C1) follows. Now consider (C4). The first and third inequalities in (C4) imply that $\delta_1 = 0$ and $\gamma_2 = 0$, hence (C2) follows from (C4).

If (C1) holds, then (12) and (13) become

$$y_1^* = \frac{1}{1 - \theta_1\theta_2}\left(x_1\beta_1 + x_2\beta_2\theta_1 + u_1 + u_2\theta_1\right) \tag{14}$$

$$y_2^* = \frac{1}{1 - \theta_1\theta_2}\left[x_1\beta_1\theta_2 + x_2\beta_2 + y_1(\gamma_2 + \delta_1\theta_2) + u_2 + u_1\theta_2\right] \tag{15}$$

If $\theta_2 = 0$, we have

$$y_1^* = x_1\beta_1 + x_2\beta_2\theta_1 + u_1 + u_2\theta_1 \tag{16}$$

$$y_2^* = x_2\beta_2 + y_1\gamma_2 + u_2 \tag{17}$$

which is our Model I. In addition, if $\theta_1 = 0$, then

$$y_1^* = x_1\beta_1 + u_1 \tag{18}$$

$$y_2^* = x_2\beta_2 + y_1\gamma_2 + u_2 \tag{19}$$

which is a special case of Model I where the health equation does not contain all the explanatory variables. This is also the actual model estimated in Sickles and Taubman (1986).

If $\gamma_2 = 0$, then (C1) implies that $\delta_1 = 0$, thus (14) and (15) become

$$y_1^* = \frac{1}{1 - \theta_1\theta_2}(x_1\beta_1 + x_2\beta_2\theta_1 + u_1 + u_2\theta_1) \tag{20}$$

$$y_2^* = \frac{1}{1 - \theta_1\theta_2}(x_1\beta_1\theta_2 + x_2\beta_2 + u_1\theta_2 + u_2) \tag{21}$$

which is the same as (7) and (8) of Model III.

If (C2) holds, then (12) and (13) become

$$y_1^* = \frac{1}{1-\theta_1\theta_2}\left[x_1\beta_1 + x_2\beta_2\theta_1 + y_2(\gamma_1 + \delta_2\theta_1) + u_1 + u_2\theta_1\right] \tag{22}$$

$$y_2^* = \frac{1}{1-\theta_1\theta_2}\left(x_1\beta_1\theta_2 + x_2\beta_2 + u_1\theta_2 + u_2\right) \tag{23}$$

If $\theta_1 = 0$, then (22) and (23) become

$$y_1^* = x_1\beta_1 + y_2\gamma_1 + u_1 \tag{24}$$

$$y_2^* = x_1\beta_1\theta_2 + x_2\beta_2 + u_1\theta_2 + u_2 \tag{25}$$

which is our Model II. In addition, if $\theta_2 = 0$, then (22) and (23) become

$$y_1^* = x_1\beta_1 + y_2\gamma_1 + u_1 \tag{26}$$

$$y_2^* = x_2\beta_2 + u_2 \tag{27}$$

which is a special case of Model II where the employment equation does not contain all the explanatory variables ($\theta_2 = 0$).

If $\gamma_1 = 0$, then (C2) implies that $\delta_2 = 0$, thus (22) and (23) become

$$y_1^* = \frac{1}{1-\theta_1\theta_2}(x_1\beta_1 + x_2\beta_2\theta_1 + u_1 + u_2\theta_1) \tag{28}$$

$$y_2^* = \frac{1}{1-\theta_1\theta_2}(x_1\beta_1\theta_2 + x_2\beta_2 + u_1\theta_2 + u_2) \tag{29}$$

which is our Model III.

By exploiting the logical consistency conditions, we are able to provide a unifying framework to derive the ad hoc models as well as their variants from the general model in a transparent and coherent way. One advantage of our derivation is that it reveals the meaning of the reduced form equations in Model I and Model II. Our procedure demonstrates how the reduced form parameters are related to the structural parameters. In addition, the model given by (18) and (19) and the model given by (26) and (27) illustrate that it is possible to tackle the endogeneity problem by means of a simultaneous equations model without imposing any exclusion restrictions on the explanatory variables.

2.3 *Likelihood Function and Estimation Method*

For Model I, we will employ the following assumption to derive the joint likelihood function:

> ε_1 and ε_2 follow a bivariate normal distribution with mean zero, unit variances, and correlation coefficient ρ.

We use maximum likelihood to estimate the models. Separate estimations of (3) and (4) will not give consistent estimates for the parameters in (4) if ε_1 and ε_2 are correlated. Therefore, we employ a joint estimation. The joint likelihood and log-likelihood functions are given by

$$\begin{aligned} L &= \prod_{i=1}^{n}\prod_{m=0}^{4}\prod_{p=0}^{1}\{\Phi_2\left[\alpha_m - x_i\pi_1, q(x_{2i}\beta_2 + y_{1i}\gamma_2), -q\rho\right] \\ &\quad - \Phi_2\left[\alpha_{m-1} - x_i\pi_1, q(x_{2i}\beta_2 + y_{1i}\gamma_2), -q\rho\right]\}^{I(y_{1i}=m, y_{2i}=p)}, \end{aligned} \tag{30}$$

and

$$\begin{aligned} \log L &= \sum_{i=1}^{n}\sum_{m=0}^{4}\sum_{p=0}^{1} I(y_{1i} = m, y_{2i} = p) \\ &\quad \times \log\{\Phi_2\left[\alpha_m - x_i\pi_1, q(x_{2i}\beta_2 + y_{1i}\gamma_2), -q\rho\right] \\ &\quad - \Phi_2\left[\alpha_{m-1} - x_i\pi_1, q(x_{2i}\beta_2 + y_{1i}\gamma_2), -q\rho\right]\}, \end{aligned} \tag{31}$$

respectively, where Φ_2 is the standard bivariate normal distribution, $p = \{0, 1\}$, $m = \{0, 1, 2, 3, 4\}$, and $q = 2p - 1$.[4] We maximize the log-likelihood (31) with respect to $(\pi_1, \beta_2, \gamma_2, \rho)$ for the model (3) and (4), and with respect to $(\beta_1, \beta_2, \gamma_2, \rho)$ for the model (18) and (19).

Perhaps because of Maddala's claim (pp.122-123, Model 6), it is generally believed that if ε_1 and ε_2 are not independent (i.e. $\rho \neq 0$), then identification requires that the x_2 in (4) do not include all the exogenous variables in (3). However, Wilde (2000) demonstrates that Maddala's proof is erroneous as he only considers the special case where both x_2 and x contain a constant only. Wilde (2000) points out that identification only requires the existence of one varying exogenous regressor.

We encounter several computational problems. First, we need to evaluate numerically bivariate integrals for the maximization procedure. As is well known, the computation is time consuming for large samples. Second, this joint likelihood function is similar to that of the bivariate probit model. As in Greene (1995), the likelihood function is not globally concave because of the correlation coefficient ρ. Therefore, the choice of the initial estimates become important. Following Greene's (1995) suggestion, we choose the maximum likelihood probit and ordered probit estimates as the initial estimates. Third, as the likelihood function is not globally concave, Newton's method does not work well. To reduce computation time, we use the BHHH algorithm (Berndt, Hall, Hall, Hausman 1974).[5] Fourth, the estimation of the variance-covariance matrix is also complicated. The likelihood function (30) is similar to the multinomial probit model. Following the discussion of Lee (1996), although it is possible to estimate the covariance matrix, the likelihood function is too flat for the parameters of the covariance matrix. Therefore, we use the cross-product of the first derivatives of the parameters to obtain a consistent estimate of the variance-covariance matrix.

[4] Note that $\Phi_2(-\infty, z, \rho) = 0$, and $\Phi_2(\infty, z, \rho) = \Phi(z)$, where Φ is the standard normal distribution.

[5] This algorithm is also used in Sickles and Taubman (1986). The advantage of this computational method is that it only requires first-order partial derivatives (Gourieroux 2000, p.11). It takes about 30 minutes to estimate a model (which will require more than 500 minutes if the maximization procedure utilizes the Broyden, Fletcher, Goldfarb, Shanno (BFGS) algorithm). The parameter estimates obtained from these two methods only differ at about 10^{-4}.

Next, consider Model II. The assumption we impose is the same as that on Model I. The joint likelihood and log-likelihood functions are similar to (30) and (31), except that $x_i\pi_1$ and $x_{2i}\beta_2 + y_{1i}\gamma_2$ are replaced by $x_{1i}\beta_1 + y_{2i}\gamma_1$ and $x_i\pi_2$, respectively. The maximum likelihood estimation is then carried out by maximizing (31) with respect to $(\beta_1, \gamma_1, \pi_2, \rho)$ for the model (5) and (6), and $(\beta_1, \gamma_1, \beta_2, \rho)$ for the model (26) and (27). Similar to Model I, as we do not assume $\rho = 0$, identification requires that there exists at least one varying exogenous regressor.

For Model III, (20) and (21) can be expressed as

$$\begin{aligned} y_1^* &= \frac{1}{1-\theta_1\theta_2}(x_1\beta_1 + x_2\beta_2\theta_1 + u_1 + u_2\theta_1) \\ &= x^*\Pi_1 + \varepsilon_1^* \end{aligned} \tag{32}$$

and

$$\begin{aligned} y_2^* &= \frac{1}{1-\theta_1\theta_2}(x_1\beta_1\theta_2 + x_2\beta_2 + u_1\theta_2 + u_2) \\ &= x^*\Pi_2 + \varepsilon_2^*, \end{aligned} \tag{33}$$

respectively, where $x^*\Pi_1 = \frac{1}{1-\theta_1\theta_2}(x_1\beta_1 + x_2\beta_2\theta_1)$, $\varepsilon_1^* = u_1 + u_2\theta_1$, $x^*\Pi_2 = \frac{1}{1-\theta_1\theta_2}(x_1\beta_1\theta_2 + x_2\beta_2)$, and $\varepsilon_2^* = u_1\theta_2 + u_2$.

We employ the following assumption to derive the joint likelihood function:

> ε_1^* and ε_2^* follow a bivariate normal distribution with mean zero, unit variances, and correlation coefficient ρ.

The joint likelihood and log-likelihood functions are similar to (30) and (31), respectively, except that $x_i\pi_1$ and $x_{2i}\beta_2 + y_{1i}\gamma_2$ are replaced by $x_i^*\Pi_1$ and $x_i^*\Pi_2$, respectively. The maximum likelihood estimation is then carried out by maximizing (31) with respect to $(\beta_1, \theta_1, \beta_2, \theta_2, \rho)$.

3 Estimation

3.*1 Description of Variables*

We estimate the models using data from The Omnibus Household Survey on Health-related Issues of the Hong Kong Population in the Third Quarter of 1999. It is a special cross-sectional data set collected by the Health and Welfare Bureau of the Hong Kong Government. The sample consists of 33,763 individuals in 10,057 households and the data are obtained by face-to-face interviews. Since employment is a main focus of our study, we consider the adult sample only. After deleting missing observations and removing individuals younger than 18, we obtain 24,418 individuals (9,569 households) in our sample. Table 1 displays a list of the variables used in the estimation. Table 2 presents some descriptive statistics of the variables. The mean and standard deviation of the self-reported health status are 2.1292 and 0.6212, respectively. Hence, on average, the respondents report more or less the same health condition as others of the same age.[6] The average

[6] The sample mean 2.1292 is not statistically different from 2 as the t-ratio is only 0.2 (0.1292/0.6212 = 0.2).

age is 43.1, hence our sample is much younger than the ones studied in the literature on health status and labor supply. For example, the mean age of the sample in Sickles and Taubman (1986) is 64.3. Tables 3 and 4 report the distribution of the variables by health status and employment status, respectively. We offer here a discussion of some possible relationships between the variables and health status.

CSSA

CSSA stands for Comprehensive Social Security Assistance, which is the major welfare program in Hong Kong. The objective of CSSA is to provide financial assistance to bring the income of needy individuals and families to a prescribed level to meet their basic and special needs. Although it is a means-tested program, it takes up a significant percentage of the government's budget. The main recipients are the elderly, the unemployed, and single parent families. For example, there were 228,015 CSSA cases in 1999, the distribution by type is 58.36% old age, 11.48% unemployment, 11.03% single parent family, 8.76% temporary disability/ill health, 3.68% mentally ill, 3.51% low earnings, and 1.24% physically disabled (Census and Statistics Department 2000, Table 13.4). It is expected that this welfare variable has a negative relationship with health.

Gender

Using adult health data from four countries, Strauss et al. (1993) find that there is a gender difference in health. Their results indicate that although females have a longer life expectancy, they report more physical problems than males of the same age. They argue that males have a higher probability of suffering from very acute symptoms (hence a higher probability of death), whereas females may suffer from chronic diseases which are less severe than males (hence longer life expectancy). Therefore, one expects that males will report a better health condition than females.

Regular Teeth Check

Regular teeth check can be regarded as an indicator of healthy life style or the willingness to invest in health. It is expected that the variable has a positive relationship with health.

Age

Many studies find that health status is negatively related with age (e.g., Anderson and Burkhauser (1985), Grossman (1972a), Kenkel (1995), Lee (1982), and Sickles and Taubman (1986)). However, Ferraro and Yu (1995) argue that there is growing evidence that people who are 75 years and older will perceive their health to be better. One can interpret this as the outcome of a selection effect because older people tend to consider themselves to be healthier, otherwise they would not have been able to survive to old age. To check whether the effect of age on health is nonlinear, dummy variables for different age groups are used.

Marrige

Marriage and health have a complex relationship. Lillard and Panis (1996) argue that marriage has two different effects on health: protective and selective. The protective effect means that married people are healthier because they receive family care. In contrast, the selective effect has two opposite components. First, because of the protective effect, people with lower health status will put more effort in seeking marriage since it will improve their health. This selective effect is negative because marriage is associated with lower health status. Second, it is easier for healthy people to get married because they may be more attractive in the marriage market, giving rise to a positive selective effect. The net impact of marriage on health cannot be determined theoretically. Unfortunately, our cross-sectional data do not allow us to separate the effects because a longitudinal data set (as in Lillard and Panis (1996)) is needed to resolve the issue. We also include an interaction term age $\times$ marriage to check whether the impact of marriage on health depends on age.

Education

Education is expected to exert a positive effect on health in at least two ways. First, education improves health because more knowledge increases the efficiency of the production of health capital. Second, education may be a proxy for time preference. More education may imply more patience and forward-looking, which means that there is a higher willingness to invest in long-term capital including health capital. We also include an interaction term age $\times$ education to investigate whether the marginal impact of education on health depends on age.

Housing

Housing is a good measure of wealth in Hong Kong. There are two main types of housing in Hong Kong: private and public. In 1998, about 35% (2.3 million) of the population in Hong Kong live in public rental housing (Hong Kong Housing Authority 1999). Housing type may affect health because people living in a better environment will probably be healthier, physically and mentally. In addition, private housing residents have on average higher income and more assets than public housing residents (since a household will not be eligible for public housing if its income or net asset value exceeds a certain upper limit). If wealth and health are positively related, then the housing variable should be able to reflect it.

Long-term Diseases/Short-term Health Problem

An objective measure of health, long-term diseases, is used in the health equation. This variable is used to test whether behavioral variables play a role in determining health. If health is purely naturally determined, then after controlling for long-term diseases and pre-determined variables like age and sex, behavioral variables such as marriage and education should have no impact on health. Using the variable "problem within 14 days," we can also control for short-term shocks which may affect the perception on health.

Finally, we discuss the issue of identification in our models. Currie and Madrian (1999, pp.3352-3353) claim that "estimates of the relationship between health and labor force outcomes vary widely and are sensitive to the identification assumptions employed" and "many of the studies either ignore endogeneity issues altogether or rely on exclusion restrictions that are not easy to justify". Our Model I and II can handle the endogeneity problem without imposing any exclusion restrictions (see the model given by (18) and (19), as well as the model given by (26), and (27), where x_1 and x_2 can be identical). On the other hand, for the rest of our models, we need identification variables to establish the empirical relationship between health and labor supply. The system of equations for our joint estimation is overidentified as more than one variable in x_2 is excluded from x_1, and also more than one variable in x_1 is excluded from x_2. The variables age $\times$ marriage, age $\times$ education, age group dummies and problem within 14 days are used to identify the employment equation (as they are only included in x_1). On the other hand, the variables age, cohorts, number of children below 6, number of children between 6 and

12, and income ratio can be used for the identification of the health equation (as they are only included in x_2). We believe that the cohort dummies are reasonable identification variables for the health equation in our study because the responses to the self-assessed health question are based on comparison with people of the same age.[7] Diagnostic checks for whether they are appropriate identification variables are performed by estimating the model with and without the cohort variables. The estimates of the cohorts are not statistically different from zero in the health equation. Both the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) indicate that there is no information gained by incorporating the cohort dummies in the health equation.[8] On the other hand, the propensity to be employed is significantly affected by the cohort dummies.

3.2 Findings

Tables 5 through 10 present the estimation results. The results in Table 5 are obtained by estimating (5) and (4) separately by maximum likelihood. Four different specifications of x_1 (the first four columns) are used for (5) while three specifications of x_2 (the last three columns) are used for (4). In Table 6, the first two columns contain the estimates for (3) and (4), whereas the last two columns are for (18) and (19). Table 7 is similar to Table 6 except that the former includes the cohort variables. In Table 8, the first two columns report the estimates for (5) and (6), whereas the last two columns are for (26) and (27). Table 9 is similar to Table 8 except that the cohort variables are added to the former. Table 10 reports the estimates for Model III: the first two columns with no cohort variables and the last two columns with the cohort variables.

The estimates show that $\widehat{\theta}_1\widehat{\theta}_2 < 1$, hence the necessary condition for identification is satisfied. Comparing Tables 5 and 10, the joint estimates do not differ much from the separate estimates except for the parameters for health and employment.[9] For proper comparison, contrast the first and sixth columns in Table 5 with the last two columns in Table 10: the estimates for employment (0.0331) and health (0.0458) are both statistically significant in Table 5, whereas the estimate for employment (0.0067) is insignificant and the estimate for health (0.1524) is significant in Table 10. Hence, separate maximum likelihood estimation of (5) and (4) produces two misleading results: employment is found to be a significant determinant of health and the impact of health on employment is substantially understated ($0.0458 < 0.1524$). This finding is supported by the results in Tables 6 through 9. The second and fourth columns in Tables 6 and 7 show that health is a significant determinant of employment, whereas the first and third columns in Tables 8 and 9 show that employment does not affect health. In sum, the joint estimation results for Models I, II and III reveal consistently that health is a significant positive determinant of employment, but employment has no significant impact on health.

[7] The means of the health status variable are 2.184, 2.165, 2.161, and 2.044 for the cohort '70 and after, '60-'69, '50-'59, and before '50, respectively. This suggests that the mean health status does not vary much with cohort. On the other hand, the means of the employment variable are 0.686, 0.779, 0.718, and 0.32 for the cohort '70 and after, '60-'69, '50-'59, and before '50, respectively. This shows that employment does vary with cohort.

[8] AIC and BIC are defined as $(-2\log L + 2p)/T$ and $(-2\log L + p\log T)/T$, respectively, where L is the likelihood of the model, T is the number of observations and p is the number of parameters.

[9] Proper comparison can only be made between the separate estimates and the estimates of the structural parameters.

Since the models are nonnested, we use the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) for model selection. Using these two criteria, Model II ($\theta_1 = 0$, $\theta_2 \neq 0$, with cohort variables) has the lowest AIC and BIC (i.e., the first two columns in Table 9, AIC = 2.1352, BIC = 2.1498). Therefore, we will focus on the this model and interpret the determinants of health.

Our results show that males have better health. As expected, those who have regular teeth check are found to be healthier. Marriage has a positive impact on health, but the effect diminishes with age. The estimates indicate that age and health have a nonlinear relationship. The youngest (age 18-25) and the oldest (age $\geq$ 85) groups report better health than the middle group (age 25-85). The oldest group reports better health probably because of a selection effect (the survival of the fittest). Those who were weak might have died when they were young, leaving those who are strong to survive to old age. The estimate for education turns out to be negative, which is a puzzling result.

Private housing residents are healthier than public housing residents. Household income has a positive effect on health. As expected, both long-term diseases and problem within 14 days are found to be negatively related to health. Comparing the values of the estimates, long-term diseases have significantly more negative impact on health.

4 Conclusion

This paper investigates empirically the interrelationship between health status and labor supply as well as their respective determinants. By exploiting the logical consistency conditions, we provide a unifying framework to develop three different simultaneous equations models to capture various possible forms of interdependence between health and labor supply. In contrast to the ad hoc models in the literature, our models are derived in a coherent and systematic way. We also demonstrate that it is possible to tackle the endogeneity problem by means of a simultaneous equations model without imposing any exclusion restrictions on the explanatory variables. We estimate the models using a large cross-sectional data set obtained from a recent survey of the Hong Kong population. Our results indicate that separate estimation of the health and labor supply equations will produce misleading results, and that the two equations should be jointly estimated. We find strong evidence that health status is a significant determinant of employment, but not vice versa.

Throughout the paper we only consider a two-equation model. However, one could argue that, in addition to labor supply, other variables (such as personal income, insurance, and marital status) may also have the endogeneity problem with health. In principle, one could add an equation for each potential endogenous variable to Models I, II, or III and jointly estimate the multiple simultaneous equations model. We have not yet made any attempt in this direction because even with modern computer technology, it is computationally difficult to handle trivariate or higher dimensional models with such a large sample. There are already too many issues that need to be resolved for a simple two-equation model. We hope that our contribution provides the first-step for this line of inquiry.

References

[1] K.H. Anderson, and R.V. Burkhauser. "The Importance of the Measure of Health in Empirical Estimates of the Labor Supply of Older Men." *Economics Letters* 16 (1984): 375-380.

[2] K.H. Anderson, and R.V. Burkhauser. "The Retirement-Health Nexus: A New Measure of an Old Puzzle." *Journal of Human Resources* 20 (1985): 315-330.

[3] E.K. Berndt; B.H. Hall; R.E. Hall; and J.A. Hausman. "Estimation and inference in nonlinear structural models." *Annals of Economic and Social Measurement* 3(4) (1974): 635-65.

[4] J. Currie, and B.C. Madrian. "Health, health insurance and the labor market." In *Handbook of Labor Economics,* Vol. 3, edited by O. Ashenfelter and D. Card. Amsterdam: Elsevier Science B.V., 1999.

[5] P. Diamond, and J. Hausman. "The retirement and unemployment behavior of older men." In *Retirement and Economic Behavior*, edited by H.Aaron and G. Burtless. Washington, D.C.: Brookings Institution, 1984.

[6] D.J. Ekerdt; L. Baden; R. Bosse; and E. Dibbs. "The effect of retirement on physical health." *American Journal of Public Health* 73 (1983): 779-783.

[7] K.F. Ferraro, and Y. Yu. "Body weight and self-ratings of health." *Journal of Health and Social Behavior* 36 (September 1995): 274-284.

[8] C. Gourieroux. *Econometrics of Qualitative Dependent Variables.* Cambridge: Cambridge University Press, 2000.

[9] L. Graur; B. West; and P. Gregory. ""How do you feel?" Self-reported health as an indicator of current physical and mental health status." *Journal of Psychosocial Nursing and Mental Health Services* 36 (June 1998): 24-30.

[10] W. Greene. *Limdep Version 7.0 User Menu.* New York: Econometric Software Inc, 1995.

[11] M. Grossman. *The Demand For Health: A Theoretical And Empirical Investigation.* New York: Columbia University Press, 1972.

[12] J. Hausman, and D. Wise. "Social security, health status, and retirement." In *Pensions, Labor, and Individual Choice*, edited by D.A. Wise. Chicago: University of Chicago Press, 1985.

[13] J.J. Heckman. "Simultaneous equation models with continuous and discrete endogenous variables and structural shifts." In *Studies in Nonlinear Estimation*, edited by S.M. Goldfeld and R.E. Quandt. Cambridge: Ballinger Publishing Company, 1976.

[14] J.J. Heckman. "Dummy endogenous variables." *Econometrica* 46 (July 1978): 931-959.

[15] Hong Kong Census and Statistics Department. *Hong Kong Annual Digest of Statistics 2000*. Hong Kong: HKSAR Government Printing Department, 2000.

[16] Hong Kong Housing Authority. *Hong Kong Housing Authority Annual Report 1998/1999*. Hong Kong: Hong Kong Housing Authority, 1999.

[17] E.L. Idler, and Y. Benyamini. "Self-rated health and mortality: A review of twenty-seven community studies." *Journal of Health and Social Behavior* 38 (1997): 21-37.

[18] D.S. Kenkel. "Should you eat breakfast? Estimates from health production functions." *Health Economics* 4 (1995): 15-29.

[19] L.F. Lee. "Health and wage: A simultaneous equation model with multiple discrete indicators." *International Economic Review* 23 (1982): 199-221.

[20] M.J. Lee. *Methods of Moments and Semiparametric Econometrics for Limited Dependent Variable Models*. New York: Springer, 1996.

[21] L.A. Lillard, and C.W.A. Panis. "Marital status and mortality: The role of health." *Demography* 33 (1996): 313-327.

[22] G.S. Maddala. *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press, 1983.

[23] D.O. Parsons. "The male labor force participation decision: Health, reported health, and economic incentives." *Economica* 49 (February 1982): 81-91.

[24] C.E. Ross, and J. Mirowsky. "Does employment affect health?" *Journal of Health and Social Behavior* 36 (September 1995): 230-243.

[25] P. Schmidt. "Constraints on the parameters in simultaneous tobit and probit models." In *Structural Analysis of Discrete Data with Econometric Applications,* edited by C.F. Manski and D. McFadden. Cambridge, Massachusetts: The MIT Press, 1981.

[26] R.C. Sickles, and P. Taubman. "An analysis of the health and retirement status of the elderly." *Econometrica* 54 (November 1986): 1339-1356.

[27] S. Stern. "Measuring the effect of disability on labor force participation." *Journal of Human Resources* 24 (Summer 1989): 361-395.

[28] J. Strauss; P. Gertler; O. Rahman; and K. Fox. "Gender and Life-Cycle Differentials in the Patterns and Determinants of Adult Health." *Journal of Human Resources* 28 (1993): 791-837.

[29] P. Taubman, and S. Rosen. "Healthiness, education and marital status." In *Economic Aspects of Health*, edited by V.R. Fuchs. Chicago and London: The University of Chicago Press, 1982.

[30] J. Wilde. "Identification of multiple equation probit models with endogenous dummy regressors." *Economics Letters* 69 (2000): 309-312.

Table 1
Description of Variables

Variables	Description
Health	Self-rated health, in comparison to others of the same age: 0=Much worse, 1=Worse, 2=More or less the same, 3=Better, 4=Much better.
Employment	1 if the respondent has a full-time or part-time job, is self-employed or working for family business without salary, 0 otherwise.
Constant	Constant 1 (Intercept).
Insurance	1 if the respondent is currently covered by medical insurance, including plans purchased by the individual, his family, employer or his household members' employers, 0 otherwise.
Regular Teeth Check	1 if the respondent has teeth check on a regular basis, 0 otherwise.
Male	1 if the respondent is male, 0 otherwise.
Age	Continuous variable: ranges from 18-99.
Married	1 if the respondent cohabitates or is married, 0 if the respondent is single, separated, divorced or widowed.
Age × Marriage	Age times Married.
Education	The highest education level (for those who completed schooling), or the class the respondent is attending: 1=No education of Kindergarten, 2=Primary, 3=Lower secondary (F.1 to F.3), 4=Upper Secondary (F. 4 to F.5), 5=Matriculation, 6=Tertiary or Above.
Age × Education	Age times Education.
Age Group	1 if the respondent falls in the respective age group, 0 otherwise.
Public Housing	1 if the respondent lives in public housing, 0 otherwise.
Private Housing	1 if the respondent lives in private housing, 0 otherwise.
Long-term Diseases	1 if the respondent has any diseases that require long-term follow-up by doctors, 0 otherwise.
Problems (14 Days)	1 if the respondent has any health-related problems in the last 14 days of the survey, 0 otherwise.
Children < 6	Number of Children age below 6 in the household.
Children 6-12	Number of Children age between 6 and 12 in the household.
Personal Income†	Monthly personal income, including all sources: 0=No Income 1=Below \$2000, 2=\$2000-\$3999, 3=\$4000-\$4999, 4=\$5000-\$5999, 5=\$6000-\$6999, 6=\$7000-\$7999, 7=\$8000-\$8999, 8=\$9000-\$9999, 9=\$10000-\$12499, 10=\$12500-\$14999, 11=\$15000-\$19999, 12=\$20000-\$24999, 13=\$25000-\$29999, 14=\$30000-\$39999, 15=\$40000-\$49999, 16=\$50000-\$59999, 17=\$60000-\$69999, 18=\$70000 or above.
Household Income	Monthly household income, including all sources. (Values are assigned in the same way as personal income)
Income Ratio	Personal Income divided by Household Income, with 0 for those with no Household Income.
CSSA	1 if the respondent is a recipient of the Comprehensive Social Security Assistance (CSSA), 0 otherwise.
Cohort : '70 and after	1 if the respondent was born after 1970, 0 otherwise.
Cohort : '60-'69	1 if the respondent was born between 1960 and 1969, 0 otherwise.
Cohort : '50-'59	1 if the respondent was born between 1950 and 1959, 0 otherwise.
Cohort : Before '50	1 if the respondent was born before 1950, 0 otherwise.

Table 2
Descriptive Statistics of Variables (N = 24,418)

	Mean	Standard Deviation	Min	Max
Health	2.1292	0.6212	0	4
Employment	0.5972	0.4905	0	1
Insurance	0.3871	0.4871	0	1
Regular Teeth Check	0.1640	0.3703	0	1
Male	0.4847	0.4998	0	1
Married	0.6457	0.4783	0	1
Education	3.4112	1.5179	1	6
Age	43.146	16.590	18	99
Public Housing	0.4041	0.4907	0	1
Private Housing	0.4234	0.4941	0	1
Long-term Diseases	0.1578	0.3646	0	1
Problems (14 Days)	0.2004	0.4003	0	1
Children < 6	0.1985	0.4758	0	3
Children 6-12	0.2614	0.5612	0	5
Personal Income	5.5621	4.8136	0	18
Household Income	11.214	4.0883	0	18
Income Ratio	0.5041	0.3766	0	1
CSSA (Individual)	0.0265	0.1606	0	1

Table 3
Number of Observations by Health Status

	Much better	Better	More or less the same	Worse	Much worse	Row total
Self-Reported Health	363	5199	16291	2360	205	24418
Economically Active	247	3503	11024	1055	57	15886
Employed	231	3243	10159	906	43	14582
Unemployed	16	260	865	149	14	1304
Non-Active	116	1696	5267	1305	148	8532
Student	14	218	727	43	0	1002
Homemakers	39	756	2369	400	32	3596
Retired Persons	60	682	1988	649	61	3440
Others	3	40	183	213	55	494
Insurance						
Yes	145	2121	6583	570	33	9452
No	218	3078	9708	1790	172	14966
Regular Teeth Check						
Yes	80	973	2665	263	23	4004
No	283	4226	13626	2097	182	20414
Problems (14 Days)						
Yes	43	715	2876	1127	133	4894
No	320	4484	13415	1233	72	19524
Sex						
Male	186	2694	7934	936	86	11836
Female	177	2505	8357	1424	119	12582
Marital Status						
Single	92	1391	4688	370	24	6565
Married$^\sharp$	224	3386	10414	1602	140	15766
Divorced*	8	125	283	75	9	500
Widowed	39	297	906	313	32	1587
Education						
Low$^\#$	129	1517	4851	1281	140	7918
Middle$^\natural$	145	2602	8090	814	50	11701
High$^\circ$	89	1080	3350	265	15	4799
Housing						
Public Rental Housing	113	1888	6547	1196	124	9868
Private Housing	173	2432	6890	783	60	10338
Others	77	879	2854	381	21	4212

$^\sharp$Include those who are cohabitated.
*Include those who are separated.
$^\#$Include kindergarten and primary.
$^\natural$Include lower secondary (F.1 to F.3) and upper secondary (F.4 to F.5).
$^\circ$Include matriculation (F.6 to F.7) and tertiary or above.

Table 3
(continued)

	Much better	Better	More or less the same	Worse	Much worst	Row total
Household Size						
1	23	248	633	193	21	1118
2	55	774	2232	435	54	3550
3	79	1020	3285	481	39	4904
4	109	1616	5103	656	44	7528
5	69	1033	3193	375	27	4697
6	15	327	1315	159	15	4697
7	11	111	318	43	5	488
8	2	34	142	14	0	192
9	0	20	41	3	0	64
10	0	3	18	1	0	22
11	0	3	10	0	0	13
14	0	10	1	0	0	11
Household Income (Monthly)						
≥40000	68	1031	3226	271	13	4609
20000-39999	115	1836	5632	666	35	8284
10000-19999	121	1465	4659	668	52	6975
5000-9999	37	518	1732	383	52	2722
<5000	22	349	1042	372	43	1828
Long-term Diseases						
Yes	28	427	2002	1232	164	3853
No	335	4772	14289	1128	41	20565
Age						
18-24	39	674	2261	155	7	3136
25-34	66	1038	3681	274	16	5075
35-44	90	1405	4356	464	20	6335
45-54	65	910	2714	429	32	4150
55-64	43	500	1384	336	46	2309
65-74	32	408	1228	426	50	2144
75-84	14	207	560	224	31	1036
≥85	14	57	107	52	3	233
Cohort						
'70 and after	64	1088	3679	249	14	5094
'60-'69	76	1191	4134	358	16	5775
'50-'59	84	1280	3781	482	27	5654
Before '50	139	1640	4697	1271	148	7895
CSSA						
Yes	6	83	348	177	33	647
No	357	5116	15943	2183	172	23771

Table 4
Number of Observations by Employment

	Employed	Unemployed	Row total
Employment	14582	9836	24418
Self Reported Health			
Much better	231	132	363
Better	3243	1956	5199
More or less the same	10159	6132	16291
Worse	906	1454	2360
Much worse	43	162	205
Insurance			
Yes	7955	1497	9452
No	6627	8339	14966
Regular Teeth Check			
Yes	2974	1030	4004
No	11608	8806	20414
Problems (14 Days)			
Yes	2708	2186	4894
No	11874	7650	19524
Sex			
Male	8595	3241	11836
Female	5987	6595	12582
Marital Status			
Single	4823	1742	6565
Married$^\sharp$	9355	6411	15766
Divorced*	272	228	500
Widowed	132	1455	1587
Education			
Low$^\#$	2704	5214	7918
Middle$^\natural$	8358	3343	11701
High$^\circ$	3520	1279	4799
Housing			
Public Rental Housing	5211	4657	9868
Private Housing	6643	3695	10338
Others	2728	1484	4212

$^\sharp$Include those who cohabitate.
*Include those who are separated.
$^\#$Include kindergarten and primary.
$^\natural$Include lower secondary (F.1 to F.3) and upper secondary (F.4 to F.5).
$^\circ$Include matriculation (F.6 to F.7) and tertiary or above.

Table 4
(Continued)

	Employed	Unemployed	Row Total
Household Size			
1	483	635	1118
2	2021	1529	3550
3	2997	1907	4904
4	4644	2884	7528
5	2932	1765	4697
6	1055	776	1831
7	287	201	488
8	107	85	192
9	37	27	64
10	13	9	22
11	6	7	13
14	0	11	11
Household Income (Monthly)			
≥40000	3610	999	4609
20000-39999	5875	2409	8284
10000-19999	3889	3086	6975
5000-9999	1046	1676	2722
<5000	162	1666	1828
Long-term Diseases			
Yes	1140	2713	3853
No	13442	7123	20565
Age			
18-24	1835	1301	3136
25-34	4201	874	5075
35-44	4617	1718	6335
45-54	2799	1351	4150
55-64	894	1415	2309
65-74	214	1930	2144
75-84	20	1016	1036
≥85	2	231	233
Cohort			
'70 and after	3496	1598	5094
'60-'69	4499	1276	5775
'50-'59	4058	1596	5654
Before '50	2529	5366	7895
CSSA			
Yes	8	639	647
No	14574	9197	23771

Table 5
Separate Maximum Likelihood Estimates for the Health and Employment equations

Variables	Health				Employment		
Constant	2.8587***	2.8760***	2.8518***	2.7027***	-2.0345***	-0.7244***	0.0448**
	(59.913)	(56.917)	(56.217)	(25.377)	(-19.165)	(-3.869)	(2.058)
Insurance	0.0326*	0.0302*	0.0334*	0.0327*	0.4378***	0.4136***	0.4023***
	(1.836)	(1.702)	(1.877)	(1.755)	(13.486)	(12.663)	(11.601)
Regular Teeth Check	0.1057***	0.1053***	0.1066***	0.1060***	-0.4267***	-0.4461***	-0.4359***
	(4.921)	(4.895)	(4.954)	(4.880)	(-9.921)	(-10.327)	(-10.223)
Male	0.0815***	0.0836***	0.0814***	0.0771***	0.1961***	0.2480***	0.2140***
	(5.023)	(5.160)	(5.021)	(4.546)	(6.414)	(7.853)	(6.432)
Married	0.1109**	0.0904*	0.1240**	0.1056*	0.2897***	0.0751**	-0.8828***
	(2.505)	(1.848)	(2.489)	(1.920)	(8.599)	(1.982)	(-6.682)
Age × Marriage	-0.0039***	-0.0037***	-0.0041***	-0.0039***	-	-	0.0181***
	(-4.198)	(-3.593)	(-3.962)	(-3.707)			(7.153)
Education	-0.0851***	-0.0825***	-0.0918***	-0.0759***	-0.1825***	-0.1775***	-0.3019***
	(-7.485)	(-5.306)	(-5.744)	(-4.168)	(-14.804)	(-13.748)	(-8.109)
Age × Education	0.0025***	0.0025***	0.0027***	0.0024***	-	-	0.0029***
	(9.883)	(7.308)	(7.705)	(6.562)			(3.683)
Age Group							
18-25	0.1572***	-	0.1294***	0.1402***	-	-	-0.7015***
	(4.891)		(3.566)	(3.473)			(-7.996)
≥85	0.2398***	-	0.2375***	0.1977***	-	-	0.2265
	(3.124)		(3.088)	(2.929)			(1.317)
Age	-	-	-	0.0022	-0.0427***	-0.0623***	-0.0837***
				(1.391)	(-34.656)	(-23.664)	(-23.673)
Public Housing	-0.0777***	-0.0778***	-0.0780***	-0.0745***	0.1919***	0.1813***	0.1696***
	(-3.528)	(-3.532)	(-3.535)	(-3.335)	(4.643)	(4.359)	(3.657)
Private Housing	0.0394*	0.0386*	0.0390*	0.0396*	-0.0727*	-0.0789**	-0.0960**
	(1.843)	(1.801)	(1.819)	(1.831)	(-1.787)	(-1.920)	(-2.133)
Long-term Diseases	-0.8735***	-0.8727***	-0.8735***	-0.8766***	-0.2573***	-0.2238***	-0.2015***
	(-37.683)	(-37.541)	(-37.569)	(-42.853)	(-6.102)	(-5.204)	(-4.926)
Problems (14 Days)	-0.5030***	-0.5020***	-0.5026***	-0.5033***	-	-	-0.0791**
	(-26.296)	(-26.247)	(-26.269)	(-28.770)			(-2.239)
Children < 6	-	-	-	0.0216	-0.1343***	-0.0959***	-0.0898***
				(1.229)	(-4.379)	(-3.014)	(-2.759)
Children 6-12	-	-	-	0.0565***	-0.0940***	-0.1244***	-0.1183***
				(3.967)	(-3.581)	(-4.569)	(-4.403)
Household Income	0.0039*	0.0035	0.0038*	0.0047**	0.2343***	0.2319***	0.2262***
	(1.752)	(1.601)	(1.737)	(1.965)	(53.067)	(51.709)	(60.323)
Income Ratio	-	-	-	0.0257	4.2709***	4.2187***	4.1766***
				(0.794)	(81.092)	(79.486)	(91.409)
CSSA (Individual)	-0.3303***	-0.3301***	-0.3292***	-0.3456***	-3.1235***	-3.2570***	-3.3251***
	(-6.830)	(-6.831)	(-6.802)	(-7.715)	(-18.713)	(-19.287)	(-24.700)
Cohort : '70 and after	-	0.0223	0.0323*	0.0810	-	-1.0559***	-0.7679***
		(1.154)	(1.671)	(1.384)		(-9.809)	(-6.446)
Cohort : '60-'69	-	0.0988*	0.0508	0.0030	-	-0.4817***	-0.4692***
		(1.943)	(0.999)	(0.070)		(-6.003)	(-6.064)
Cohort : '50-'59	-	-0.0150	-0.0051	0.0110	-	-0.0887	-0.1037*
		(-0.390)	(-0.136)	(0.344)		(-1.531)	(-1.823)
Employment	0.0331*	0.0041	0.0073	0.0304	-	-	-
	(1.721)	(0.145)	(0.256)	(1.149)			
Health	-	-	-	-	0.0352	0.0458**	0.0394*
					(1.546)	(1.986)	(1.722)
Threshold Parameters							
α_1	1.3682***	1.3668***	1.3681***	1.3690***	-	-	-
	(46.285)	(46.297)	(46.286)	(45.772)			
α_2	3.5628***	3.5606***	3.5629***	3.5651***	-	-	-
	(111.671)	(111.734)	(111.673)	(110.607)			
α_3	5.0354***	5.0319***	5.0357***	5.0386***	-	-	-
	(134.401)	(134.482)	(134.402)	(134.396)			
Mean log-likelihood	-0.863269	-0.863675	-0.863210	-0.862797	-0.209225	-0.205616	-0.202495
Number of Parameters	20	21	23	27	16	19	24
AIC	1.7282	1.7291	1.7283	1.7278	0.4198	0.4128	0.4070
BIC	1.7348	1.7360	1.7359	1.7368	0.4251	0.4191	0.4149

*** 1 % Level of Significance ** 5 % Level of Significance * 10 % Level of Significance
Numbers in parenthesis are t-values.

Table 6
Joint Maximum Likelihood Estimates for Model I (No Cohort Variable)

	$\theta_1 \neq 0, \theta_2 = 0$		$\theta_1 = \theta_2 = 0$	
Variables	Health	Employment	Health	Employment
Constant	2.8169***	-2.4874***	2.8672***	-2.5541***
	(32.954)	(-10.748)	(60.960)	(-11.190)
Insurance	0.0327*	0.4315***	0.0382**	0.4288***
	(1.766)	(12.841)	(2.128)	(12.730)
Regular Teeth Check	0.1030***	-0.4335***	0.1045***	-0.4340***
	(4.747)	(-10.393)	(4.845)	(-10.443)
Male	0.0782***	0.1823***	0.0900***	0.1793***
	(4.614)	(5.795)	(5.770)	(5.694)
Married	0.0528	0.2968***	0.1080**	0.2944***
	(0.994)	(8.479)	(2.455)	(8.466)
Age × Marriage	-0.0030***	-	-0.0040***	-
	(-2.977)		(-4.586)	
Education	-0.0784***	-0.1903***	-0.0807***	-0.1901***
	(-4.602)	(-15.746)	(-7.213)	(-15.767)
Age × Education	0.0024***	-	0.0025***	-
	(7.055)		(10.308)	
Age Group				
18-25	0.1465***	-	0.1267***	-
	(4.158)		(3.755)	
≥85	0.2257***	-	0.2253***	-
	(3.475)		(3.695)	
Age	0.0003	-0.0428***	-	-0.0427***
	(0.208)	(-35.954)		(-35.690)
Public Housing	-0.0745***	0.2001***	-0.0771***	0.2013***
	(-3.336)	(4.410)	(-3.468)	(4.451)
Private Housing	0.0391*	-0.0780*	0.0381*	-0.0785*
	(1.807)	(-1.780)	(1.760)	(-1.797)
Long-term Diseases	-0.8755***	-0.1422**	-0.8788***	-0.1248*
	(-42.883)	(-2.014)	(-43.738)	(-1.772)
Problems (14 Days)	-0.5047***	-	-0.5033***	-
	(-28.844)		(-28.794)	
Children < 6	0.0200	-0.1365***	-	-0.1345***
	(1.165)	(-4.445)		(-4.409)
Children 6-12	0.0525***	-0.1001***	-	-0.0940***
	(3.747)	(-3.941)		(-3.741)
Household Income	0.0057***	0.2310***	0.0046**	0.2303***
	(2.687)	(50.576)	(2.264)	(48.958)
Income Ratio	0.0442*	4.2224***	-	4.2136***
	(1.820)	(67.137)		(64.634)
CSSA (Individual)	-0.3550***	-3.0411***	-0.3398***	-3.0263***
	(-8.166)	(-21.813)	(-7.938)	(-21.602)
Employment	-	-	-	-
Health	-	0.2709**	-	0.3041***
		(2.304)		(2.625)
Threshold Parameters				
α_1	1.3701***	-	1.3694***	-
	(45.686)		(45.708)	
α_2	3.5657***	-	3.5636***	-
	(110.353)		(110.447)	
α_3	5.0381***	-	5.0353***	-
	(134.240)		(134.383)	
ρ	-0.1652**		-0.1886**	
	(-2.012)		(-2.330)	
Mean log-likelihood	-1.07209		-1.07246	
Number of Parameters	40		36	
AIC	2.1475		2.1479	
BIC	2.1607		2.1598	

*** 1 % Level of Significance ** 5 % Level of Significance * 10 % Level of Significance
Numbers in parenthesis are t-values. Likelihood ratio test statistic for $\theta_1 = 0$ is 18.069, $\chi^2_{0.01,4} = 13.28$.

Table 7

Joint Maximum Likelihood Estimates for Model I (Include Cohort Variables)

	$\theta_1 \neq 0, \theta_2 = 0$		$\theta_1 = \theta_2 = 0$	
Variables	Health	Employment	Health	Employment
Constant	2.7459***	-1.1465***	2.8641***	-1.1189***
	(25.987)	(-4.007)	(60.896)	(-3.806)
Insurance	0.0343*	0.4081***	0.0391*	0.4086***
	(1.851)	(12.000)	(2.179)	(12.003)
Regular Teeth Check	0.1042***	-0.4524***	0.1051***	-0.4522***
	(4.800)	(-10.791)	(4.872)	(-10.767)
Male	0.0771***	0.2352***	0.0898***	0.2355***
	(4.547)	(7.227)	(5.759)	(7.203)
Married	0.0814	0.0822**	0.1079**	0.0815**
	(1.483)	(2.108)	(2.454)	(2.088)
Age × Marriage	-0.0034***	-	-0.0039***	-
	(-3.251)		(-4.481)	
Education	-0.0819***	-0.1786***	-0.0806***	-0.1786***
	(-4.531)	(-14.602)	(-7.199)	(-14.585)
Age × Education	0.0025***	-	0.0024***	-
	(6.887)		(10.226)	
Age Group				
18-25	0.1181***	-	0.1380***	-
	(2.931)		(4.086)	
≥85	0.2099***	-	0.2339***	-
	(3.115)		(3.838)	
Age	0.0012	-0.0623***	-	-0.0622***
	(0.818)	(-25.116)		(-25.009)
Public Housing	-0.0742***	0.1891***	-0.0775***	0.1888***
	(-3.320)	(4.116)	(-3.486)	(4.101)
Private Housing	0.0388*	-0.0837*	0.0385*	-0.0832*
	(1.791)	(-1.889)	(1.778)	(-1.875)
Long-term Diseases	-0.8775***	-0.1187*	-0.8782***	-0.1294*
	(-42.958)	(-1.674)	(-43.699)	(-1.819)
Problems (14 Days)	-0.5046***	-	-0.5039***	-
	(-28.827)		(-28.818)	
Children < 6	0.0218	-0.0993***	-	-0.0969***
	(1.238)	(-3.165)		(-3.097)
Children 6-12	0.0561***	-0.1305***	-	-0.1243***
	(3.941)	(-4.941)		(-4.745)
Household Income	0.0059***	0.2291***	0.0047**	0.2297***
	(2.727)	(50.429)	(2.265)	(52.431)
Income Ratio	0.0495**	4.1772***	-	4.1897***
	(2.033)	(68.355)		(74.161)
CSSA (Individual)	-0.3564***	-3.1827***	-0.3384***	-3.1943***
	(-8.197)	(-22.081)	(-7.905)	(-22.380)
Employment	-	-	-	-
Health	-	0.2614**	-	0.2398**
		(2.190)		(2.000)
Cohort : '70 and after	0.0881	-1.0520***	-	-1.0453***
	(1.506)	(-10.063)		(9.962)
Cohort : '60-'69	0.0021	-0.4750***	-	-0.4759***
	(0.048)	(-6.222)		(-6.235)
Cohort : '50-'59	0.0110	-0.0904*	-	-0.0893*
	(0.344)	(-1.668)		(-1.647)
Threshold Parameters				
α_1	1.3700***	-	1.3688***	-
	(45.671)		(45.706)	
α_2	3.5659***	-	3.5632***	-
	(110.331)		(110.451)	
α_3	5.0387***	-	5.0351***	-
	(134.259)		(134.381)	
ρ	-0.1512*		-0.1361	
	(-1.814)		(-1.626)	
Mean log-likelihood	-1.06838		-1.06890	
Number of Parameters	46		39	
AIC	2.1405		2.1410	
BIC	2.1558		2.1539	

*** 1 % Level of Significance ** 5 % Level of Significance * 10 % Level of Significance

Numbers in parenthesis are t-values. Likelihood ratio test statistic for $\theta_1 = 0$ is 25.39472, $\chi^2_{0.01,7} = 18.48$.

Table 8
Joint Maximum Likelihood Estimates for Model II (No Cohort Variable)

	$\theta_1 = 0, \theta_2 \neq 0$		$\theta_1 = \theta_2 = 0$	
Variables	Health	Employment	Health	Employment
Constant	2.8592***	-0.1894	2.8587***	-1.9643***
	(60.840)	(-1.136)	(60.843)	(-21.899)
Insurance	0.0339*	0.4041***	0.0327*	0.4383***
	(1.799)	(11.862)	(1.735)	(13.056)
Regular Teeth Check	0.1057***	-0.4221***	0.1057***	-0.4241***
	(4.902)	(-9.907)	(4.906)	(-10.059)
Male	0.0830***	0.2001***	0.0816***	0.1978***
	(4.931)	(6.094)	(4.838)	(6.398)
Married	0.1124**	-0.7245***	0.1109**	0.2881***
	(2.526)	(-6.085)	(2.485)	(8.220)
Age × Marriage	-0.0040***	0.0173***	-0.0039***	-
	(-4.469)	(7.023)	(-4.419)	
Education	-0.0847***	-0.3834***	-0.0851***	-0.1887***
	(-7.440)	(-10.783)	(-7.483)	(-15.506)
Age × Education	0.0025***	0.0046***	0.0025***	-
	(10.432)	(6.237)	(10.456)	
Age Group				
18-25	0.1559***	-0.7708***	0.1571***	-
	(4.546)	(-10.734)	(4.601)	
≥85	0.2389***	-0.0453	0.2398***	-
	(3.920)	(-0.231)	(3.935)	
Age	-	-0.0733***	-	-0.0426***
		(-24.367)		(-35.661)
Public Housing	-0.0776***	0.1754***	-0.0777***	0.1904***
	(-3.491)	(3.794)	(-3.494)	(4.184)
Private Housing	0.0393*	-0.0954**	0.0394*	-0.0722
	(1.817)	(-2.125)	(1.820)	(-1.635)
Long-term Diseases	-0.8741***	-0.2332***	-0.8735***	-0.2738***
	(-43.084)	(-5.922)	(-43.078)	(-7.025)
Problems (14 Days)	-0.5030***	-0.0834**	-0.5030***	-
	(-28.765)	(-2.413)	(-28.755)	
Children < 6	-	-0.1446***	-	-0.1331***
		(-4.550)		(-4.307)
Children 6-12	-	-0.1247***	-	-0.0927***
		(-4.730)		(-3.647)
Household Income	0.0040*	0.2281***	0.0039*	0.2344***
	(1.892)	(61.519)	(1.830)	(64.483)
Income Ratio	-	4.1715***	-	4.2711***
		(91.539)		(97.158)
CSSA (Individual)	-0.3317***	-3.2684***	-0.3304***	-3.1264***
	(-7.695)	(-24.670)	(-7.667)	(-24.286)
Employment	0.0272	-	0.0328	-
	(1.098)		(1.315)	
Health	-	-	-	-
Threshold Parameters				
α_1	1.3680***	-	1.3681***	-
	(45.706)		(45.700)	
α_2	3.5626***	-	3.5627***	-
	(110.410)		(110.420)	
α_3	5.0352***	-	5.0352***	-
	(134.306)		(134.322)	
ρ	0.0079		0.0004	
	(0.398)		(0.021)	
Mean log-likelihood	-1.06730		-1.07255	
Number of Parameters	41		36	
AIC	2.1380		2.1480	
BIC	2.1516		2.1600	

*** 1 % Level of Significance ** 5 % Level of Significance * 10 % Level of Significance
Numbers in parenthesis are t-values. Likelihood ratio test statistic for $\theta_2 = 0$ is 256.389, $\chi^2_{0.01,5} = 15.09$.

Table 9
Joint Maximum Likelihood Estimates for Model II (Include Cohort Variables)

	$\theta_1 = 0, \theta_2 \neq 0$		$\theta_1 = \theta_2 = 0$	
Variables	Health	Employment	Health	Employment
Constant	2.8596***	0.5269**	2.8594***	-0.6356***
	(60.839)	(2.489)	(60.849)	(-3.658)
Insurance	0.0348*	0.4031***	0.0348*	0.4142***
	(1.850)	(11.628)	(1.846)	(12.186)
Regular Teeth Check	0.1056***	-0.4333***	0.1056***	-0.4432***
	(4.899)	(-10.142)	(4.899)	(-10.469)
Male	0.0843***	0.2156***	0.0843***	0.2502***
	(5.005)	(6.481)	(4.996)	(7.831)
Married	0.1137**	-0.8199***	0.1148**	0.0739*
	(2.553)	(-6.662)	(2.569)	(1.900)
Age × Marriage	-0.0040***	0.0180***	-0.0040***	-
	(-4.502)	(7.126)	(-4.506)	
Education	-0.0844***	-0.3034***	-0.0846***	-0.1769***
	(-7.411)	(-8.158)	(-7.448)	(-14.388)
Age × Education	0.0025***	0.0029***	0.0025***	-
	(10.409)	(3.750)	(10.435)	
Age Group				
18-25	0.1549***	-0.6982***	0.1561***	-
	(4.515)	(-7.964)	(4.565)	
≥85	0.2383***	0.2696	0.2384***	-
	(3.909)	(1.334)	(3.911)	
Age	-	-0.0836***	-	-0.0621***
		(-23.660)		(-24.956)
Public Housing	-0.0776***	0.1681***	-0.0776***	0.1793***
	(-3.489)	(3.621)	(-3.489)	(3.899)
Private Housing	0.0393*	-0.0955**	0.0393*	-0.0780*
	(1.814)	(-2.120)	(1.817)	(-1.748)
Long-term Diseases	-0.8747***	-0.2186***	-0.8746***	-0.2456***
	(-43.115)	(-5.522)	(-43.139)	(-6.242)
Problems (14 Days)	-0.5030***	-0.0893**	-0.5027***	-
	(-28.760)	(-2.572)	(-28.732)	
Children < 6	-	-0.0894***	-	-0.0949***
		(-2.746)		(-3.012)
Children 6-12	-	-0.1176***	-	-0.1233***
		(-4.376)		(-4.681)
Household Income	0.0041*	0.2263***	0.0041*	0.2320***
	(1.945)	(60.249)	(1.944)	(62.338)
Income Ratio	-	4.1763***	-	4.2182***
		(91.316)		(94.819)
CSSA (Individual)	-0.3328***	-3.3292***	-0.3327***	-3.2626***
	(-7.720)	(-24.709)	(-7.718)	(-24.393)
Employment	0.0226	-	0.0229	-
	(0.914)		(0.919)	
Health	-	-	-	-
Cohort : '70 and after	-	-0.7657***	-	-1.0530***
		(-6.425)		(-10.024)
Cohort : '60-'69	-	-0.4950***	-	-0.4806***
		(-6.047)		(-6.276)
Cohort : '50-'59	-	-0.1035*	-	-0.0881
		(-1.818)		(-1.615)
Threshold Parameters				
α_1	1.3679***	-	1.3680***	-
	(45.701)		(45.697)	
α_2	3.5625***	-	3.5625***	-
	(110.393)		(110.392)	
α_3	5.0351***	-	5.0351***	-
	(134.292)		(134.287)	
ρ	0.0143		0.0138	
	(0.722)		(0.697)	
Mean log-likelihood	-1.06581		-1.06896	
Number of Parameters	44		39	
AIC	2.1352		2.1411	
BIC	2.1498		2.1541	

*** 1 % Level of Significance ** 5 % Level of Significance * 10 % Level of Significance
Numbers in parenthesis are t-values. Likelihood ratio test statistic for $\theta_2 = 0$ is 153.8334, $\chi^2_{0.01,5} = 15.09$.

Table 10
Joint Maximum Likelihood Estimates for the Health and Employment equations (Model III)

	No Cohorts Variable		Include Cohorts Variables	
Variables	Health	Employment	Health	Employment
Constant	2.8851***	-2.4595***	2.8794***	-1.0296***
	(59.020)	(-13.140)	(58.869)	(-4.281)
Insurance	0.0296	0.4342***	0.0319*	0.4110***
	(1.539)	(12.841)	(1.665)	(12.039)
Regular Teeth Check	0.1048***	-0.4427***	0.1056***	-0.4574***
	(4.865)	(-10.369)	(4.901)	(-10.709)
Male	0.0788***	0.1787***	0.0797***	0.2342***
	(4.536)	(5.672)	(4.565)	(7.217)
Married	0.0903**	0.3008***	0.0910**	0.0846**
	(2.014)	(8.413)	(2.016)	(2.140)
Age × Marriage	-0.0037***	-	-0.0036***	-
	(-4.135)		(-3.967)	
Education	-0.0830***	-0.1934***	-0.0821***	-0.1804***
	(-7.276)	(-15.617)	(-7.202)	(-14.449)
Age × Education	0.0025***	-	0.0025***	-
	(10.366)		(10.209)	
Age Group				
18-25	0.1259***	-	0.1404***	-
	(3.698)		(4.095)	
≥85	0.2305***	-	0.2433***	-
	(3.783)		(3.979)	
Age	-	-0.0434***	-	-0.0626***
		(-34.871)		(-24.955)
Public Housing	-0.0772***	0.2073***	-0.0778***	0.1928***
	(-3.473)	(4.487)	(-3.500)	(4.136)
Private Housing	0.0388*	-0.0818*	0.0391*	-0.0855*
	(1.791)	(-1.844)	(1.806)	(-1.911)
Long-term Diseases	-0.8752***	-0.0878	-0.8746***	-0.1006
	(-43.092)	(-1.205)	(-43.041)	(-1.379)
Problems (14 Days)	-0.5024***	-	-0.5038***	-
	(-28.756)		(-28.823)	
Children < 6	-	-0.1362***	-	-0.0974***
		(-4.413)		(-3.092)
Children 6-12	-	-0.0940***	-	-0.1238***
		(-3.694)		(-4.694)
Household Income	0.0035	0.2330***	0.0037*	0.2309***
	(1.614)	(63.232)	(1.683)	(61.403)
Income Ratio	-	4.2641***	-	4.2147***
		(96.375)		(94.462)
CSSA (Individual)	-0.3248***	-3.0541***	-0.3236***	-3.2060***
	(-7.346)	(-23.213)	(-7.286)	(-23.482)
Employment	0.0077	-	0.0067	-
	(1.390)		(1.221)	
Health	-	0.1952***	-	0.1524**
		(2.980)		(2.315)
Cohort: '70 and after	-	-	-	-1.0490***
				(-9.992)
Cohort: '60-'69	-	-	-	-0.4781***
				(-6.243)
Cohort: '50-'59	-	-	-	-0.0896
				(-1.642)
Threshold Parameters				
α_1	1.3678***	-	1.3678***	-
	(45.733)		(45.730)	
α_2	3.5622***	-	3.5623***	-
	(110.512)		(110.499)	
α_3	5.0348***	-	5.0348***	-
	(134.429)		(134.419)	
ρ	0.0157		0.0248	
	(1.014)		(1.602)	
Mean log-likelihood	-1.07240		-1.06885	
Number of Parameters	37		40	
AIC	2.1478		2.1410	
BIC	2.1601		2.1543	

*** 1 % Level of Significance   ** 5 % Level of Significance   * 10 % Level of Significance
Numbers in parenthesis are t-values.

Appendix: Logical Consistency of the General Model

In this appendix, we prove the following proposition.

Proposition 2 For the model

$$y_1^* = \frac{1}{1-\theta_1\theta_2}\left[y_1(\delta_1+\gamma_2\theta_1)+y_2(\gamma_1+\delta_2\theta_1)+x_1\beta_1+x_2\beta_2\theta_1+u_1+u_2\theta_1\right] \tag{12}$$

$$y_2^* = \frac{1}{1-\theta_1\theta_2}\left[y_1(\gamma_2+\delta_1\theta_2)+y_2(\delta_2+\gamma_1\theta_2)+x_1\beta_1\theta_2+x_2\beta_2+u_1\theta_2+u_2\right] \tag{13}$$

to be logically consistent, either (C3) or (C4), or both, must hold:

$$\delta_1+\gamma_2\theta_1=0,\ \ \delta_2+\gamma_1\theta_2=0,\ \text{ and } \gamma_1+\delta_2\theta_1=0 \tag{C3}$$

$$\delta_1+\gamma_2\theta_1=0,\ \ \delta_2+\gamma_1\theta_2=0,\ \text{ and } \gamma_2+\delta_1\theta_2=0 \tag{C4}$$

Proof: For convenience, let us rewrite (12) and (13) as

$$y_1^* = x\Pi_1 + y_1\lambda_1 + y_2\eta_1 + v_1, \tag{34}$$

and

$$y_2^* = x\Pi_2 + y_2\lambda_2 + y_1\eta_2 + v_2, \tag{35}$$

respectively, where $x\Pi_1 = \frac{1}{1-\theta_1\theta_2}(x_1\beta_1 + x_2\beta_2\theta_1)$, $x\Pi_2 = \frac{1}{1-\theta_1\theta_2}(x_1\beta_1\theta_2 + x_2\beta_2)$, $\lambda_1 = \frac{1}{1-\theta_1\theta_2}(\delta_1+\gamma_2\theta_1)$, $\eta_1 = \frac{1}{1-\theta_1\theta_2}(\gamma_1+\delta_2\theta_1)$, $\lambda_2 = \frac{1}{1-\theta_1\theta_2}(\delta_2+\gamma_1\theta_2)$, $\eta_2 = \frac{1}{1-\theta_1\theta_2}(\gamma_2+\delta_1\theta_2)$, $v_1 = \frac{1}{1-\theta_1\theta_2}(u_1+u_2\theta_1)$, and $v_2 = \frac{1}{1-\theta_1\theta_2}(u_1\theta_2+u_2)$. We divide our proof into two parts.
Part 1. We prove that a model of the form

$$y_1^* = x\Pi_1 + y_1\lambda_1 + v_1 \tag{36}$$

$$y_2^* = x\Pi_2 + y_2\lambda_2 + v_2 \tag{37}$$

cannot be logically consistent unless $\lambda_1 = 0$ and $\lambda_2 = 0$. For the proof that $\lambda_2 = 0$, see Proposition 2 in Heckman (1976). For the proof that $\lambda_1 = 0$, Heckman's Proposition 2 is not strictly applicable here because y_1 is a polychotomous variable (instead of a binary variable). We offer a brief proof here. Substituting (36) into (1),

$$y_1 = \begin{cases} 4 & \text{if} \quad \alpha_3 - x\Pi_1 - y_1\lambda_1 < v_1 < \alpha_4 \\ 3 & \text{if} \quad \alpha_2 - x\Pi_1 - y_1\lambda_1 < v_1 \le \alpha_3 - x\Pi_1 - y_1\lambda_1 \\ 2 & \text{if} \quad \alpha_1 - x\Pi_1 - y_1\lambda_1 < v_1 \le \alpha_2 - x\Pi_1 - y_1\lambda_1 \\ 1 & \text{if} \quad -x\Pi_1 - y_1\lambda_1 < v_1 \le \alpha_1 - x\Pi_1 - y_1\lambda_1 \\ 0 & \text{if} \quad \alpha_{-1} < v_1 \le -x\Pi_1 - y_1\lambda_1 \end{cases},$$

then the probability $P(y_1 = m)$, $m = 0, 1, ..., 4$, will be given by

$$\begin{aligned}
P(y_1 = 0) &= F(-x\Pi_1), \\
P(y_1 = 1) &= F(\alpha_1 - x\Pi_1 - \lambda_1) - F(-x\Pi_1 - \lambda_1), \\
P(y_1 = 2) &= F(\alpha_2 - x\Pi_1 - 2\lambda_1) - F(\alpha_1 - x\Pi_1 - 2\lambda_1), \\
P(y_1 = 3) &= F(\alpha_3 - x\Pi_1 - 3\lambda_1) - F(\alpha_2 - x\Pi_1 - 3\lambda_1), \\
P(y_1 = 4) &= 1 - F(\alpha_3 - x\Pi_1 - 4\lambda_1),
\end{aligned}$$

where $F(.)$ denotes the distribution function of v_1. Assume that $v_1 \in (-\infty, \infty)$ and $F(.)$ is continuous, then the sum of the above five probabilities will be equal to one if and only if $\lambda_1 = 0$ (see Figure 1a and Figure 1b below). Therefore, the model given by (36) and (37) is not logically consistent unless both $\lambda_1 = 0$ and $\lambda_2 = 0$.

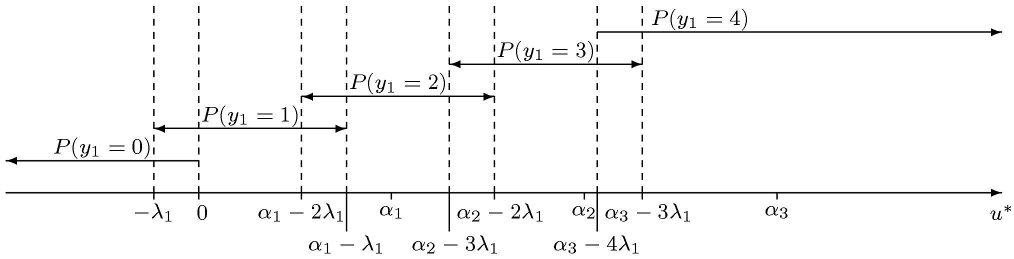

Figure 1a. $\lambda_1 > 0$ and $x^*\Pi_1 = 0$.

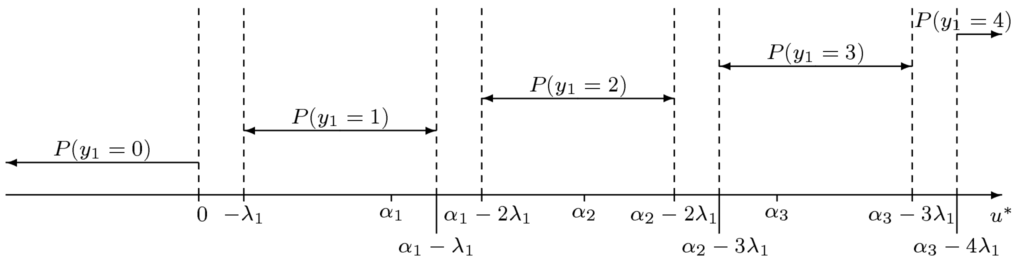

Figure 1b. $\lambda_1 < 0$ and $x^*\Pi_1 = 0$.

Part 2. We prove that a model of the form

$$y_1^* = x_1\beta_1 + y_2\eta_1 + v_1 \tag{38}$$

$$y_2^* = x_2\beta_2 + y_1\eta_2 + v_2 \tag{39}$$

is not logically consistent unless $\eta_1\eta_2 = 0$ (i.e., either $\eta_1 = 0$ or $\eta_2 = 0$, or both). First, let $G(w_1, w_2) = \int_{-\infty}^{w_1}\int_{-\infty}^{w_2} g(v_1, v_2)dv_1dv_2$ be the distribution function of v_1 and v_2, $G_1(w_1) = \int_{-\infty}^{w_1}\int_{-\infty}^{\infty} g(v_1, v_2)dv_2dv_1$ be the marginal distribution of v_1, and $G_2(w_2) = \int_{-\infty}^{w_2}\int_{-\infty}^{\infty} g(v_1, v_2)dv_1dv_2$ be the marginal distribution of v_2, where $g(v_1, v_2)$ is the joint probability density function of v_1 and v_2. For simplicitly, assume that $v_1 \in (-\infty, \infty)$, $v_2 \in (-\infty, \infty)$, $G(., .)$ is continuous, and $g(., .)$ is symmetric. The joint probabilities of

y_1 and y_2 are given by

$$\begin{aligned}
P(y_1 = 0, y_2 = 0) &= G(-x_1\beta_1, -x_2\beta_2), \\
P(y_1 = 0, y_2 = 1) &= G(-x_1\beta_1 - \eta_1, x_2\beta_2), \\
P(y_1 = 1, y_2 = 0) &= G(\alpha_1 - x_1\beta_1, -x_2\beta_2 - \eta_2) - G(-x_1\beta_1, -x_2\beta_2 - \eta_2), \\
P(y_1 = 1, y_2 = 1) &= G(\alpha_1 - x_1\beta_1 - \eta_1, x_2\beta_2 + \eta_2) - G(-x_1\beta_1 - \eta_1, x_2\beta_2 + \eta_2), \\
P(y_1 = 2, y_2 = 0) &= G(\alpha_2 - x_1\beta_1, -x_2\beta_2 - 2\eta_2) - G(\alpha_1 - x_1\beta_1, -x_2\beta_2 - 2\eta_2), \\
P(y_1 = 2, y_2 = 1) &= G(\alpha_2 - x_1\beta_1 - \eta_1, x_2\beta_2 + 2\eta_2) - G(\alpha_1 - x_1\beta_1 - \eta_1, x_2\beta_2 + 2\eta_2), \\
P(y_1 = 3, y_2 = 0) &= G(\alpha_3 - x_1\beta_1, -x_2\beta_2 - 3\eta_2) - G(\alpha_2 - x_1\beta_1, -x_2\beta_2 - 3\eta_2), \\
P(y_1 = 3, y_2 = 1) &= G(\alpha_3 - x_1\beta_1 - \eta_1, x_2\beta_2 + 3\eta_2) - G(\alpha_2 - x_1\beta_1 - \eta_1, x_2\beta_2 + 3\eta_2), \\
P(y_1 = 4, y_2 = 0) &= G_2(-x_2\beta_2 - 4\eta_2) - G(\alpha_3 - x_1\beta_1, -x_2\beta_2 - 4\eta_2), \\
P(y_1 = 4, y_2 = 1) &= G_2(x_2\beta_2 + 4\eta_2) - G(\alpha_3 - x_1\beta_1 - \eta_1, x_2\beta_2 + 4\eta_2).
\end{aligned}$$

It can be easily shown that if both η_1 and η_2 are not zero, the model will not be logically consistent because the sum of the probabilities will not be equal to one. Figure 2 illustrates the case where η_1 and η_2 are both negative. The other cases are just similar.

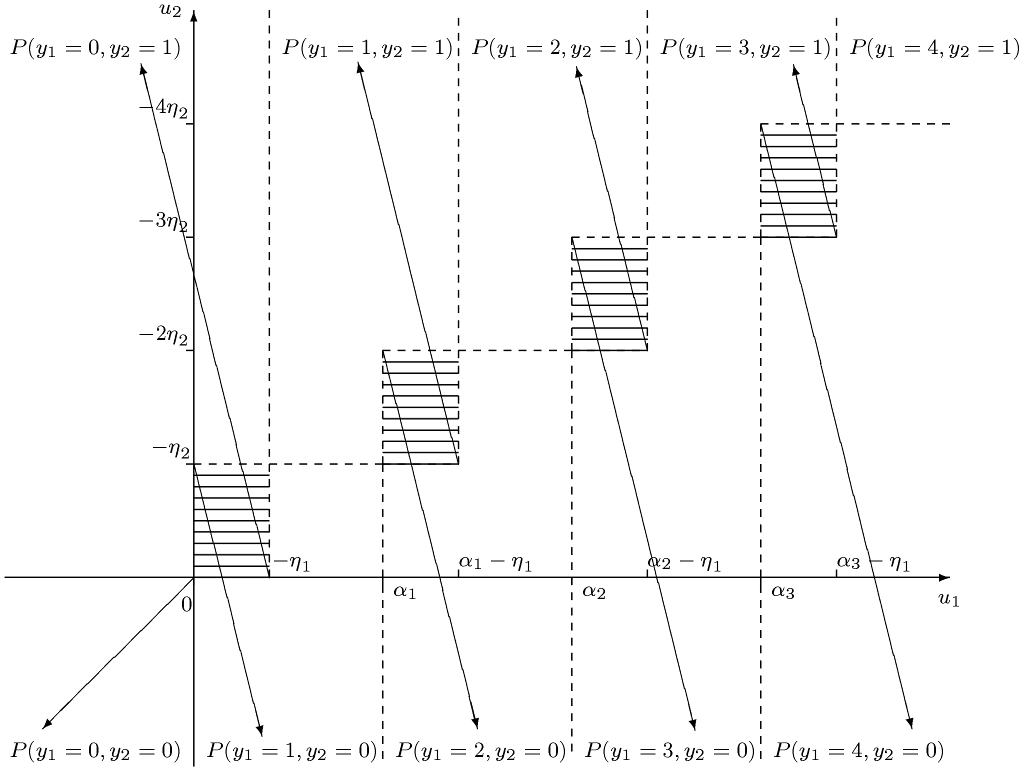

Figure 2. $\eta_1 < 0$, $\eta_2 < 0$, $x_1\beta_1 = 0$, and $x_2\beta_2 = 0$.

If $\eta_1 = 0$ or $\eta_2 = 0$ (or both), then the probabilities will sum to one. For example, suppose $\eta_1 = 0$, then

$$\begin{aligned}
P(y_1 = 0, y_2 = 0) &= G(-x_1\beta_1, -x_2\beta_2), \\
P(y_1 = 0, y_2 = 1) &= G(-x_1\beta_1, x_2\beta_2), \\
P(y_1 = 1, y_2 = 0) &= G(\alpha_1 - x_1\beta_1, -x_2\beta_2 - \eta_2) - G(-x_1\beta_1, -x_2\beta_2 - \eta_2),
\end{aligned}$$

$$
\begin{aligned}
P(y_1 = 1, y_2 = 1) &= G(\alpha_1 - x_1\beta_1, x_2\beta_2 + \eta_2) - G(-x_1\beta_1, x_2\beta_2 + \eta_2), \\
P(y_1 = 2, y_2 = 0) &= G(\alpha_2 - x_1\beta_1, -x_2\beta_2 - 2\eta_2) - G(\alpha_1 - x_1\beta_1, -x_2\beta_2 - 2\eta_2), \\
P(y_1 = 2, y_2 = 1) &= G(\alpha_2 - x_1\beta_1, x_2\beta_2 + 2\eta_2) - G(\alpha_1 - x_1\beta_1, x_2\beta_2 + 2\eta_2), \\
P(y_1 = 3, y_2 = 0) &= G(\alpha_3 - x_1\beta_1, -x_2\beta_2 - 3\eta_2) - G(\alpha_2 - x_1\beta_1, -x_2\beta_2 - 3\eta_2), \\
P(y_1 = 3, y_2 = 1) &= G(\alpha_3 - x_1\beta_1, x_2\beta_2 + 3\eta_2) - G(\alpha_2 - x_1\beta_1, x_2\beta_2 + 3\eta_2), \\
P(y_1 = 4, y_2 = 0) &= G_2(-x_2\beta_2 - 4\eta_2) - G(\alpha_3 - x_1\beta_1, -x_2\beta_2 - 4\eta_2), \\
P(y_1 = 4, y_2 = 1) &= G_2(x_2\beta_2 + 4\eta_2) - G(\alpha_3 - x_1\beta_1, x_2\beta_2 + 4\eta_2).
\end{aligned}
$$

It follows that

$$
\begin{aligned}
P(y_1 = 0, y_2 = 0) + P(y_1 = 0, y_2 = 1) &= G_1(-x_1\beta_1), \\
P(y_1 = 1, y_2 = 0) + P(y_1 = 1, y_2 = 1) &= G_1(\alpha_1 - x_1\beta_1) - G_1(-x_1\beta_1), \\
P(y_1 = 2, y_2 = 0) + P(y_1 = 2, y_2 = 1) &= G_1(\alpha_2 - x_1\beta_1) - G_1(\alpha_1 - x_1\beta_1), \\
P(y_1 = 3, y_2 = 0) + P(y_1 = 3, y_2 = 1) &= G_1(\alpha_3 - x_1\beta_1) - G_1(\alpha_2 - x_1\beta_1), \\
P(y_1 = 4, y_2 = 0) + P(y_1 = 4, y_2 = 1) &= 1 - G_1(\alpha_3 - x_1\beta_1).
\end{aligned}
$$

The sum of probabilities is now equal to one. Hence, the model given by (38) and (39) is logically consistent if $\eta_1\eta_2 = 0$.

Now we can combine Part 1 and Part 2. Since the model given by (36) and (37) and the model given by (38) and (39) are special cases of (34) and (35), therefore the model given by (34) and (35) is not logically consistent unless $\lambda_1 = 0$, $\lambda_2 = 0$, and $\eta_1\eta_2 = 0$. Consequently, for the model (9) and (10) to be logically consistent, either (C3) or (C4), or both, must hold.