HEALTH AND WAGE: A SIMULTANEOUS EQUATION
MODEL WITH MULTIPLE DISCRETE INDICATORS

by
Lung-Fei Lee

Discussion Paper No. 79-127, December 1979

Center for Economic Research
Department of Economics
University of Minnesota
Minneapolis, Minnesota 55455

HEALTH AND WAGE:

A SIMULTANEOUS EQUATION MODEL WITH MULTIPLE DISCRETE INDICATORS

by

Lung-Fei Lee¹

1. Introduction

In this article, we propose an econometric approach for estimating the simultaneous effects of health and wages in a structural equations framework. In the theoretical works of Grossman [1972a, 1972b], health is viewed as a capital stock which is determined endogenously a priori. The demand for health should rise with wage rate as the rate of return to health investment increases with wage. On the other hand, better health should raise market productivity and hence wage. Since health capital is an abstract theoretical concept, it can hardly be measured quantitatively; however, most population surveys include various health indicators. These indicators are usually discrete variables. In the empirical work of Grossman and Benham [1974], those qualitative indicators are treated as continuous variables and two stage least squares estimation techniques are used. This specification raises various econometric problems (see, e.g. Goldberger [1964, pp. 248-255], Domencich and McFadden [1975], and Nerlove and Press [1973]). A satisfactory specification should take into account the discrete nature of indicators.

Our specification generalizes the dummy variables simultaneous equations system of Heckman [1978] and measurement models of Joreskog and Goldberger

¹I appreciate having financial support from the National Science Foundation under Grant SOC-78-07304. I should like to thank Richard Hoffbeck for being my research assistant.

[1975] and is closely related to the models of Muthén [1976a, 1976b].

Muthén suggests using full information maximum likelihood procedures for estimation. Except for very simple cases, this does not generally yield feasible or tractable procedures. In this article, we introduce a tractable procedure for statistical estimation and testing. This specification and the estimation procedure is then applied to study the joint dependence of health and wages. Another departure of our approach from others is that we will test and correct the possible incompatibility of the various health capital indicators. In the population surveys, one may have subjective health indicators, such as self-evaluation of health, which may reflect different populations' perception of health. We propose a statistical procedure for testing measurement compatibility and correcting this problem.

This article is divided into two parts. In the first part, we consider model specification, estimation and statistical tests in a general simultaneous equation model with multiple discrete indicators. In the second part of the paper, we consider specification of the simultaneous equation model of health and wages. The model is then estimated and tested by the proposed procedures. Implications of the empirical results are discussed.

2. Model Specification, Normalization and Identification

Consider the following simultaneous equation model

$$(2.1) \quad \begin{array}{ccc} y_i^* B & = & x_i \Gamma + u_i \\ 1 \times G & G \times G & 1 \times K \quad K \times G \quad 1 \times G \end{array} \quad i=1, \dots, N$$

where $u_i \sim N(0, \Sigma)$, $E(u_i | x) = 0$, $E(u_i u_j | x) = 0$ for $i \neq j$ and B is a nonsingular matrix. Without loss of generality, we assume that the first m components $y_{i1}, \dots, y_{im}$ of y_i^* are observable continuous variables and the last $G-m$ components are unobservables. This model is similar to the classical simultaneous equation models except that some of the endogenous variables in y_i^* are unobservable. For the unobservable variables, we assume that there are indicators available which provide indirect measures of the underlying unobservable variables. Models with continuous indicators have been analyzed extensively (see e.g. Joreskog and Goldberger [1975]). Recently, Heckman [1978] and Muthén [1976] considered the case of discrete indicators. Heckman considered models with a single qualitative indicator, while Muthén considered the case of multiple dichotomous indicators.

Suppose $y_{i\ell}^*$ is an unobservable dependent variable and several (J_ℓ) qualitative indicators I_{ij}^ℓ , $j=1, \dots, J_\ell$ are available. The indicators are related to the underlying unobservables $y_{i\ell}^*$ by some measurement relations. Suppose I_{ij}^ℓ is a dichotomous variable with values 0 and 1, the measurement relation can be specified as,

$$(2.2) \quad P(I_{ij}^\ell = 1 | y_{i\ell}^*) = F(\alpha_{\ell j} + \lambda_{\ell j} y_{i\ell}^*; \sigma_{\ell j}^2),$$

where $F(z; \sigma^2) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{u}{2\sigma^2}) du$ is the normal distribution with variance σ^2 , evaluated at z . Obviously, not all the parameters in (2.2) are identifiable. The parameters in (2.1) and (2.2) need to be

normalized. It will be convenient to normalize (reparameterize) the parameters in (2.1) and (2.2) in the following way.

To simplify notations, let us suppress the subscript i which refers to sample i . The measurement equation can be rewritten as

$$(2.3) \quad z_{lj}^* = \alpha_{lj} + \lambda_{lj} y_{lj}^* + \epsilon_{lj} \quad \begin{matrix} j=1, \dots, J_l \\ l=m+1, \dots, G \end{matrix}$$

where z_{lj}^* is an unobservable index, $\epsilon_{lj} \sim N(0, \sigma_{\epsilon_{lj}}^2)$ and are mutually independent for all i and are independent with u_i . The index z_{lj}^* determines the observed discrete indicators as follows. Suppose I_{lj} is an ordered polychotomous indicator with values $1, \dots, K_{lj}$, I_{lj} and z_{lj}^* are related as follows:

$$(2.4) \quad \begin{aligned} I_{lj} &= 1 && \text{if } z_{lj}^* < 0 \\ I_{lj} &= 2 && \text{if } 0 \leq z_{lj}^* < \mu_{1lj} \\ &\vdots && \vdots \\ I_{lj} &= K_{lj} && \text{if } \mu_{K_{lj}-2, lj} \leq z_{lj}^* < \infty \end{aligned}$$

where $\mu_{s, lj}$, $s=1, \dots, K_{lj}-2$ are unknown constants. Obviously, not all of the parameters in (2.3) and (2.4) are identifiable; therefore, we normalize the parameters as follows.²

Let $y^* = (y^0, y_{m+1}^*, \dots, y_G^*)$ where $y^0 = (y_1, \dots, y_m)$ is the vector of observable dependent variables. The reduced form equations of (2.1) are

$$(2.5) \quad y^* = x\pi + \eta$$

²When there is only one indicator for each unobservable variable y_{lj}^* , normalizing $\alpha_{lj} = 0$, $\lambda_{lj} = 1$ and $\eta_{lj} = 0$, provides the system specified in Heckman [1978].

Denote the corresponding reduced form equations of y^0 in (2.5) as

$$(2.6) \quad y^0 = x\pi^0 + \eta^0$$

Substituting $y_\ell^* = x\pi_\ell + \eta_\ell$, $\ell = m+1, \dots, G$ into (2.3),

$$(2.7) \quad z_{\ell j}^* = \alpha_{\ell j} + \lambda_{\ell j} x\pi_\ell + \lambda_{\ell j} \eta_\ell + \varepsilon_{\ell j} \quad \begin{matrix} j=1, \dots, J_\ell, \\ \ell=m+1, \dots, G \end{matrix}$$

Let $\Omega_0 = E(\eta^0 \eta^0)$ and $\sigma_{\ell j}^0 = E((\lambda_{\ell j} \eta_\ell + \varepsilon_{\ell j}) \eta^0)$. Conditional on y^0 , equation (2.7) implies

$$(2.8) \quad z_{\ell j}^* = \alpha_{\ell j} + \lambda_{\ell j} x\pi_\ell + \sigma_{\ell j}^0 \Omega_0^{-1} (y^0 - x\pi^0) + \xi_{\ell j}$$

where $\xi_{\ell j} = \lambda_{\ell j} \eta_\ell + \varepsilon_{\ell j} - \sigma_{\ell j}^0 \Omega_0^{-1} \eta^0$ which are independent with η^0 .

Let us denote $\sigma_{\ell j}^2 = E(\xi_{\ell j}^2)$ the variance of $\xi_{\ell j}$. Equations in (2.8) can be rewritten as

$$(2.8)' \quad \begin{aligned} \frac{z_{\ell j}^*}{\sigma_{\ell j}} &= \frac{\alpha_{\ell j}}{\sigma_{\ell j}} + \frac{\lambda_{\ell j}}{\sigma_{\ell j}} x\pi_\ell + \frac{\sigma_{\ell j}^0}{\sigma_{\ell j}} \Omega_0^{-1} (y^0 - x\pi^0) + \frac{\xi_{\ell j}}{\sigma_{\ell j}} \\ &= \frac{\alpha_{\ell j}}{\sigma_{\ell j}} + \left(\frac{\lambda_{\ell j}}{\sigma_{\ell j}} \frac{\sigma_{\ell 1}}{\lambda_{\ell 1}} \right) x \left(\frac{\lambda_{\ell 1}}{\sigma_{\ell 1}} \pi_\ell \right) + \frac{\sigma_{\ell j}^0}{\sigma_{\ell j}} \Omega_0^{-1} (y^0 - x\pi^0) + \frac{\xi_{\ell j}}{\sigma_{\ell j}} \end{aligned}$$

where, implicitly, $\lambda_{\ell 1} \neq 0$, $\sigma_{\ell j} \neq 0$ for j and ℓ . Denote

$$D = \begin{bmatrix} 1 & & & \\ & \ddots & & \\ & & \lambda_{m+1,1}/\sigma_{m+1,1} & \\ & & & \ddots \\ & & & & \lambda_{G1}/\sigma_{G1} \end{bmatrix}$$

be a $G \times G$ diagonal matrix.

$$\text{Let } B^* = D^{-1}B, \quad \alpha_{\ell j}^* = \alpha_{\ell j}/\sigma_{\ell j}, \quad \lambda_{\ell j}^* = \frac{\lambda_{\ell j}}{\sigma_{\ell j}} \frac{\sigma_{\ell 1}}{\lambda_{\ell 1}}, \quad \mu_{s\ell j}^* = \mu_{s\ell j}/\sigma_{\ell j},$$

$$y^{**} = y^* D, \quad z_{\ell j}^{**} = z_{\ell j}^*/\sigma_{\ell j}, \quad \varepsilon_{\ell j}^* = \varepsilon_{\ell j}/\sigma_{\ell j} \quad \text{and} \quad \sigma_{\ell j}^{0*} = \sigma_{\ell j}^0/\sigma_{\ell j}.$$

Under this reparameterization, the structural equations are

$$(2.9) \quad y^{**} B^* = x\Gamma + u$$

and the measurement equations are

$$(2.10) \quad z_{lj}^{**} = \alpha_{lj}^* + \lambda_{lj}^* y_l^{**} + \varepsilon_{lj}^*$$

The indicators are determined as follows,

$$(2.11) \quad \begin{aligned} I_{lj} &= 1 && \text{if } z_{lj}^{**} < 0 \\ I_{lj} &= 2 && \text{if } 0 \leq z_{lj}^{**} < \mu_{1lj}^* \\ &\vdots && \vdots \\ I_{lj} &= K_{lj} && \text{if } \mu_{K_{lj}-2, lj}^* \leq z_{lj}^{**} \end{aligned}$$

From this construction, we have $\lambda_{l1}^* = 1$, $l = m+1, \dots, G$. The structural equations (2.9) imply the following reduced form equations

$$(2.12) \quad \begin{aligned} Y^{**} &= x\Gamma B^{*-1} + uB^{*-1} \\ &= x\pi^* + \eta^* \end{aligned}$$

where $\pi^* = \Gamma B^{*-1}$, $\eta^* = uB^{*-1}$. The advantage of this reparameterization (or normalization) is that the variances of ξ_{lj}^* in the reduced form measurement equations

$$(2.13) \quad z_{lj}^{**} = \alpha_{lj}^* + \lambda_{lj}^* x\pi_l^* + \lambda_{lj}^* \sigma_{lo}^* \Omega_o^{-1} (y^o - x\pi^o) + \xi_{lj}^*$$

are unity, where $\sigma_{lo}^* = E(\eta_l^* \eta_o^o)$, $\xi_{lj}^* = \lambda_{lj}^* \eta_l^* + \varepsilon_{lj}^* - \lambda_{lj}^* \sigma_{lo}^* \Omega_o^{-1} \eta_o^o$.

Furthermore, since x contains a constant term, we normalize $\alpha_{l1}^* = 0$ for $l = m+1, \dots, G$.

As in the usual simultaneous equation model, the structural parameters are not identifiable without imposing restrictions on the structure. We assume, as usual, that the rank conditions (see, e.g. Theil [1971]) are satisfied for the equations in (2.9). Under the rank condition, once the reduced form parameters π^* are identified, the structural parameters B and Γ will be identifiable. Since the equations (2.13) are probit regression equations with $\alpha_{l1}^* = 0$ and $\lambda_{l1}^* = 1$, the reduced form parameters π_l^* are identifiable under the general condition that $\text{plim}_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N x_i' x_i$ exists and is positive definite (c.f. Domencich and McFadden [1975]). Similarly, the parameters α_{lj}^* and λ_{lj}^* are identifiable. The identification of the remaining parameters, in the covariance matrix Σ and the variances $\sigma_{\epsilon_{lj}^*}^2$ of ϵ_{lj}^* depends crucially on the joint distribution of ϵ^* and u . For each l , let

$$\epsilon_l^{**} = \begin{pmatrix} \lambda_{l1}^* & & & \\ & \ddots & & \\ & & \lambda_{lJ_l}^* & \\ & & & \end{pmatrix} \begin{pmatrix} \tilde{\eta}_l^* \\ \vdots \\ \tilde{\eta}_l^* \end{pmatrix} + \begin{pmatrix} \epsilon_{l1}^* \\ \vdots \\ \epsilon_{lJ_l}^* \end{pmatrix}$$

where $\tilde{\eta}_l^* = \eta_l^* - \sigma_{\epsilon_l^*}^{-1} \eta_l^* 0$

This implies a factor structure,

$$(2.13) \quad \Sigma_{\epsilon_l^{**}} = \sigma_{\eta_l^*}^2 \Lambda_l \Lambda_l + \Sigma_{\epsilon_l^*}$$

where $\Lambda_l = \text{diag}(\lambda_{l1}^*, \dots, \lambda_{lJ_l}^*)$, $\sigma_{\eta_l^*}^2 = E(\tilde{\eta}_l^{*2})$ and $\Sigma_{\epsilon_l^*}$ is the variance matrix of $\epsilon_l^* = (\epsilon_{l1}^*, \dots, \epsilon_{lJ_l}^*)$.

Evidently the covariance matrix $\Sigma_{\epsilon_l^{**}}$ is identifiable from the multivariate probit equations (2.13). Since Λ_l are identifiable, the remaining number of unknown parameters on the right hand side is

$J_l(J_l + 1)/2 + 1$. Thus, the parameters $\sigma_{\eta_l^*}^2$ and $\Sigma_{\epsilon_l^*}$ are underidentified.

In Muthén (1976), he assumes that ϵ_{ls}^* , ϵ_{lq}^* , $s \neq q$ are mutually independent. Under this assumption, $\Sigma_{\epsilon_l}^*$ is a diagonal matrix and all the parameters can be identified; however, this assumption may be too strong for general purposes. Since our main interests are in the estimation of the structural coefficients B and Γ , these restrictive assumptions are unnecessary and will not be imposed.

3. Estimation and Hypothesis Test

We assume that the rank conditions are satisfied for our model to achieve identification of structural coefficients B and Γ . In addition to this identification condition and the conditions in (2.1), we assume that the explanatory variables x_i are random samples drawn from a distribution with a finite positive definite variance matrix

$$\Sigma_x. \text{ This implies } \text{plim}_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N x_i' x_i = \Sigma_x.$$

To estimate this model, the maximum likelihood procedure is computationally complicated and does not result in general or feasible analysis procedure. Alternative procedures which are computationally tractable are needed. For the model with a single indicator for each dependent variable, consistent and tractable estimation procedures have been proposed in Maddala and Lee [1976], Heckman [1978] and Amemiya [1979]. It has been shown in Amemiya [1979] and Lee [1980] that the estimates derived in Amemiya are more efficient than the others. For the multiple discrete indicators model, our procedure proposed derives from the application of the minimum $-\chi^2$ principle of estimation³, c.f. Neyman [1949], Taylor [1953] and Ferguson [1958].

³Other applications of this principle of estimation may include Zellner's [1962] approach to generalized least squares estimation in multivariate regression models with linear constraints on the regression coefficients; Malinvaud's [1970, ch. 9] approach to cover nonlinear constraints on the regression coefficients and Jöreskog and Goldberger's [1972] generalized least squares approach to factor analysis and others. Amemiya's [1979] approach is also related to this principle.

To simplify the expressions, we illustrate the procedure in a two equation model with one unobservable dependent variable and two ordered polychotomous indicators for this variable. The procedure is quite flexible and can immediately be extended to models with an arbitrary number of equations and discrete indicators. Specifically, the structural equations are

$$(3.1) \quad y_1 + \beta_{12}y_2^* = x_1\gamma_1 + u_1$$

$$\beta_{21}y_1 + y_2^* = x_2\gamma_2 + u_2$$

and the measurement equations for y_2^* are

$$(3.2) \quad z_j^* = \alpha_j + \lambda_j y_2^* + \epsilon_j \quad j=1,2$$

where $\alpha_1 = 0$ and the unknown coefficients have been normalized as in previous sections.

Let I_j be the ordered polychotomous indicator related to z_j^* with K_j categories as specified in (2.11). Equations (3.1) and (3.2) imply the following set of block reduced form equations,

$$(3.3) \quad y_1 = x\pi_1 + \eta_1$$

$$(3.4) \quad z_1^* = x\pi_2 + \sigma_{12}\sigma_1^{-2}(y_1 - x\pi_1) + \xi_1$$

$$(3.5) \quad z_2^* = \alpha_2 + \lambda_2 x\pi_2 + \lambda_2 \sigma_{12}\sigma_1^{-2}(y_1 - x\pi_1) + \xi_2$$

where $\text{var}(\xi_1) = \text{var}(\xi_2) = 1$, $\sigma_1^2 = E(\eta_1^2)$ and $\sigma_{12} = E(\eta_1 \eta_2)$.

The estimation procedure is divided into three stages:

Stage 1: Rearrange equations (3.4) and (3.5) into

$$(3.4)' \quad z_1^* = y_1 \delta_{11} + x \delta_{12} + \xi_1$$

$$(3.5)' \quad z_2^* = y_1 \delta_{21} + x \delta_{22} + \xi_2$$

Equation (3.3) can be estimated by ordinary least squares (OLS) which gives

$$(3.6) \quad \hat{\pi}_1 = (\sum_{i=1}^N x_i' x_i)^{-1} \sum_{i=1}^N x_i' y_{1i}$$

with variance $E(\hat{\pi}_1 - \pi_1)(\hat{\pi}_1 - \pi_1)' = \sigma_{\eta_1}^2 E(\sum_{i=1}^N x_i' x_i)^{-1}$. $\sigma_{\eta_1}^2$ can be

estimated by $\hat{\sigma}_{\eta_1}^2 = \frac{1}{N} \sum_{i=1}^N (y_{1i} - x_i' \hat{\pi}_1)^2$. Each equation in (3.4)' and (3.5)' is an ordered polychotomous probit equation, which can be estimated by maximum likelihood procedures (c.f. McKelvey and Zaviona [1975]). Corresponding to equation (3.4)', define K_1 dichotomous variables for each sample i ,

$$d_{i1j} = \begin{cases} 1 & \text{if } I_{1i} = j \\ 0 & \text{otherwise} \end{cases} \quad j=2, \dots, K_1$$

and $d_{i11} = 1 - d_{i12} - \dots - d_{i1K_1}$.

Let $\phi(v)$ be the standard normal distribution function evaluated at v . The log likelihood function is

$$(3.7) \quad \ln L_1(\theta_1) = \sum_{i=1}^N \sum_{j=1}^{K_1} d_{i1j} \ln(\phi(\mu_{j-1,1} + r_i) - \phi(\mu_{j-2,1} + r_i))$$

where $r_i = -(y_{i1} \delta_{11} + x_i' \delta_{12})$, $\theta_1' = (\delta_{11}, \delta_{12}', \mu_{11}, \dots, \mu_{K_1-2,1})$,

$\mu_{-1,1} = -\infty$, $\mu_{01} = 0$ and $\mu_{K_1-1,1} = \infty$.

$$(3.8) \quad \ell_{1i} = \sum_{j=1}^{K_1} d_{i1j} \ln(\phi(\mu_{j-1,1} + r_i) - \phi(\mu_{j-2,1} + r_i))$$

The estimates of θ_1 can be derived by maximizing (3.7). This procedure generalizes the probit maximum likelihood procedure (Finney [1947]).

Under the assumptions that the parameter spaces are compact, the vector of true parameters is an interior point and the other assumptions already specified, one can easily show (see, e.g. Jennrich [1969]) that

the maximum likelihood estimates $\hat{\theta}_1$ of θ_1 from (3.7) are asymptotically normally distributed,

$$(3.9) \quad \sqrt{N}(\hat{\theta}_1 - \theta_1) \xrightarrow{D} N(0, \Omega_{\theta_1})$$

where $\Omega_{\theta_1} = \text{plim}_{N \rightarrow \infty} \left(\frac{1}{N} \sum_{i=1}^N \frac{\partial \ell_{1i}}{\partial \theta_1} \frac{\partial \ell_{1i}}{\partial \theta_1} \right)^{-1}$ and

$$(3.10) \quad \hat{\delta}_1 - \delta_1 \xrightarrow{D} J_1 \left(\sum_{i=1}^N \frac{\partial \ell_{1i}}{\partial \theta_1} \frac{\partial \ell_{1i}}{\partial \theta_1} \right)^{-1} \sum_{i=1}^N \frac{\partial \ell_{1i}}{\partial \theta_1}$$

where " $\xrightarrow{D}$ " means that both sides have the same asymptotic distribution and J_1 is defined from $\delta_1 = J_1 \theta_1$. Similarly we can estimate equation (3.5)'. Let δ_2 denote the vector of coefficient in (3.5)' and $\hat{\delta}_2$ be the ordered probit estimates. We have

$$(3.11) \quad \hat{\delta}_2 - \delta_2 \xrightarrow{D} J_2 \left(\sum_{i=1}^N \frac{\partial \ell_{2i}}{\partial \theta_2} \frac{\partial \ell_{2i}}{\partial \theta_2} \right)^{-1} \sum_{i=1}^N \frac{\partial \ell_{2i}}{\partial \theta_2}$$

Let $\hat{\omega}' = (\hat{\pi}_1', \hat{\delta}_1', \hat{\delta}_2')$ be the estimates of $\omega' = (\pi_1', \delta_1', \delta_2')$ derived above. Evidently, $\sqrt{N}(\hat{\omega} - \omega) \xrightarrow{D} N(0, \Omega_\omega)$. By the strong law of large number, Ω_ω can be consistently estimated by W .

$$(3.12) \quad W = \begin{pmatrix} W_{\pi_1} & 0 & 0 \\ 0 & W_{\delta_1} & W_{\delta_1 \delta_2} \\ 0 & W_{\delta_2 \delta_1} & W_{\delta_2} \end{pmatrix}$$

where $W_{\pi_1} = \hat{\sigma}_\eta^2 \left(\frac{1}{N} \sum_{i=1}^N x_i' x_i \right)^{-1}$, $W_{\delta_1} = J_1 \left(\frac{1}{N} \sum_{i=1}^N \frac{\partial \ell_{1i}}{\partial \theta_1} \frac{\partial \ell_{1i}}{\partial \theta_1} \right)^{-1} J_1$,

$$W_{\delta_2} = J_2 \left(\frac{1}{N} \sum_{i=1}^N \frac{\partial \ell_{2i}}{\partial \theta_2} \frac{\partial \ell_{2i}}{\partial \theta_2} \right)^{-1} J_2 \quad \text{and}$$

$$W'_{\delta_1 \delta_2} = W_{\delta_2 \delta_1} = J_2 \left(\frac{1}{N} \sum_{i=1}^N \frac{\partial \ell_{2i}}{\partial \theta_2} \frac{\partial \ell_{2i}}{\partial \theta_1'} \right)^{-1} \frac{1}{N} \sum_{i=1}^N \frac{\partial \ell_{2i}}{\partial \theta_2} \frac{\partial \ell_{1i}}{\partial \theta_1'} \left(\frac{1}{N} \sum_{i=1}^N \frac{\partial \ell_{1i}}{\partial \theta_1} \frac{\partial \ell_{1i}}{\partial \theta_1'} \right)^{-1} J_1$$

evaluated at $\hat{\omega}$.

Stage 2. In this stage, we will estimate the reduced form parameters π .

Let $\tau' = (\pi_1', \pi_2', \alpha_2, \lambda_2, \sigma_{12}\sigma_1^{-2})$. ω defined above is a function of τ , i.e.

$$(3.13) \quad \omega = h(\tau)$$

Explicitly, these relations are

$$(3.14) \quad \begin{aligned} \pi_1 &= \pi_1 \\ \delta_{11} &= \sigma_{12}\sigma_1^{-2} \\ \delta_{12} &= \pi_2 - \sigma_{12}\sigma_1^{-2}\pi_1 \\ \delta_{21} &= \lambda_2\sigma_{12}\sigma_1^{-2} \\ \delta_{22} &= \alpha_2 e_1 + \lambda_2\pi_2 - \lambda_2\sigma_{12}\sigma_1^{-2}\pi_1 \end{aligned}$$

where $e_1' = (1, 0, \dots, 0)$. To estimate τ , we suggest the following nonlinear least squares procedure (NLS),

$$\min_{\tau} (\hat{\omega} - h(\tau))' W^{-1} (\hat{\omega} - h(\tau))$$

It can be easily shown as in Jennrich [1969] that $\hat{\tau}$ is strongly consistent and

$$\sqrt{N}(\hat{\tau} - \tau) \xrightarrow{D} N(0, (\nabla h'(\tau) \Omega_{\omega}^{-1} \nabla h(\tau))^{-1})$$

where $\nabla h(\tau) = \partial h(\tau) / \partial \tau'$. An initial consistent estimate $\tilde{\tau}$ can be derived as follows. From the first stage estimation, we have already derived estimates $\hat{\pi}_1, \hat{\delta}_{11}$. The remaining parameters can be estimated as (if $\hat{\delta}_{11} \neq 0$),

$$\tilde{\pi}_2 = \hat{\delta}_{12} + \hat{\delta}_{11}\hat{\pi}_1$$

$$\tilde{\lambda}_2 = \hat{\delta}_{21}/\hat{\delta}_{11}$$

$$\tilde{\alpha}_2 = e_1'(\hat{\delta}_{22} - \tilde{\lambda}_2\hat{\delta}_{12}).$$

These estimates can then be used as initial estimates for the iterations.⁴ In general, the second stage estimates of π_1 will improve on the first stage estimate of π_1 with the following exception. For the case in which there are no reduced form constraints on π_2 , the second stage estimate of π_1 will be the same as the first stage estimate $\hat{\pi}_1$ of π_1 because the equations $\hat{\delta}_1, \hat{\pi}_1$ are uncorrelated and there is no information about π_1 that can be attained from δ 's.

Step 3. In this stage, we will estimate the structural parameters B and Γ from $\pi = [\pi_1, \pi_2]$. From the second stage estimation, we have an estimate of $\text{vec}'\pi \equiv (\pi_1', \pi_2')$. Let L be the matrix such that $\text{vec}\pi = L\tau$ and let $\text{vec}\hat{\pi} = L\hat{\tau}$. Define a quadratic function

⁴A relatively simpler procedure is to linearize $h(\tau)$ at an initial consistent $\tilde{\tau}$ and estimate τ by generalized least squares. Linearizing $h(\tau)$ at $\tilde{\tau}$, it implies

$$\hat{\omega} - h(\tilde{\tau}) = \nabla h(\tilde{\tau})\tau + (\hat{\omega} - \omega)$$

and τ can be estimated as

$$\hat{\tau} = (\nabla h'(\tilde{\tau})W^{-1}\nabla h(\tilde{\tau}))^{-1}\nabla h'(\tilde{\tau})W^{-1}(\hat{\omega} - h(\tilde{\tau}))$$

This procedure is known as the method of scoring. The estimate $\hat{\tau}$ have the same limiting distribution as $\hat{\tau}$.

$$(3.16) \quad Q_V = \text{vec}'(\hat{\pi}B - \Gamma)V^{-1}\text{vec}(\hat{\pi}B - \Gamma)$$

where V is a positive definite matrix. If V is a known matrix, minimizing Q_V with respect to B and Γ will give estimates linearly in $\hat{\pi}$. Chosen

$$(3.17) \quad V = (B \otimes I)(Vh'(\tilde{\tau})W^{-1}Vh(\tilde{\tau}))^{-1}(B' \otimes I),$$

it implies

$$(3.18) \quad \begin{aligned} Q_V &= \text{vec}'(\hat{\pi}B - \Gamma)(B^{-1} \otimes I)(Vh'(\tilde{\tau})W^{-1}Vh(\tilde{\tau}))(B'^{-1} \otimes I) \text{vec}(\hat{\pi}B - \Gamma) \\ &= \text{vec}'(\hat{\pi} - \Gamma B^{-1})(Vh'(\tilde{\tau})W^{-1}Vh(\tilde{\tau}))\text{vec}(\hat{\pi} - \Gamma B^{-1}). \end{aligned}$$

As V in (3.17) contains unknown parameters B , Q_V in (3.18) is highly nonlinear. Instead of minimizing this highly nonlinear function, we suggest a modified procedure. First, we estimate B and Γ by minimizing Q_I ,

$$(3.19) \quad Q_I = \text{vec}'(\hat{\pi}B - \Gamma)\text{vec}(\hat{\pi}B - \Gamma)$$

subject to the fixed parameters and exclusion constraints imposed on B and Γ .

Denote this estimate of B as $\tilde{B}$. Substituting $\tilde{B}$ into (3.17), we have $\tilde{V} = (\tilde{B} \otimes I)(Vh'(\tilde{\tau})W^{-1}Vh(\tilde{\tau}))^{-1}(\tilde{B}' \otimes I)$ and then minimize

$$(3.20) \quad Q_V^* = \text{vec}'(\hat{\pi}B - \Gamma)\tilde{V}^{-1}\text{vec}(\hat{\pi}B - \Gamma)$$

w.r.t. B and Γ subject to the fixed parameters and exclusion constraints imposed on B and Γ .

The estimation procedure in (3.19) is similar to procedures suggested in Amemiya for Heckman's model. The estimation procedure in (3.20) generalizes his procedure to the joint estimation of a system of equations. These similarities can be shown as follows. Define matrices L_1 and L_2

from $x_1 = xL_1$ and $x_2 = xL_2$. It follows $\Gamma = [L_1\gamma_1, L_2\gamma_2]$ and

$$(3.21) \quad [\pi_1, \pi_2] \begin{pmatrix} 1 & \beta_{21} \\ \beta_{12} & 1 \end{pmatrix} = [L_1\gamma_1, L_2\gamma_2]$$

Equivalently,

$$(3.22) \quad \hat{\pi}_1 = -\hat{\pi}_2 \beta_{12} + L_1\gamma_1 + \xi_1$$

$$(3.23) \quad \hat{\pi}_2 = -\hat{\pi}_1 \beta_{21} + L_2\gamma_2 + \xi_2$$

where $\xi = (\xi_1', \xi_2')' = (B' \otimes I) \begin{pmatrix} \hat{\pi}_1 - \pi_1 \\ \hat{\pi}_2 - \pi_2 \end{pmatrix}$. The asymptotic variance matrix of ξ is

$$(3.24) \quad V_\xi = (B' \otimes I) \frac{1}{N} Vh'(\tau) \Omega_\omega^{-1} Vh(\tau) (B \otimes I)$$

Amemiya's suggestion amounts to estimating equations (3.22) and (3.23)

by ordinary least squares. Obviously this is equivalent to minimizing

Q_I . A more efficient procedure is to estimate the equations (3.22) and

(3.23) jointly by generalized least squares (GLS) with an estimated $\tilde{V}_\xi$,

i.e. $\tilde{V}_\xi = (\tilde{B}' \otimes I) \frac{1}{N} Vh'(\tilde{\tau}) \tilde{W}^{-1} Vh(\tilde{\tau}) (\tilde{B} \otimes I)$, substituted for V_ξ . Evidently,

this is the same procedure as minimizing $Q_{\tilde{V}}$ in (3.20). Let

$$G = \begin{pmatrix} -\hat{\pi}_2 & L_1 & 0 & 0 \\ 0 & 0 & -\hat{\pi}_1 & L_2 \end{pmatrix}$$

The GLS estimate $\hat{\theta}$ of $\theta' = (\beta_{12}', \gamma_1', \beta_{21}', \gamma_2')$ is

$$(3.25) \quad \hat{\theta} = (G' \tilde{V}_\xi^{-1} G)^{-1} G' \tilde{V}_\xi^{-1} \text{vec}(\hat{\pi})$$

which is asymptotic normal $N(\theta, \Sigma)$. The asymptotic variance matrix

$\frac{1}{N}\Sigma$ of $\hat{\theta}$ can be estimated by $\frac{1}{N}\hat{\Sigma} = (G' \tilde{V}_\xi^{-1} G)^{-1}$.

Using these estimates and their asymptotic distributions, hypothesis testing is straightforward. Suppose we would like to test the null

hypothesis, $H_0: g(\theta) = 0$, where $g(\theta)$ is a r vector and $\partial g(\theta)/\partial \theta'$ has rank r at the true parameter, the statistic

$$(3.26) \quad Ng'(\hat{\theta}) \left[\frac{\partial g(\hat{\theta})}{\partial \theta'} \hat{\Sigma} \frac{\partial g'(\hat{\theta})}{\partial \theta} \right]^{-1} g'(\hat{\theta})$$

can be used. As $\hat{\theta}$ is asymptotically normal, the statistic in (3.26) is asymptotic chi-square with r degrees of freedom. In addition to testing the significance of coefficients in the model, we may be interested in testing the compatibility of the measurement equations. In the specification of the measurement equations (2.10), we have implicitly assumed that the measurement errors are nonsystematic. If there are systematic measurement errors present in the indicator, the measurement equations will not be compatible. Specifically, in addition to y_{lj}^{**} , if there is a systematic measurement bias related to x , the measurement equation will be

$$(3.27) \quad z_{lj}^{**} = \alpha_{lj}^* + \lambda_{lj}^* y_{lj}^{**} + x_{lj} \mu_{lj} + u.$$

If the specified measurement equations in (3.2) are compatible, they imply cross equations constraints as in (3.4) and (3.5). If the measurement equations are equations as in (3.27), the cross equation constraints in (3.4) and (3.5) will not hold. Thus to test the compatibility of the measurement equations, we suggest testing the validity of the cross equations constraints. The constraints in this model are expressed as $\omega = h(\tau)$ in (3.13). Thus the null hypothesis for measurement compatibility is $H_0: \omega = h(\tau)$. To test this null hypothesis, we can use the statistic

$$(3.28) \quad N(\hat{\omega} - \omega^*)' W^{-1} (\hat{\omega} - \omega^*)$$

where $\hat{\omega}$ is the unconstrained estimates derived from stage 1 and $\omega^* = h(\hat{\tau})$ constrained estimate of ω with the estimates $\hat{\tau}$ derived from the second stage estimation. Under the null hypothesis H_0 ,

$$\begin{aligned}\sqrt{N}(\hat{\tau}-\tau) &\stackrel{D}{=} [\nabla' h(\tau) \Omega_{\omega}^{-1} \nabla h(\tau)]^{-1} \nabla' h(\tau) \Omega_{\omega}^{-1} (\hat{\omega}-h(\tau)) \\ &\stackrel{D}{=} [\nabla' h(\tau) \Omega_{\omega}^{-1} \nabla h(\tau)]^{-1} \nabla' h(\tau) \Omega_{\omega}^{-1} (\hat{\omega}-\omega)\end{aligned}$$

The Taylor expansion about $h(\tau)$ yields $h(\hat{\tau}) = h(\tau) + \nabla h(\tau^+)(\hat{\tau}-\tau)$, where τ^+ lies between $\hat{\tau}$ and τ . Under the null hypotheses, τ^+ is consistent estimate of τ . Hence

$$\begin{aligned}\sqrt{N}(\hat{\omega} - \omega^*) &= \sqrt{N}(\hat{\omega} - h(\hat{\tau})) \\ &\stackrel{D}{=} \sqrt{N}(\hat{\omega} - \omega - \nabla h(\tau)(\hat{\tau} - \tau)) \\ &\stackrel{D}{=} [I - \nabla h(\tau)(\nabla' h(\tau) \Omega_{\omega}^{-1} \nabla h(\tau))^{-1} \Delta' h(\tau) \Omega_{\omega}^{-1}] \sqrt{N}(\hat{\omega}-\omega),\end{aligned}$$

and asymptotically,

$$\begin{aligned}N(\hat{\omega} - \omega^*)' W^{-1}(\hat{\omega} - \omega^*) \\ &\stackrel{D}{=} \sqrt{N}(\hat{\omega} - \omega)' \Omega_{\omega}^{-1/2} [I - \Omega_{\omega}^{-1/2} \nabla h(\tau)(\nabla' h(\tau) \Omega_{\omega}^{-1} \nabla h(\tau))^{-1} \nabla' h(\tau) \Omega_{\omega}^{-1/2}] \\ &\quad \Omega_{\omega}^{-1/2} \sqrt{N}(\hat{\omega} - \omega)\end{aligned}$$

is $\chi^2(q-p)$ where q is the dimension of the vector ω and p is the number of free parameters in τ . (See Graybill [1961]).

To correct for incompatibility of measurements, one need to identify the potential explanatory variables which characterize the measurement biases. If there is a systematic measurement bias related to x , the measurement equations should be corrected as in (3.27). In general, even the explanatory variables x in (3.27) contain explanatory variables in the structural equations, the reduced form parameters π and hence the structural equations parameters may still be identifiable unless there are measurement biases in every measurement equation. Accordingly, the second and third estimations should be modified.

4. An Empirical Application -- Health and Wage

4.1 Model and Data.

To illustrate the estimation procedures, we apply them to study a simultaneous equation model with health and wages jointly dependent. In Grossman [1972a, 1972b], health is viewed as an endogenously determined capital stock. The stock of health is a theoretical concept which is difficult to quantify empirically; however, various indicators are available. These include an individual's evaluation of the state of his health, health limitations,⁵ etc. In most of the existing empirical works, e.g. Grossman and Benham [1974], Luft [1975, 1978] and Grossman [1976] these indicators are treated as continuous proxy variables for the unobservable health stock and models are estimated by ordinary least squares or two stage least squares. These approaches are computationally simple but involve some specification problems. Those indicators are either discrete or limited dependent. It is well known (see, e.g. Goldberger [1964], pp. 248-255) that linear regression approaches create various specification and estimation problems. These specifications may become more problematic for simultaneous equation models. Another problem is the quality of the various health indicators. The self-evaluation health indicators may represent information bias about health as well as measurement of health capitals. Health economists recognize this problem (c.f. Luft [1975]) but no satisfactory solutions have been proposed. These problems, however, can be easily handled by our approach.

⁵Other indicators may be of work-loss weeks due to illness, number of symptoms, etc.

The health model we consider consists of two structural equations with unobservable health capital and wage rate jointly dependent. Let H^* be the unobservable health stock of an individual and W be the natural logarithmic wage rate. The two structural equations in the model are as follows,

$$(4.1) \quad H^* = W\beta_{12} + x_1\gamma_1 + u_1$$

$$(4.2) \quad W = H^*\beta_{21} + x_2\gamma_2 + u_2$$

where H^* and W are endogenous variables and x_1, x_2 are vectors of explanatory variables. A complete list of variables used in this model is given in table 1.

TABLE 1
LIST OF VARIABLES USED IN THE MODEL

<u>Variables</u>	<u>Definition</u>
$x_{11} = x_{21}$	Constant 1 (CONSTANT)
$x_{12} = x_{22}$	Highest grade completed in years, 1966 (EDUCATION)
$x_{13} = x_{23}$	Age of respondent, 1966 (AGE)
$x_{14} = x_{24}$	Marital status of respondent, 1966 (MARRIED);
	$x_{15} = 1$ if married and spouse present,
	$x_{15} = 0$ otherwise
x_{15}	Total net family asset in thousand dollars, 1966 (ASSET)
x_{25}	Square of experience; $x_{25} = (\text{AGE} - \text{EDUCATION} - 6)^2/100$ (EXPSQ)
x_{26}	Current residence in SMSA, 1966; $x_{26} = 1$ if in SMSA, $x_{26} = 0$ otherwise (SMSA)
x_{27}	Southern region dummy; $x_{27} = 1$ if in South Atlantic,

Table 1 (continued)

<u>Variables</u>	<u>Definition</u>
	E.S. Central, or W.S. Central
	$x_{27} = 0$ otherwise (SOUTH)
x_{28}	Race; $x_{28} = 1$ if white, $x_{28} = 0$ for others (RACE)
W	Hourly rate of pay at current or last job in nature logarithm scale of cents per hour, 1966
H^*	Unobservable health stock
I_2	Health in comparison to other men the age of respondent, self-rating, 1966 (INDICATOR 2); $I_2 = 1$ if poor, $I_2 = 2$ if fair, $I_2 = 3$ if good, $I_2 = 4$ if excellent
I_1	(INDICATOR 1); $I_1 = 1$ if health does not affect kind or amount of work, 1966; $I_1 = 0$ no health limitation.

The data we used are from the National Longitudinal Survey of Men 45 - 59, Survey year 1966. The sample used is restricted to middle age males, 45 to 59 years of age, with positive earnings in 1966. Samples with incomplete observations in our list of explanatory variables are omitted. The final sample size is 2876. In this sample, two health indicators are available. One indicator is the individuals' self-evaluation of his general health as excellent, good, fair or poor as compared to other men of the same age. This indicator is an ordered polychotomous variable normalized with respect to age. The other health indicator is related to health limitation. In this survey the individual is asked whether his health prevents working, limits the kind of work, limits the amount of work or imposes no limitations, mutually exclusively. Since our sample is limited to individuals with positive earnings, individuals who are unable to work have been excluded. This might create a selection problem; however, we

don't expect this to create a serious problem since these individuals compose only 3.7 percent of the survey (see the National Longitudinal Survey [1966]).⁶ This indicator is treated as a dichotomous variable.⁷

⁶This potential problem might deserve some comments. Trying to incorporate the unable to work individuals into the model creates some conceptual and specification difficulties. In Heckman [1974], he has formulated a labor supply model to correct for potential wage selection bias for the female labor supply. However, his approach is not quite appropriate for modeling the "unable to work individuals" labor supply. Potential market wages cannot be meaningfully defined for the men who are unable to work since "health prevents working" is unlikely to be a temporary event. In the NLS male 45-59 samples, the mean years of duration of "health preventing work" is about ten years. Another difficulty is the specification of the market experience variable (see footnote 9). The market experience for the unable to work individuals may be specified as (age-schooling-6-D) where D is the year of duration of health preventing work. In retaining unable to work individuals in the sample, the health limitation indicators I_1 will be a three level, ordered, discrete variable. Suppose $I_1 = 0$ denote the unable to work category. However, since $D \neq 0$ if and only if $I_1 = 0$, the inclusion of variable D in the model would create tautological problems. Fitting an ordered probit equation with I_1 as dependent variable and D as explanatory variable, the coefficient of D will be negative infinity. Due to these difficulties and the small proportion of samples in this category, such observations are excluded from our sample.

⁷One may treat this indicator as an ordered polychotomous variable. However, we are not quite convinced that "health affects kind of work" would represent a poorer health condition than "health affects amount of

The explanatory variables in the health structural equation consist of CONSTANT, EDUCATION, AGE, MARRIED and ASSET. The explanatory variables in the health equation are variables that may affect both the demand for health as well as the production of health (Grossman and Benham [1974]). On the production side, health is expected to respond positively to the "years of schooling completed" as schooling may increase the production efficiency of health and fall with age as the rate of health depreciation may rise with age for middle age individuals. On the demand side, the demand for health should rise with the wage rate as the rate of return to health investment should be positively related to the wage rate. The demand for health will rise with "net family assets" since good health is expected to increase individual's well being. On the production side, it is a proxy for inputs in the production of health.⁸ The explanatory variables in the wage generating function consist of CONSTANT, EDUCATION, AGE, MARRIED, EXPSQ, SMSA, SOUTH and RACE. The specification of the wage generating function is standard in the human capital literature (c.f. Becker [1967], Mincer [1970]). Wage rates depend on investments in human capital as measured by investment in formal schooling and on the job training. Hence schooling and market experience have positive effects on

7 (continued) work" or vice versa. Treating it as a dichotomous variable may result in loss of efficiency but avoids the problems created by classification errors.

⁸The net family asset is used as a proxy for inputs, in the production of health which may include utilization of preventive medical services, health insurance coverage, recreation, quality of the living environment, etc. Specific information about these variables are not available from the NLS survey.

wages.⁹ The marital status variable is also related to human capital investment since married men specialize in the labor market more than do single men and accordingly have a greater investment in human capital (c.f. Kenny et. al. [1979]). The regional variables, SOUTH and SMSA are proxies for the differential cost of living in these regions. The cost of living is usually higher in more densely large populated regions; consequently, the coefficients are expected to be negative for southern regions and positive for SMSA. The race variable is related to racial discrimination in labor markets. Health capital is included as an explanatory variable for wages as it should raise market productivity and hence wages. From this specification, the structural health equation is exactly identified as the only explanatory variable excluded from this equation is the net family asset variable. On the other hand, the structural health equation is overidentified as the variables EXPSQ, SMSA, SOUTH and RACE are excluded from the wage generating equation.

The measurement equation corresponding to INDICATOR 1 is specified as in (3.2) with $\alpha_1 = 0$ and $\lambda_1 = 1$. The measurement equation corresponding to INDICATOR 2 needs to take into account that individual's health self-evaluations are normalized with respect to individual's age. Specifically, we specify the first measurement equation as

⁹ Actual market experiences for individual workers is hardly available in census surveys. As a conventional rule, one may use (age-schooling-6) as measure of market experience. In our wage equation, the age variable is entered instead of experience. The coefficient of age in the wage equation does represent the effect of experience, but the coefficient of schooling will be the difference of the effects of education and market experience.

$$(4.3) \quad z_1^* = H_1^* + \epsilon_1$$

and the second measurement equation as

$$(4.4) \quad z_2^* = \alpha_2 + \lambda_2(H^* - x_{13}\gamma_{13}) + \epsilon_2$$

where γ_{13} is the coefficient of the age variables x_{13} in the structural health equation. The second health indicator is "self-evaluation of health status" which may be subject to perceptual biases about health for different populations. To correct for this measurement problem, one needs to identify the potential explanatory variables which characterizes such measurement differentials. Once the variables R_2 are identified, they can be introduced into the measurement equation as

$$(4.5) \quad z_2^* = \alpha_2 + \lambda_2(H^* - x_{13}\gamma_{13}) + R_2\delta + \epsilon_2$$

4.2 Empirical Results

The structural equations (4.1), (4.2) with measurement equations (4.3) and (4.4) are estimated by the three stages procedure described in the previous sections. The results from the first stage estimation are presented in TABLE 2. In the second stage estimation, we have used the nonlinear least squares procedure.¹⁰ The results of the second stage estimation are presented in TABLE 3. In this estimation, we have taken into account the specific structure of the measurement equation in (4.4) and hence there is an estimate of the coefficient of age variable γ_{13}

¹⁰ In all the second stage estimation, we use Newton's iteration procedure with consistent initial estimates. The procedure converges rapidly at 7 or 8 iterations. The first iteration estimates are the estimates of method of scoring. It provides close approximates to the convergent estimates. However the derived chi-squares statistics are usually unnecessary large to be useful.

in the structural health equation. The signs of the coefficients of reduced form wage equation in the first and second stage estimations are the same as expected. The standard errors of the estimates are in the brackets. The third stage estimates are reported in TABLE 4.

It is interesting to note that the estimated coefficient for the wage rate in the structural health equation in TABLE 4 is significantly positive; that is, the wage rate increases the demand for health. Another significant coefficient in the health equation is the effect of aging on health. The health capital stock of a middle age male decreases as the individual grows older. The effect of education is positive but only marginally significant at the 10 percent level of significance. The net family asset variable has the expected positive effect on the demand as well as production of health capital but is not significantly different from zero. However, even though the effect is not strong in the structural equation of health, its net effects as derived in the reduced form equations in TABLE 3 are quite strong. The coefficient of marital status is negative but not statistically different from zero. Comparing the t values, wage rate has the stronger effect on health than education. Our results thus differ from the conclusions of Grossman and Benham [1974] where schooling had stronger effect on health than wages.---

¹¹ It should be noted that the estimated coefficient of schooling (EDUCATION) would be biased upward if past health capital during schooling ages and highest grade completed were positive correlated, and past health had an effect on the current health equation (Grossman [1976]). Thus the effect of schooling might even be weaker. Information on workers' past health during schooling ages was not available in the NLS sample.

TABLE 2

FIRST STAGE ESTIMATION*

Ordered Probit Maximum Likelihood Estimates

	INDICATOR 1	INDICATOR 2	W
CONSTANT	-.53588 (1.16108)	1.93876 (.91447)	5.65524 (.34153)
EDUCATION	-.02109 (.04468)	.08985 (.03519)	.07265 (.01369)
AGE	.00527 (.04467)	-.07187 (.03523)	-.03569 (.01376)
MARRIED	-.01307 (.08204)	-.09425 (.06465)	.25786 (.02533)
ASSET	.00051 (.00076)	.00125 (.00060)	.00175 (.00018)
EXPSQ	-.03926 (.06024)	.08062 (.04764)	.04121 (.01869)
SMSA	.09369 (.05998)	.15338 (.04745)	.26089 (.01836)
SOUTH	.11098 (.06276)	.04347 (.04893)	-.16206 (.01914)
RACE	-.21916 (.06713)	-.00952 (.05185)	.20278 (.02042)
W	.32937 (.05899)	.31092 (.04721)	
		$\mu_1 = 1.03300$ (.04569)	$\sigma^2 = .1957$
		$\mu_2 = 2.28526$ (.05101)	
-2 Log Likelihood Function	-1467.9491	-3157.2675	
-2 Log Likelihood Ratio (with 9 degrees of freedom)	80.3307	207.1084	

* Standard errors are in brackets.

TABLE 3
SECOND STAGE ESTIMATION

	Reduced H [*]	Reduced W
CONSTANT	2.43045 (.55707)	5.65520 (.34154)
EDUCATION	.05701 (.02158)	.07265 (.01369)
AGE	-.05576 (.02083)	-.03569 (.01376)
MARRIED	-.00565 (.03743)	.25786 (.02533)
ASSET	.00102 (.00035)	.00175 (.00018)
EXPSQ	.04307 (.02821)	.04121 (.01869)
SMSA	.12960 (.03295)	.26089 (.01836)
SOUTH	-.00299 (.02851)	-.16206 (.01914)
RACE	.02673 (.03047)	.20278 (.02042)
	$\gamma_{13} = -.01542 \text{ (.00581)}$ $\sigma_{12}\sigma_1^{-2} = .18146 \text{ (.03925)}$ $\alpha = -.77216 \text{ (.70973)}$ $\lambda = 1.69786 \text{ (.24399)}$	

Test of measurement compatibility $\chi^2(7) = 28.345$

TABLE 4

THIRD STAGE ESTIMATION

	Structural H [*]	Structural W
CONSTANT	.49066 (.50306)	1.42018 (1.50026)
EDUCATION	.00914 (.00553)	.00036 (.03348)
AGE	-.01922 (.00567)	.03890 (.03387)
MARRIED	-.06512 (.04245)	.27341 (.06520)
ASSET	.00032 (.00034)	
EXPSQ		.00388 (.02597)
SMSA		.09531 (.07785)
SOUTH		-.09770 (.02295)
RACE		.11079 (.03058)
W	.22402 (.07549)	
H [*]		2.22232 (.56691)

For the estimates of the structural wage rate equation, the coefficients of health capital, marital status, race and south variables are significant. The signs of these effects are correct as theoretically predicted. Better health increases market productivity and hence wages. Married males with spouse present receive 27 percent higher wages. White workers receive 11 percent higher wages than nonwhite with the same years of schooling. Workers in southern regions receive 9.8 percent lower wages than workers in other regions. The remaining explanatory variables are not statistically significant. Workers living in SMSA receive higher wages. One more year of market experience increases wages by 3.9 percent as measured by the coefficient of age. Schooling has essentially the same rate of return as market experience. The rate of return for this middle age cohort is less than the rate of return for younger cohorts in other studies. In Kenny et. al. [1979], the estimates rate of return for high school graduates in 1960 is about eleven percent. The differences for the two cohorts might be explained by differences of quality of education received by the cohorts. The coefficient of squared market experience is essentially zero. As the dependent variable W is measured in logarithmic scale, the estimated earning profile is increasing and concave. In summary, these empirical results are compatible with the theoretical implications of human capital theories (Grossman [1972b] and Becker [1967]).

Let us now consider the problem of compatibility of the two health indicators. From the estimates of λ , at the bottom of TABLE 3, the estimated scale coefficient λ is significant. Thus the self-evaluation health indicator does contain valuable information for the unobservable health capital. However, the Chi-square test of measurement compatibility, $\chi^2 = 28.345$ with seven degrees of freedom, does indicate that there is a systematic measurement bias in self-evaluation of health. This measurement

bias may reflect perceptual problems about health. We suspect that populations of different race or with different levels of education may have different perceptions about health.¹² To correct for this, we modify the measurement equation (4.4) into equation (4.5). This results in modification of the second and third stage estimations for our model. The results are reported in TABLE 5 and TABLE 6. After controlling for measurement differences in health evaluation by education and race, the Chi-square test of measurement compatibility is $\chi^2 = 7.58$ with five degrees of freedom. This is not significantly different from zero and hence the modified measurement equations are compatible. The estimated scale coefficient λ which is again significant and is not different from unity, indicates that the health self-evaluation indicator contains valuable information about health capital in addition to the health limitation indicator.

Comparing TABLE 4 and TABLE 6, the notable differences are that education and wages have weaker effects (in term of t values) in the demand and production for health capital. Other than that there are no significant differences. Health capital has similar effect on wages when the potential health self-evaluation measurement bias are corrected. The variables that have significant coefficients are the same in both sets of estimates in TABLE 4 and TABLE 6. Married workers with spouse present have 26 percent higher wages than others. White workers receive 19 percent more wages than nonwhites. Workers living in southern regions have 11 percent lower wages. Except for these slight differences in estimated magnitudes, the qualities and the implications of the estimated effects are hardly changed.

¹²As previously indicated, we have chosen the health limitation indicator as the standard one since we are interested in the aspect of health capital on market productivity. Implicitly, we assume that the health limitation indicator is an objective measure and health self-evaluation is a subjective measure of health.

TABLE 5
SECOND STAGE ESTIMATION
(Measurement Compatibility Correlated)

	Reduced H [*]	Reduced W
CONSTANT	2.92587 (.77817)	Same as Table 3
EDUCATION	.06399 (.03004)	
AGE	-.06975 (.03018)	
MARRIED	.00476 (.05342)	
ASSET	.00141 (.00046)	
EXPSQ	.05936 (.04089)	
SMSA	.19820 (.04530)	
SOUTH	.00717 (.04120)	
RACE	-.14994 (.06422)	
$\gamma_{13} = -.01402 \text{ (.00596)}$ $\sigma_{12}^{-2} \sigma_1^2 = .28574 \text{ (.05243)}$ $\lambda = 1.13667 \text{ (.15554)}$		
Systematic Reporting in Self-Evaluation		
CONSTANT α	-.12446 (.45062)	
EDUCATION	.02042 (.00880)	
RACE	.22121 (.06980)	
$\chi^2(5) = 7.58$		

TABLE 6

STAGE 3 ESTIMATION

(Measurement Compatibility Corrected)

	Structural H [*]	Structural W
CONSTANT	.36472 (.62771)	2.23857 (1.23747)
EDUCATION	.00165 (.01012)	.02251 (.03016)
AGE	-.02002 (.00566)	.02717 (.03043)
MARRIED	-.07153 (.06017)	.25906 (.06744)
ASSET	.00052 (.00047)	
EXPSQ		.00509 (.02798)
SMSA		.09998 (.08463)
SOUTH		-.11204 (.02742)
RACE		.19232 (.07099)
W	.26748 (.11353)	
H [*]		1.60510 (.40095)

5. Conclusions

In this article, we have proposed an econometric approach for estimating the simultaneous effects of health and wages in a structural equations framework. Our approach takes into account that health capital is an unobservable and the health indicators are discrete variables. Specifically, we generalize Heckman's Dummy Endogenous Variable Simultaneous Equation System [1978] to an equation system with multiple ordered polychotomous indicators. An estimation procedure based on minimum χ^2 principle is suggested. This procedure is computationally tractible and results in generally feasible analysis procedure. We have also considered the problem of measurement compatibility of the indicators. An χ^2 test is proposed for testing the compatibility of the indicators. Our approach has the flexibility that the disturbances in the measurement equations need not be independent.

Our empirical specification of the structural joint dependence of health and wages is based on the theoretical framework of health capital developed in Grossman [1972a, 1972b]. Two ordered discrete variables are used as health indicators. One is based on health limitations and another is based on the individual's self-evaluation of health. We have found that individuals with different levels of schooling and of difference races have different perceptions of health in their self-evaluation. The estimated results are compatible with theoretical implications of human capital theory. We can conclude that wages and health capital are strongly jointly dependent. Wages have strong positive effect on the demand for good health and good health raises market productivity and hence wages. It was found that schooling has less effect on the production and demand for health capital than wages.

References

- Amemiya, T., "The Estimation of a Simultaneous Equation Generalized Probit Model", Econometrica, 46, (Sept. 1978), 1193-1205.
- Becker, G. S., Human Capital and the Personal Distribution of Income: An Analytical Approach, (Ann Arbor: University of Michigan Press, 1967).
- Domencich, T. A., and D. McFadden, Urban Travel Demand, (Amsterdam: North Holland Publishing Co. 1975).
- Ferguson, T. S., "A Method of Generating Best Asymptotically Normal Estimates with Application to the Estimation of Bacterial Densities", Annals of Mathematical Statistics, 29, (Dec. 1958), 1049-1062.
- Finney, D. J., Probit Analysis, (Cambridge: Cambridge University Press, 1947).
- Goldberger, A. S., Econometric Theory. (New York: John Wiley and Sons, 1964).
- Graybill, F. A., An Introduction to Linear Statistical Models, vol I, (New York: McGraw-Hill, 1961).
- Grossman, M., The Demand for Health: A Theoretical and Empirical Investigation, (New York: Columbia University Press, 1972b).
- Grossman, M., "On the Concept of Health Capital and the Demand for Health", Journal of Political Economy, 80, (March 1972b), 223-255.
- Grossman, M. and L. Benham, "Health, Hours and Wages", in Mark Perlman, ed., The Economics of Health and Medical Care, (London: Macmillan, 1974), 205-233.
- Grossman, M., "The Correlation Between Health and Schooling", in N.E. Terleckyj, ed., Household Production and Consumption (New York, Columbia University Press, 1976), 147-223.

- Heckman, J., "Shadow Prices, Market Wages and Labor Supply", Econometrica, 42, (July 1974), 679-694.
- Heckman, J., "Dummy Endogenous Variable in a Simultaneous Equation System", Econometrica, 46, (July 1978), 931-939.
- Jenrich, R., "Asymptotic Properties of Nonlinear Least Squares Estimates", Annals of Mathematical Statistics, 40, (April 1969), 633-643.
- Jöreskog, K. G. and A. S. Goldberger, "Factor Analysis by Generalized Least Squares", Psychometrika, 37 (September 1972), 243-260.
- Jöreskog, K. G. and A. S. Goldberger, "Estimation of a Model with Multiple Indicators and Multiple Causes of a Single Latent Variable", Journal of the American Statistical Association, 79, (September 1975), 631-639.
- Kenny, L. W., L. F. Lee, G. S. Maddala and R. P. Trost, "Return to College Education: An Investigation of Self-Selection Bias Based on the Project Talent Data", forthcoming in International Economic Review, 1979.
- Lee, L. F., "Identification and Estimation in Binary Choice Models with Limited (Censored) Dependent Variables", Econometrica, 47, (July 1979), 977-996.
- Lee, L. F., "Simultaneous Equations Models with Discrete and Censored Variables", in C. Manski and D. McFadden, eds., Structural Analysis of Discrete Data, (Mass: MIT Press, 1980), Chapter 9.
- Luft, H. S., "The Impact of Poor Health on Earnings", Review of Economics and Statistics 57, (February 1975), 43-57.
- Luft, H. S., Poverty and Health: Economic Causes and Consequences of Health Problems, (Cambridge, Mass.: Ballinger, 1978).
- Maddala, G. S. and L. F. Lee, "Recursive Models with Qualitative Endogenous Variables", Annals of Economic and Social Measurement, 5, (December 1976), 525-545.

- Malinvaud, E., Statistical Methods of Econometrics, (Amsterdam: North-Holland, 1970).
- McKelvey, R. D. and W. Zavoina, "A Statistical Model for the Analysis of Ordinal Level Dependent Variables", Journal of Mathematical Sociology, 4, (December 1975), 103-120.
- Mincer, J., "The Distribution of Labor Incomes: A Survey with Special Reference to the Human Capital Approach", Journal of Economic Literature, VIII, (March 1970), 1-26.
- Muthén, B., "Structural Equation Models with Dichotomous Dependent Variables", Research Report 76-17, Department of Statistics, University of Uppsala (1976a).
- Muthén, B., "Structural Equation Models with Dichotomous Dependent Variables: A Sociological Analysis Problem Formulated by O. D. Duncan", Research Report 76-19, Department of Statistics, University of Uppsala (1976b).
- Nerlove, M. and S. J. Press, Univariate and Multivariate Log-Linear and Logistic Models, (Santa Monica: Rand R1306-NIH/EDA, 1973).
- Neyman, J., "Contribution to the Theory of χ^2 Test", in J. Neyman, ed., Proceedings of the First Berkeley Symposium on Mathematical Statistics and Probability, (Berkeley: University of California Press, 1949), 239-273.
- Taylor, W. F., "Distance Functions and Regular Best Asymptotically Normal Estimates", Annals of Mathematical Statistics 24, (March 1953), 85-92.
- Theil, H., Principles of Econometrics, (New York: John Wiley and Sons, Inc, 1971).
- The National Longitudinal Survey of Men 45-59, 1966 to 1975, (Ohio: Center for Human Resource Research, 1976).
- Zellner, A., "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias", Journal of the American Statistical Association, 57, (June 1962), 348-368.