Interrelated Dynamics of Health and Poverty in Australia

Hielke Buddelmeyer *and* **Lixin Cai**

Melbourne Institute of Applied Economic and Social Research
University of Melbourne
Contact: cail@unimelb.edu.au

This version on 20 June 2009

Keywords: Socio-economic status, poverty, health, recursive models, panel data
JEL codes: I12; I32; C35

Acknowledgement: Both authors acknowledge financial support provided by the Faculty of Economics and Commerce of the University of Melbourne. The paper uses the data in the confidentialised unit record file from the Department of Families, Housing, Community Services and Indigenous Affairs' (FaHCSIA) Household, Income and Labour Dynamics in Australia (HILDA) Survey, which is managed by the Melbourne Institute of Applied Economic and Social Research (MIAESR). The findings and views reported in the paper, however, are those of the authors and should not be attributed to either the FaHCSIA or the MIAESR.

Interrelated Dynamics of Health and Poverty in Australia

Abstract

Using the Households, Income and Labour Dynamics in Australia (HILDA) survey, this study examines the joint dynamics of health and poverty of Australian families. Taking advantage of panel data, the modelling approach used in this study allows a better estimation of the causal relation between health and poverty. The results indicate that the causality runs both ways and the relationship could be confounded by unobserved heterogeneity. In particular, it is found that families headed by a person in ill-health are more likely to be in poverty compared with families headed by a person with good health. On the other hand, a family head whose family is in poverty in the current year is more likely to be in ill-health in the next year compared with a family head whose family is not in poverty. In addition, there is evidence that health and poverty are affected by correlated unobservables, causing health to be endogenous to poverty even in the absence of a reserve effect. Consequently, treating health as exogenous in a poverty equation would produce a biased estimate for its effect.

1. Introduction

Both health and poverty are important measures of personal wellbeing and they are closely related in their evolvement. Understanding how health and poverty are determined and evolve over time has important policy implications. Numerous international studies have documented a close association between socio-economic status (SES), often measured by income, and health (see Adams et al., 2003 and references therein). The association is found to hold for different populations and various measures of health (Goldman, 2001). The association between poverty and health has also long been noticed in Australia. For example, in the mid-1970s, the Poverty Commission identified poor health as a condition that greatly increased the risk of poverty (Commission of Inquiry into Poverty, 1975). The Australian Council of Social Service (ACOSS) described poverty as being both a consequence of poor health and a health hazard of its own (ACOSS, 1993; Mitchell, 1993). In a recent study, Saunders (1998) showed that Australians under and at the margin of the poverty line are more likely to experience financial and emotional stress in their lives than better-off Australians.

Theoretically the causality between SES and health can run either way. On the one hand, low SES (e.g. income poverty) may cause poor health due to malnutrition and/or less access to medical services. Health risk behavior, such as smoking, alcoholism and drug use, is also more likely to be found among people with low income than among those with high income (Stronks et al., 1996). On the other hand, ill-health may lead to low income and thus poverty because ill-health reduces the ability to work. Despite a close association between SES and health has long been observed, the direction of causality remains an open issue that attracts researchers from both social and medical sciences (Smith, 1999, 2004; Fuchs, 2004; Meer et al., 2003; Deaton, 2002; Frijters et al., 2005). From a policy-maker' viewpoint, knowing the correlation between SES and health are not enough because policy design aimed at improving general health or narrowing health inequality requires understanding the direction of causality.

Medical scientists and researchers in the public health area tend to believe that the pathway is from SES to health (Smith, 1999, 2004). For example, there is a growing research interest in the socio-economic determinants of health in the public health literature, where it is emphasized that the determination of health disparities goes beyond medical treatments and health care services, which are traditionally regarded as the most important determinants of health, to socio-economic factors, such as income, employment status, environment and even income distributions (Wilkinson and Marmot, 1998; Marmot and Wilkinson, 1999). On the other hand, economists seem to be more interested in the effect of health on SES, particularly the effect of health on labour supply and wages (or earnings), with the general finding that people with better health have a higher labour force participation rate and earn higher wages (Cai and Kalb, 2006; Cai, 2009a,b; Cai and Cong, 2009; Stern, 1989; Haveman, 1994; Lee, 1982; Grossman and Benham, 1974).[1]

This study has two main objectives regarding enhancing our understanding of the relationship between health and poverty: (a) to disentangle the causal relationship between health and poverty; and (b) to identify whether and to what extent intertemporal persistence exists in health and poverty. For such purposes, we explore the panel nature of the Households, Income and Labour Dynamics in Australia (HILDA) survey and model jointly the dynamics of health and poverty.

The rest of the paper is organised as follows. Section 2 discusses the identification strategy regarding the causal relation between health and poverty and describes the statistical model and estimation methods. Section 3 describes the data source and key variables and provides some descriptive results. Section 4 presents the model estimation results and Section 5 sets out the conclusions.

2. Identification strategy, statistical model and estimation method

2.1. Identification strategy

In this study we model the joint dynamics of health and poverty by exploring the panel feature of the data. The advantage of panel data combined with the modelling approach allows us to better address the causality issue between health and poverty. The reasons are as follows. Although the causal effect between health and poverty may run either way, the effect of poverty on health and the effect of health on poverty are likely to occur with a time difference, rather than simultaneously. Change in health is often slow in nature, implying that a change in income and thus poverty status is unlikely to lead to an immediate change in

[1] However, it should be acknowledged that in his pioneered work on health production theory, Grossman, an economist, noted the causal effect of SES on health (1972). By Grossman’s theory, health is a form of human capital that can be maintained or improved through investment. Because health investment depends on both time and economic resources, health capital is affected by individuals’ SES.

health, even if there is a causal effect from income to health. The effect of income change on health would occur some time later. On the other hand, a deterioration of health would have an immediate impact on productivity and labour supply and thus on income. The timing differences in the two effects suggest that longitudinal data are most suitable for identifying the causal effects between health and poverty. The HILDA data we use for this research are longitudinal data covering six years.

However, longitudinal data itself is not sufficient for identifying the causal effects if unobserved determinants of health and poverty are correlated. The joint dynamic model estimated in this study controls for the correlation between the unobserved determinants of health and poverty and therefore allows us to better identify the causal effects between health and poverty. To our best knowledge, there does not seem to be a study in the literature that has used a joint modelling approach to analysing the interrelationship between health and poverty or between SES and health.

2.2. The statistical model

The above discussion suggests that a two-way causality between health and poverty might exist, and correlated unobserved determinants of health and poverty might lead to a spurious relationship between health and poverty if the correlation is not accounted for. To disentangle the complicated relationship between health and poverty, we set up a two equation system that consists of the determination of health and poverty. The first equation specifies how health is determined,

$$(1)\quad H_{it}^{*} = \alpha_{H1} H_{i,t-1} + \alpha_{H2} P_{i,t-1} + X_{H,it}\beta_H + e_{H,it}, \text{ with } H_{it} = \begin{cases} =1 & \text{(Ill-health)} & \text{if } H_{it}^{*}>0 \\ =0 & \text{(Good health)} & \text{otherwise} \end{cases}.$$

Where H_{it}^{*} refers to latent health of individual *i* in time *t*; H_{it} is observed health status. P_{it} refers to observed poverty status in time *t*; $X_{H,it}$ is a vector of observed variables that may affect health; and $e_{H,it}$ is an error term summarising unobserved determinants of health. $H_{i,t-1}$ on the right-hand side of equation (1) is meant to capture the intertemporal persistence of health status.

A similar equation for the determination of poverty status is specified as,

(2) $P_{it}^{*} = \alpha_{P1} H_{it} + \alpha_{P2} P_{i,t-1} + X_{P,it}\beta_P + e_{P,it}$, with $P_{it} = \begin{cases} =1 & \text{(In poverty)} \quad \text{if } P_{it}^{*} > 0 \\ =0 & \text{(Non-poverty)} \quad \text{otherwise} \end{cases}$.

Similarly, the lagged poverty status $P_{i,t-1}$ in equation (2) is used to estimate the degree of intertemporal persistence of poverty status.

Equations (1) and (2) consist of the equation system governing the joint dynamics of health and poverty. Since current health status affects current poverty status, but current poverty does not affect current health, the system is often called a recursive model. However, the casual effect of poverty on health can be assessed by the estimate on the lagged poverty variable in the health equation. As discussed earlier, such a specification of the model can be justified by the observation that health status often evolves slowly and it takes time for the effect of a change in poverty status on health to realise. On the other hand, it should be current health that matters in terms of impacting on current income and thus poverty status.

For the estimate on the health variable in the poverty equation and the estimate on the poverty variable in the health equation to be interpreted as a causal effect, we also need to account for the potential correlation between the two error terms in the two equations. This is implemented by further taking advantages of panel data to assume that each of the error terms in the two equations has two components: a time-invariant component and a time-variant component:

(3) $e_{m,it} = \mu_{m,i} + \varepsilon_{m,it}$, for $m = P, H$,

with $\varepsilon_m \sim N(0,1)$, $cor(\varepsilon_{H,t}, \varepsilon_{P,t}) = \rho_{\varepsilon}$, and $\text{cov}(\varepsilon_{m,t}, \varepsilon_{m,s}) = 0$ for $s \neq t$.

In models with a latent dependent variable such as ours, the time invariant error component is often assumed to be random (Hsiao 2003), implying that time invariant individual heterogeneity is not correlated with the observed variables included in the model. To relax this assumption, we adopt the Mundlak (1978) approach to allow the unobserved individual effect to be correlated with observed controls through the form,[2]

(4) $\mu_{m,i} = \bar{X}_{m,i}\gamma_m^{'} + \eta_{m,i}$ for $m = P, H$.

[2] Such models are often termed as correlated random effects models.

Where $\bar{X}_{m,i} = \frac{1}{T}\sum_{t=1}^{T} X_{m,it}$, $\eta_m \sim N(0,\sigma^2_{m,\eta})$ and is uncorrelated with observed control variables, and $\text{cov}(\eta_H, \eta_P) = \sigma_\eta$.

In our estimation we allow both η_m and ε_m to be correlated across the two equations. We will also estimate models that assume that the two random error components are uncorrelated across equations to see how the estimates change. This is equivalent to assume that the health variable in the poverty equation is exogenous.

With a joint normal assumption on the time variant error components, conditional on observed and unobserved determinants, the joint probability of observing *H=h* and *P=p* (for *h,p=0,1*) is given by

(5) $\Pr(H_{it} = h, P_{it} = p \mid \eta_H, \eta_P) = \Phi_2\{(2h-1)\Pi_{H,it}, (2p-1)\Pi_{P,it}, (2h-1)(2p-1)\rho_\varepsilon\}$,

where $\Pi_{H,it} = \alpha_{H1}H_{i,t-1} + \alpha_{H2}P_{i,t-1} + X_{H,it}\beta_H + \bar{X}_{H,i}\gamma'_H + \eta_H$, and $\Pi_{P,it} = \alpha_{P1}H_{it} + \alpha_{P2}P_{i,t-1} + X_{P,it}\beta_P + \bar{X}_{P,i}\gamma'_P + \eta_P$. Then the conditional probability of observing a sequence of health and poverty status over the period *1,…,T* is

(6) $L_i(\eta_H, \eta_P) = \prod_{t=1}^{T} \Pr(H_{it} = h, P_{it} = p \mid \eta_H, \eta_P)$.

2.3. The Initial condition problem

The dynamic nature of model implies that current health status and poverty status depend on the initial health status and poverty status, which for most of the families in the sample at hand predate the start of the data collection. Estimates of the system (i.e. the coefficient parameters in equations (1) and (2)) are consistent under the assumption of exogenous initial conditions, i.e. if the first observation of health status and poverty status in the data is independent of all previous health status and poverty status. This is a restrictive assumption and is very likely to be violated, given that health and poverty are affected by time invariant unobserved individual effects. One solution, originally suggested by Heckman (1981), is to approximate the unknown initial conditions of health and poverty with two reduced form static equations that utilize information from the first wave of panel data. The reduced form equations for initial health and poverty status can be expressed as,

(7) $H_{i0}^{*} = X_{H,i0}\beta_H + \bar{X}_{H,i}\gamma_{H,0}^{'} + \theta_H\eta_H + e_{H,i0}$, with $H_{i0} = \begin{cases} =1 & \text{(Ill-health)} \quad \text{if } H_{i0}^{*}>0 \\ =0 & \text{(Good health)} \quad \text{otherwise} \end{cases}$, and

(8) $P_{i0}^{*} = \alpha_{H0}H_{i0} + X_{P,i0}\beta_H + \bar{X}_{P,i}\gamma_{P,0}^{'} + \theta_P\eta_P + e_{P,i0}$, with $P_{i0} = \begin{cases} =1 & \text{(Poverty)} \quad \text{if } P_{i0}^{*}>0 \\ =0 & \text{(Nonpoverty)} \quad \text{otherwise} \end{cases}$.

The time variant error terms in the initial condition equations have the same distribution as in the dynamic equations.

Conditional on observed and unobserved heterogeneity, the probability of observing the initial health status and poverty status is defined similarly to equation (5). The conditional (on observed and unobserved heterogeneity) probability of observing the initial condition and a sequence of health and poverty status is

(9) $L_i^{'}(\eta_H, \eta_P) = \Pr(H_{i0} = h, P_{i0} = p \mid \eta_H, \eta_P)\prod_{t=1}^{T}\Pr(H_{it} = h, P_{it} = p \mid \eta_H, \eta_P)$.

The probability conditional on observed variables, but not on unobserved heterogeneity, is then obtained by integrating out the unobserved heterogeneity,

(10) $L_i^{'} = \iint L_i^{'}(\eta_H, \eta_P)dG(\eta_H, \eta_P)$.

Where $G(.)$ is the joint distribution of η_H and η_P. The integration is approximated using numerical simulation,

(11) $\tilde{L}_i = \frac{1}{R}\sum_{r=1}^{R} L_i^{'}(\eta_H^r, \eta_P^r)$.

Where η_H^r and η_P^r are the r^{th} random draws from the joint distribution of η_H and η_P, which is assumed to be the bivariate normal distribution; R is the total number of random draws used in simulating the likelihood function. For the results reported later we use 50 Halton sequence draws in simulating the likelihood function. It has been shown that Halton sequence draws perform better than simple random draws in terms of approximating the objective function (Train 2003).

The sample log-likelihood function is obtained by summing up the log of equation (11) over the families in the sample.

3. Data, variables, and descriptive analysis

3.1. Data and variables

The empirical analysis is based on the Household, Income and Labour Dynamics in Australia (HILDA) survey. The survey is a national household panel survey focusing on families, income, employment and well-being.[3] The HILDA survey contains detailed information on individuals' current labour market activity including labour force status, earnings and hours worked, and employment and unemployment history. The first wave was conducted between August and December 2001. Then, 7,683 households representing 66 per cent of all in-scope households were interviewed, generating a sample of 15,127 persons 15 years or older and eligible for interview. Of them, 13,969 were successfully interviewed. Subsequent interviews for later waves were conducted about one year apart.

At the time of undertaking this study, there were six waves of the HILDA (2001-2006) surveys available. However, since poverty is defined based on financial year family disposable income and financial year income in a year is asked in the survey of the following year, for the six years covered by the data, we can only define poverty for five years (i.e. year 2001-2005).[4] In other words, after the sixth wave data are used to define poverty status for wave 5, the sixth wave data are excluded from the analysis. Consequently, we have a five-year panel to carry out our analysis.

We use a balanced panel for our analysis. Our working sample includes 1.751 families headed by a person aged between 18 and 64 years. We excluded families headed by full-time students and with missing family disposable income. We also excluded families with missing information on the explanatory variables unless the missing value can be incorporated into our analysis.

Financial year disposable income of a family is derived as the sum of financial year individual disposable income of family members in the family. To define poverty, we use OECD equivalence scale to calculate equivalised family disposable income.[5] A family is defined to be poor if its equivalised family income is less than a half of the median in the

---

[3] Detailed documentation of the survey is in Wooden, Freidin and Watson (2002).
[4] The first wave data can be used to define poverty for 2000, but information on other variables, such as health, in 2000 is not available.
[5] This scale gives a weight of 0.5 for the second and subsequent adults and a weight of 0.3 for each child in the family.

sample. As a result the unit of analysis in this study is families, not individuals. Using this definition 10.6 per cent families in the sample are classified as poor, with the poverty rate varying from 10.3 to 11.1 during the five year period 2001-2005, as shown in the third column in Table 1.

Information relating to individual health was collected in both the personal interviews and self-completion questionnaires. In the personal interviews, individuals were asked whether they had a long-term condition, impairment or disability that restricted everyday activities and had lasted or was likely to last for six months or more. In the self-completion questionnaire, the Short Form 36 (SF-36) health status questions were asked. The SF-36 is a measure of general health and wellbeing, and produces scores for eight dimensions of health (Ware *et al.,* 2000). The first question in the SF-36 is the standard self-reported health status question, asking: "*In general, would you say your health is excellent, very good, good, fair or poor?*". This variable is used to define our measure of general health in this study. As can be seen from the question, health from this question has five levels. For ease of modelling and interpretation, we recode the original five levels of health into a binary variable: good health *vs* ill-health. Good health refers to the original good, very good and excellent health; ill-health refers to the original poor and fair health. Since our units of analysis are families and health is measured at individual levels, we use health of the family reference person ('family head' thereafter) to represent health of the family.[6]

3.2. Descriptive analysis

The second column in Table 1 shows the proportion of families headed by ill-health persons in the sample. In the pooled sample from waves 1 to 5, 17.6 per cent of families are headed by persons with ill-health. This proportion increases from 15.3 in 2001 to 19.4 per cent in 2005, perhaps reflecting the ageing of the sample.

Tables 2 and 3 present descriptive information on intertemporal persistence of health and poverty status. On a year-by-year basis both health status and poverty status show substantial intertemporal persistence. For example, of those in good health in a year, 92.6 per cent are expected to remain in good health in the next year; of those in ill-health in a year, 70.3 per cent are expected to remain in ill-health in the next year. For poverty status, of those not in

[6] For a couple family reference person refers to the oldest male in the family; for other families, reference person refers to the oldest person in the family.

poverty in a year, 95.2 are expected to remain non-poor in the next year, and of those being poor in a year, 60.8 per cent are expected to remain poor in the next year.

Table 1: Poverty and health status by wave

Year/wave[(a)]	Ill-health	In poverty
2001	15.31	10.34
2002	16.96	10.34
2003	18.62	10.22
2004	17.76	11.08
2005	19.36	11.14
All years/waves	17.6	10.62

Note: The number of observations in each wave is 1,751.

Table 2: Year-by-year transitions of health/poverty status (row %)

		health/poverty status at t+1		Number of observations
		Good health	Ill-health	
Health status at t	Good health	92.62	7.38	5,802
	Ill-health	29.70	70.30	1,202
		Non-poverty	Poverty	
Poverty status at t	Non-poverty	95.18	4.82	6,269
	Poverty	39.18	60.82	735

Note: Pooled data from waves 1 to 5.

Table 3 shows the persistence of health and poverty from a different angle by looking at the distribution of the number of years in poverty or in ill-health over the five year period. From the table, 67.5 per cent of families in the sample were never in poverty in the five years; 11.2 per cent were poor for only one year in the five years; 5.4 per cent were poor in two out of the five years. About 6.7 per cent were poor for all the five years, and 4.7 poor for four out of the five years. For health status, 76.8 per cent in the sample were always in good health over the five years; 10.5 per cent were in ill-health for one year only; 4.7 per cent were in ill-health for two of the five years. 3.5 percent were in ill-health for all the five years, and two per cent were in ill-health for four out of the five years.

Table 3: Distribution of the number of years in ill-health/poverty

Number of years in poverty /ill-health	Poverty (%)	Ill-health (%)
0	67.45	76.76
1	11.19	10.45
2	5.37	4.68
3	4.57	2.63
4	4.74	2.00
5	6.68	3.48
No. of obs.	1,751	

As for the relationship between health and poverty, Table 4 indicates a clear positive association between ill-health and poverty. From the table 25.1 per cent of those in ill-health are also in poverty. In contrast, only 7.5 per cent of those in good health are in poverty.

Table 4: Cross-tabulation of health and poverty status (row %)

	Non-poverty	Poverty	No.obs.
Good health	92.47	7.53	7,214
Ill-health	74.89	25.11	1,541

The intertemporal persistence of and the association between health and poverty demonstrated in the above tables are descriptive, since these results have not controlled for observed and unobserved heterogeneity among families. The model estimation results presented in the next section take care of these issues.

Table 5 presents the summary statistics for the variables used in the model and the grouping of the variables also shows the model specification. In Panel A are the two endogenous variables. Note that the lagged values of the two endogenous variables are also included in the right-hand side of the model. The variables in Panel B are included as the control variables in both the health and poverty equations. The variables in Panel C are only included in the health equation, while those in Panel D are only included in the initial condition equations of both health and poverty. The triangular feature of the model requires instrumental variables for health for identification purposes, although the non-linearity of the model also helps with identification. The two variables in Panel C serve as instruments for health. Lack of physical activity is a health risk variable and physical functioning index is a summary measure of physical health, As such, it is reasonable to believe that the two variables have no direct effect on poverty status, rather their effects on poverty operate through general health.

Table 5: Summary statistics and model specification

Variables in dynamic equation	Mean	sd
A. Endogenous variables		
Ill-health	0.1760	0.3809
Poor	0.1062	0.3081
B. Variables included in both equations		
Aged 18-25	0.0226	0.1487
Aged 26-35	0.1436	0.3507
Aged36-45	0.2260	0.4183
Aged 46-55	0.3381	0.4731
Aged 56+	0.2697	0.4438
Degree	0.2322	0.4223
Other post-sch qualification	0.3865	0.4870
Year 12	0.1059	0.3077
Year 11 or below	0.2754	0.4467
Australian born	0.7653	0.4238
Overseas En-speak country	0.1342	0.3409
Overseas Non-En country	0.1005	0.3007
Number of children	0.2336	0.6326
Number of workers	1.4176	0.9110
Couple with dependent children	0.2699	0.4439
Couple without dependent children	0.3612	0.4804
Sole parent	0.0332	0.1793
Lone person	0.3357	0.4723
Urban	0.6102	0.4877
Inner region	0.2505	0.4333
Outer region	0.1143	0.3182
Remote area	0.0250	0.1562
Extroversion	4.0033	1.4425
Agreeableness	4.8943	1.5429
Conscientiousness	4.8357	1.3956
Emotional stability	4.8881	1.6364
Openness to experience	4.0287	1.4574
Missing personality	0.0640	0.2447
C. Variables in health equation only		
Lack physical activity	0.1017	0.3022
Physical functioning index	0.8434	0.2181
D. Variables in initial condition		
Prop in emp	0.8815	0.1895
Father white collar	0.3033	0.4598
Father other white collar	0.2204	0.4147
Father blue collar	0.4443	0.4970
Father occupation unknown	0.0320	0.1760
Number of families (observations)	1751 (8755)	

4. Model estimation results

4.1. Goodness-of-fit of the model

Before presenting the model estimation results, we first compare in Table 6 the model predicted (joint and marginal) distributions of health and poverty with that observed directly from the data. This would provide us with a measure about how the model fits the data. The figures in the table suggest that overall the model fits the data quite well. For example, while 76.2 per cent of the family heads in the sample is observed to be in good health and not in poverty, the model predicts 74.9 per cent belonging to this category. While 4.4 per cent of the family heads in the sample are observed to be in ill-health and in poverty, the model predicts this to be 3.7 per cent. The model slightly under-predicts the probabilities of being in good health and not in poverty and being in ill-health and in poverty, and slightly over-predicts the probabilities of falling in the other two groups. In terms of the marginal distribution, the predicted probabilities are even closer to the observed ones. For example, the model predicted probability of being in good health is 82.0 per cent, while the observed probability is 82.4 per cent. For poverty the model predicted and observed probabilities are 10.8 per cent and 10.6 per cent respectively. The Pearson goodness-of-fit statistic also indicates a good fit of the model to the data.

Table 6: Observed and model predicted distributions of health and poverty

	Non-poverty	Poverty	Column sum
	Observed		
Good health	0.7619	0.0620	0.8240
Ill-health	0.1318	0.0442	0.1760
Row sum	0.8938	0.1062	1.0000
	Model predicted		
Good health	0.7494	0.0708	0.8202
Ill-health	0.1430	0.0369	0.1798
Row sum	0.8924	0.1076	1.0000
The Pearson Goodness-of-fit statistic			0.0036

4.2. Coefficient estimates

Table 7 presents the coefficient estimates of the model, together with the estimates for the covariance matrix of the composite errors. Due to the non-linear nature of the model, the coefficient estimates do not represent marginal effects. However, the sign of the estimate on an explanatory variable indicates the impact direction of the variable on the dependent variables. For example, a positive sign on a variable implies that an increase in the variable, *ceteris paribus*, raises the probability of being in poverty (or in ill-health).

First, let's look at the estimates on the relation between health and poverty. In the poverty equation ill-health is estimated to be positive and significant, indicating that compared with a family headed by a person in good health, families headed by a person in ill-health are more likely to be in poverty. In the health equation lagged poverty is estimated to be positive and significant at the 10 per cent significant level, implying that compared with a family head that was not poor in the previous financial year, a head of family that was poor in the previous year is more likely to be in ill health in this year. These effects may be interpreted as causal effects since the endogeneity of the outcome variables have been taken into account in the model estimation.

Turn to the intertemporal persistence of health and poverty. The lagged health variable is positive and significant in the health equation, suggesting that those who were in ill-health in the previous year are also more likely to be in ill-health in the currently year. Similarly, the estimate on the lagged poverty variable indicates that those families who were poor in the previous year are more likely to be poor in this year. That is, there is evidence on intertemporal persistence for both health and poverty even after controlling for observed and unobserved heterogeneity among families.

As for the other control variables, for the health equation it is found that family heads with more children and more workers in the family are associated with a lower probability of being in ill-health. Compared with family heads in couple-with-children families, family heads in the other types of families are less likely to be in ill-health.The heads of families in outer regions are more likely to be in ill-health compared with those living in a city. The two health indicator variables, *lack physical activity* and *physical functioning index*, are significant and have the expected sign. Four of the six personality variables are also found to have a significant effect on health.

Table 7: Coefficient estimates

	Health equation		Poverty equation	
	Coef.	S.e.	Coef.	S.e.
Ill-health			0.8470***	0.1668
Lagged ill-health	0.4142***	0.1000		
Lagged poor	0.2397*	0.1324	0.6825***	0.1141
Male	-0.0822	0.1439	-0.0333	0.1290
Aged 18-25	-0.9205**	0.4073	-0.5367	0.4931
Aged 26-35	-0.0915	0.1348	-0.3031*	0.1582
Aged 46-55	0.1842	0.1236	0.0727	0.1146
Aged 56+	-0.1591	0.1372	0.2132*	0.1235
Degree	0.0738	0.1415	-0.4553***	0.1387
Other post-sch qualification	-0.0793	0.1118	-0.1793*	0.1031
Year 12	-0.0653	0.1674	-0.0213	0.1499
Overseas En-speak country	0.1679	0.1338	-0.1037	0.1343
Overseas Non-En country	0.1807	0.1629	0.3222**	0.1330
Number of children	-0.3073**	0.1464	0.2120	0.1747
Number of workers	-0.2090**	0.0924	-0.2321***	0.0797
Couple without dependent children	-0.2571*	0.1373	-0.2384*	0.1364
Sole parent	-0.6839**	0.3137	-0.2120	0.2906
Lone person	-0.5267***	0.1781	-0.0810	0.1710
Inner region	0.1014	0.1061	0.2982***	0.0985
Outer region	0.3886***	0.1338	0.4822***	0.1136
Remote area	0.0314	0.3631	0.5513*	0.2955
Lack physical activity	0.2446**	0.1207		
Physical functioning index/100	-2.3666***	0.2367		
Extroversion	-0.1617***	0.0493	0.0419	0.0405
Agreeableness	-0.0252	0.0594	0.0249	0.0532
Conscientiousness	-0.1895***	0.0521	-0.0248	0.0479
Emotional stability	-0.1376***	0.0504	0.0057	0.0422
Openness to experience	0.0478	0.0496	0.0089	0.0401
Missing personality	-2.6043***	0.4531	0.6412*	0.3539
Mean no. children	0.0790	0.1735	0.1321	0.1933
Mean no. workers	-0.2472**	0.1092	-0.5409***	0.1068
Mean lack of physical activity	0.4958**	0.2497		
Mean physical functioning/100	-2.3336***	0.4180		
Constant	5.3065***	0.6396	-1.4539***	0.4342
Variance: time invariant err in health	1.3111***	0.2274		
Variance: time invariant err in poverty	0.7460***	0.1837		
Covariance of time invariant errors	-0.1828**	0.0798		
Correlation of time variant errors	-0.3273***	0.1167		
Log-likelihood	-4104.19			
Number of individuals	1,751			

***, **, * indicate estimates are significant at 1%, 5% and 10% respectively.

For the poverty equation it is found that higher levels of education reduce the probability of being in poverty. Families headed by immigrants from non-English speaking countries are more likely to be in poverty compared with families head by persons born in Australian. An increase in the number of workers reduces the probability of being in poverty. Families living in non-urban areas are more likely to be in poverty compared with those living in urban area. Personality does not seem to matter in terms of affecting family poverty status.

The estimates for the covariance matrix of the error terms indicate that time invariant unobserved heterogeneity accounts for about 57 per cent of the total variance for the health equation and 43 per cent for the poverty equation. Both the time invariant and time variant error components are estimated to be significantly correlated between the two equations, suggesting that endogeneity of (current) health in the poverty equation can not be rejected.

4.3. Mean marginal effects

To assess the magnitude of the effects of the explanatory variables on the outcome variables, we calculate the mean marginal effects (MME) based on the model estimates reported in Table 7. The MMEs are obtained by first calculating the marginal effect of a variable on an outcome variable for each observation in the sample and then taking the mean of the marginal effect over the whole sample. In addition, we further decompose the effect on poverty of a variable that is included in both the health and poverty equations into the direct effect and the indirect effect (i.e. the effect through health). The formulas used for computing the MMEs can be found in the Appendix. The MME estimates are reported in Table 8, together with the standard errors calculated using the Delta method (Greene, 2000).

Focusing on the relationship between health and poverty first, the MME estimates show that compared with families headed by a person with good health, families headed by a person with ill-health have a probability of being poor that is two percentage points higher. If a family was poor in the previous financial year, the probability of the family head being in ill-health in the current year is about three percentage points higher compared with a family that was not poor in the previous year.

Table 8: Mean marginal effect estimates

	Health equation		Poverty equation			
	MME	S.e.[(a)]	MME: total	S.e. [(a)]	MME: direct	MME: indirect
Ill-health			0.0193*	0.0116		
Lagged ill-health	0.0509***	0.0149				
Lagged poor	0.0281*	0.0163	0.0906***	0.0204	0.0879	0.0027
Male	-0.0093	0.0164	-0.0044	0.0132	-0.0034	-0.0010
Aged 18-25	-0.0808***	0.0287	-0.0503*	0.0300	-0.0414	-0.0089
Aged 26-35	-0.0099	0.0146	-0.0262**	0.0118	-0.0252	-0.0010
Aged 46-55	0.0214	0.0142	0.0075	0.0084	0.0053	0.0022
Aged 56+	-0.0170	0.0146	0.0124	0.0086	0.0142	-0.0018
Degree	0.0084	0.0162	-0.0338***	0.0092	-0.0347	0.0009
Other post-sch qualification	-0.0087	0.0124	-0.0115*	0.0059	-0.0106	-0.0009
Year 12	-0.0072	0.0183	-0.0026	0.0133	-0.0019	-0.0007
Overseas En-speak country	0.0191	0.0156	-0.0072	0.0114	-0.0091	0.0020
Overseas Non-En country	0.0206	0.0192	0.0346**	0.0145	0.0324	0.0021
Number of children	-0.0519**	0.0247	-0.0117	0.0119	0.0077	-0.0194
Number of workers	-0.0353**	0.0156	-0.0217***	0.0065	-0.0085	-0.0132
Couple without dependent children	-0.0308*	0.0168	-0.0191**	0.0084	-0.0160	-0.0032
Sole parent	-0.0739**	0.0299	-0.0267	0.0237	-0.0189	-0.0078
Lone person	-0.0591***	0.0201	-0.0102	0.0081	-0.0040	-0.0062
Inner region	0.0112	0.0118	0.0246***	0.0082	0.0234	0.0011
Outer region	0.0460***	0.0167	0.0548***	0.0132	0.0501	0.0046
Remote area	0.0034	0.0396	0.0663	0.0404	0.0659	0.0004
Lack physical activity	0.0288*	0.0149				
Physical functioning index/100	-0.3995***	0.0410				
Extroversion	-0.0273***	0.0084	-0.0087**	0.0034	0.0015	-0.0102
Agreeableness	-0.0043	0.0100	-0.0007	0.0042	0.0009	-0.0016
Conscientiousness	-0.0320***	0.0089	-0.0129***	0.0038	-0.0009	-0.0120
Emotional stability	-0.0232***	0.0085	-0.0085**	0.0036	0.0002	-0.0087
Openness to experience	0.0081	0.0084	0.0033	0.0034	0.0003	0.0030
Missing personality	-0.1673***	0.0158	0.0525	0.0468	0.0697	-0.0172
Mean no. children	0.0133	0.0293	0.0098	0.0134	0.0048	0.0050
Mean no. workers	-0.0417**	0.0184	-0.0354***	0.0083	-0.0197	-0.0156
Mean lack of physical activity	0.0837**	0.0422				
Mean physical functioning/100	-0.3939***	0.0733				

***, **, * indicate estimates are significant at 1%, 5% and 10% respectively.

(a): standard errors are calculated using the Delta method (Greene, 2000).

As for the intertemporal persistence of health, the probability of being in ill-health in the current year is about five percentage points higher for a family head who was in ill-health in the previous year than for a family head who was in good health in the previous year. If a family was poor in the previous financial year, the probability that the family is poor in the current year is about nine percentage points higher compared with a family who was not poor in the previous year. The degree of intertemporal persistence in poverty appears to be larger that in health. The effect of lagged poverty on current poverty is largely through the direct effect (8.8 per cent), as opposed to the indirect effect (0.3 per cent).

4.4. Alternative model results

The model results reported above treated current health as endogenous and the two equations are estimated jointly. Since the results on the correlation between the two equations suggest that current health could be endogenous due to that both time invariant and time variant unobservables are correlated between the two equations, it would be interesting to assess the bias of a model that treats current health as exogenous. This model is equivalent to estimating the dynamic model of health and poverty separately, while unobserved heterogeneity and initial condition are still accounted for in the respective model. Panel (a) in Table 9 reports the MME estimates on the key variables from such a model. Panel (b) in the table presents model results when the lagged dependent variables are treated as exogenous. That is, the initial condition problems are ignored in such a model, while the time variant and time invariant errors are still allowed to be correlated between the two equations. The simplest model that one can estimate is to ignore unobserved individual effects, the initial condition issue and the correlation across equations altogether. This is equivalent to estimating a separate simple dynamic probit model for health and poverty by pooling the panel data. The results from such a model are reported in panel (c) in Table 9.

First, compared with the MME estimate in Table 8, the MME estimate in panel (a) on (current) ill-health in the poverty equation is larger when health is treated as exogenous; the difference is statistically significant at the 5 percent level. On the other hand, the MME estimates on the lagged poor and lagged health variables are smaller when health is treated as exogenous compared with the estimates where health is treated as endogenous, but the differences of the estimates on these variables between the two models do not appear to be statistically significant. Second, in panel (b) when initial condition issue is ignored, the MME estimate on health in the poverty equation becomes smaller and insignificant; the lagged poverty variable has a marginal effect on health that is similar to the one shown in Table 8; but the estimates on intertemporal persistence of health and poverty become much larger than those reported in Table 8. Third, like the results in panel (b), the simple dynamic models also produce much larger estimates for intertemporal persistence of the dependent variables; the estimate on the health variable in the poverty equation becomes much larger as well, while the lagged poverty variable becomes smaller and insignificant in the health equation.

Table 9: MME estimates of alternative models

	Health equation		Poverty equation	
	MME	S.e.	MME	S.e.
(a). Uncorrelated dynamics with RE & IC				
Ill-health			0.0442***	0.0105
Lagged ill-health	0.0356***	0.0135		
Lagged poor	0.0226*	0.0133	0.0675***	0.0188
(b). Correlated dynamics with RE, without IC				
Ill-health			0.0121	0.0107
Lagged ill-health	0.1207***	0.0185		
Lagged poor	0.0293*	0.0175	0.2641***	0.0286
(c). Uncorrelated dynamics without RE & IC				
Ill-health			0.0348***	0.0083
Lagged ill-health	0.3648***	0.0175		
Lagged poor	0.0119	0.0112	0.325***	0.0197

***, **, * indicate estimates are significant at 1%, 5% and 10% respectively.

5. Conclusion

Using the Households, Income and Labour Dynamics in Australia (HILDA) survey, this study examines the joint dynamics of health and poverty of Australian families. The joint modelling approach taken in this study, combined with advantage of panel data, allows us to better estimate the causal effects between health and poverty. Therefore, this study not only helps our understanding about how health and poverty are determined and evolve over time among Australian families, it also contributes to the literature on the relationship between SES and health.

The estimation results indicate that the causality indeed runs both ways and the relationship between health and poverty could be confounded by unobserved heterogeneity. In particular, it is found that families headed by a person in ill-health are more likely to be in poverty compared with families headed by a person with good health. On the other hand, a family head whose family is in poverty in this year is more likely to be in ill-health in the next year compared with a family head whose family is not in poverty. In addition, it is found that the unobserved determinants of health are significantly and negatively correlated with the unobserved determinants of poverty, suggesting that health should not be treated as exogenous in the poverty equation. It is also found that intertemporal persistence exists in both health and poverty even after controlling for observed and unobserved heterogeneity, and the degree of persistence is larger in poverty than in health.

Appendix

Table a1: Estimates for initial condition equations

	Health equation		Poverty equation	
	Coef.	S.e.	Coef.	S.e.
Lagged poor	0.1317	0.2255		
Ill-health			0.9156***	0.2438
Male	-0.0870	0.2407	0.2794	0.2434
Aged 18-25	-0.3331	0.4120	-0.3371	0.5663
Aged 26-35	-0.1146	0.2708	-0.1593	0.2370
Aged 46-55	-0.0854	0.1974	-0.0424	0.1920
Aged 56+	-0.0455	0.2274	0.0727	0.2242
Degree	-0.0864	0.2414	-0.6865***	0.2475
Other post-sch qualification	0.0214	0.1802	-0.4439**	0.1932
Year 12	0.0467	0.2947	-0.3867	0.2587
Overseas En-speak country	0.2876	0.2194	0.0152	0.2245
Overseas Non-En country	-0.3201	0.2590	0.2195	0.2417
Number of children	-0.0955	0.3588	0.2883	0.3239
Number of workers	-0.1943	0.1615	-0.4540***	0.1440
Couple without dependent children	-0.2852	0.2379	-0.3047	0.2534
Sole parent	-0.9480**	0.4379	0.4207	0.3833
Lone person	-0.5109*	0.2875	0.1188	0.2606
Inner region	0.2620	0.1807	0.3650**	0.1773
Outer region	0.4059*	0.2461	0.5482**	0.2226
Remote area	0.4091	0.4882	0.8265**	0.3621
Prop life-time emp	-0.0057	0.0036	-0.0045	0.0042
Father other white collar	0.2938	0.2052	-0.3162	0.2144
Father blue collar	0.0557	0.1708	-0.1863	0.1638
Father occupation unknown	0.4458	0.3939	-0.2763	0.4609
Lack physical activity	0.5368**	0.2721		
Physical functioning index	-2.2371***	0.5359		
Extroversion	-0.1505*	0.0771	0.0266	0.0738
Agreeableness	-0.0343	0.0880	0.0147	0.0846
Conscientiousness	-0.3465***	0.0835	0.0046	0.0774
Emotional stability	-0.0942	0.0890	-0.0255	0.0785
Openness to experience	0.1868**	0.0783	0.0480	0.0768
Missing personality	-1.9656***	0.7268	0.7756	0.6882
Mean no. children	-0.2925	0.3499	0.2559	0.3487
Mean no. workers	-0.2666	0.1854	-0.2849	0.2016
Mean lack of physical activity	0.7012*	0.4238		
Mean physical functioning	-2.5187***	0.6342		
Constant	5.6423***	0.9751	-1.0913	0.8295
Coe. on random effects	0.9484***	0.1422	1.1095***	0.2476

References

Adams, P., Hurd, M.D., McFadden, D., Merrill, A. and Ribeiro, T. (2003) "Healthy, Weathy, and Wise? Tests for Direct Causal Paths between Health and Socioeconomic Status", *Journal of Econometrics*, Vol. 112, p3-56.

ACOSS (1993) "Poverty Is a Health Hazard", *ACOSS Research Paper No. 7*, Sydney: Australian Council of Social Service.

Cai, L. and Kalb, G. (2006) "Health Status and Labour Force Participation: Evidence from Australia", *Health Economics*, Vol. 15, p241-261.

Cai, L. (2009a) "Effects of Health on Wages of Australian Men", *Economic Record* (in press).

--- (2009b) "The Relationship between Health and Labour Force Participation: Evidence from a Panel Data Simultaneous Equation Model", *Labour Economics* (in press).

--- and Cong, C. (2009) "Effects of Health and Chronic Diseases on Labour Force Participation of Older Working-Age Australians", *Australian Economic Papers* (in press)

Commission of Inquiry into Poverty (1075) *First Main Report: Poverty in Australia*, Canberra: Australian Government Publishing Service.

Currie, J. and Madrian, B. C. (1999). "Health, Health Insurance and the Labor Market", in *Handbook of Labor Economics*, ed. by Ashenfelter, O. and Card, D., Elsevier Science B.V. 3: p3310-3415.

Deaton, A. (2002) "Policy Implications of the Gradient of Health and Wealth", *Health Affair*, Vol. 21 (2), p13-30.

Frijters, P., Shields, M.A., and J.P. Haisken-DeNew (2005) "The Effect of Income on Health: Evidence from a Large Scale Natural Experiment", *Journal of Health Economics* Vol. 24, p 997-1017.

Fuchs, V.R. (2004) "Reflections on the Socio-economic Correlates of Health", *Journal of Health Economics*, Vol. 23, p653-661.

Goldman, N. (2001) “Social Inequalities in Health: Disentangling the Underlying Mechanisms”, in Weinstein, M., and Hermalin, A. (Eds.), *Strengthening the Dialogue between Epidemiology and Demography*, Annals of the New York Academy of Sciences.

Greene, W.H. (2000) *Econometric Analysis*, 4th edition, Prentice_Hall, Inc.

Hsiao, C. (2003) *Analysis of Panel Data*, 2nd edition, Cambridge.

Meer, J., Miller, D. and Rosen, H. (2003) “Exploring the Health-Wealth Nexus”, *Journal of Health Economics,* Vol. 22, p713-730.

Mitchell, M. (1993) “Poverty Is a Health Hazard: Address to the Behavioral Medicine Conference”, *ACOSS Paper No. 62*. Sydney: Australian Council of Social Service.

Saunders, P. (1998) “Poverty and Health: Exploring the Links between Financial Stress and Emotional Stress in Australia”, *Australian and New Zealand Journal of Public Health*, Vol. 22 (1), p11-16.

Smith, J. (1999) “Healthy Bodies and Thick Wallets: the Dual Relation between Health and Economic Status”, *Journal of Economic Perspectives*, Vol. 13, p146-166.

---- (2004) “Unraveling the SES Health Connection”, *The Institute for Fiscal Studies Working Paper 04/02*.

Stronks, K., Van de Mheen, H., Looman, C. W. N. and Mackenbach, J. P. (1996) “Behavioral and structural factors in the explanation of socio-economic inequalities in health: An empirical analysis”, *Sociology of Health and Illness*, Vol. 18, p653-674.

Train, K.E. (2003) *Discrete Choice Methods with Simulation*, Cambridge.