DISCUSSION PAPER SERIES

IZA DP No. 4602

Interrelated Dynamics of Health and Poverty in Australia

Hielke Buddelmeyer
Lixin Cai

November 2009

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Interrelated Dynamics of Health and Poverty in Australia

Hielke Buddelmeyer
MIAESR, University of Melbourne and IZA

Lixin Cai
MIAESR, University of Melbourne

Discussion Paper No. 4602
November 2009

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 4602
November 2009

ABSTRACT

Interrelated Dynamics of Health and Poverty in Australia*

Using the Household, Income and Labour Dynamics in Australia (HILDA) Survey, this study examines the joint dynamics of health and poverty in Australian families. Taking advantage of panel data, the modelling approach used in this study allows a better estimation of the causal relationship between health and poverty. The results indicate that the causality between health and poverty runs both ways and the relationship is 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 reverse effect from poverty on health. Consequently, treating health as exogenous in a poverty equation would produce biased estimates.

JEL Classification: I12, I32, C35

Keywords: socio-economic status, poverty, health, recursive models, panel data

Corresponding author:

Hielke Buddelmeyer
Melbourne Institute of Applied Economic & Social Research
University of Melbourne
Alan Gilbert Building, 7th floor
161 Barry Street
Carlton 3053 VIC
Australia
E-mail: hielkeb@unimelb.edu.au

* The authors acknowledge financial support from 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.

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) shows 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 long observed close association between SES and health, 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's viewpoint, knowing the correlation between SES and health is 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, 2009a,b; Cai and Cong, 2009; Cai and Kalb, 2006; 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 HILDA Survey and model the dynamics of health and poverty jointly.

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

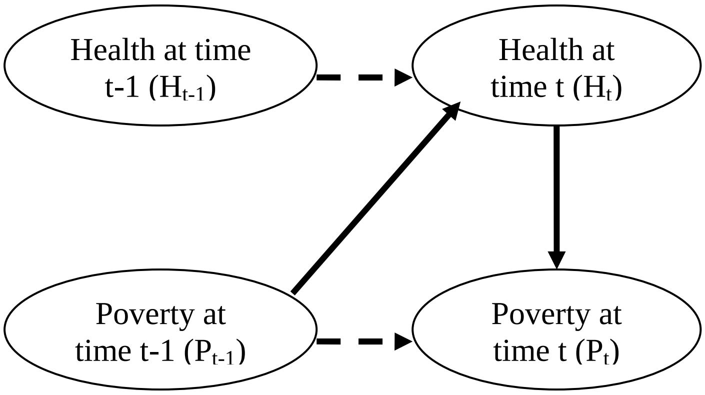
2.1. Identification strategy

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 employed here 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 in theory, 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 health, even if there is a causal effect from income on health. In other

[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.

words, it takes time for the effect of income changes on health to manifest itself. On the other hand, a deterioration of health would have an immediate impact on productivity and/or labour supply and thus on income. Abstracting from other observed and unobserved factors that also affect health and poverty status, the potential causal relationship between health and poverty can be illustrated by the solid arrow lines in Figure 1. That is, while it is current health status that affects current poverty status, it is lagged poverty status that affects current health status.

Figure 1: The relationship between health and poverty



The timing differences in the two effects suggest that longitudinal data are most suitable to identify the causal effects between health and poverty. The HILDA data used 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. This, together with the panel data, allows us to better identify the causal effects between health and poverty.

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 causal relationship between health and poverty if the correlation is not accounted for. To disentangle the complicated relationship between health and poverty, we estimate a two equation system that consists of the determination of health and poverty. First, equation (1) specifies how health is determined,

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

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 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, as illustrated by the dashed line in Figure 1.

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

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

Similarly, the lagged poverty status $P_{i,t-1}$ in equation (2) is used to capture 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 status, such a 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 manifest itself. 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, representing unobserved individual fixed effects; and a time-variant component, representing unobserved transitory shocks to health and income, respectively:

(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 equations (1) and (2), the time invariant error component is often assumed to be random, implying that unobserved individual heterogeneity is not correlated with the observed variables included in the model. This is a maintained assumption in this study.[2]

In our estimation we allow both μ_m and ε_m to be correlated across the two equations. With a joint normal assumption on the time invariant error components (i.e. $\mu_m \sim N(0, \sigma^2_{m,\mu})$ and $\text{cov}(\mu_H, \mu_P) = \delta_\mu$), conditional on observed and unobserved determinants, the joint probability of observing *H=h* and *P=p* (for *h,p=0,1*) is given by

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

where $\Pi_{H,it} = \alpha_{H1} H_{i,t-1} + \alpha_{H2} P_{i,t-1} + X_{H,it}\beta_H + \mu_H$, and $\Pi_{P,it} = \alpha_{P1} H_{it} + \alpha_{P2} P_{i,t-1} + X_{P,it}\beta_P + \mu_P$. Then, the conditional probability of observing a sequence of health and poverty status over the period *1,…,T* is

(5) $L_i(\mu_H, \mu_P) = \prod_{t=1}^{T} \Pr(H_{it} = h, P_{it} = p \mid \mu_H, \mu_P)$.

2.3. The Initial condition problem

The dynamic nature of the 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. The parameter estimates of the system (i.e. the coefficients 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

[2] Mundlak (1978) proposes to relax this assumption by specifying that the unobserved individual effects are a linear function of the means of the observed time varying variables. Although most of the variables included in our models are time varying, the variation over time of the variables for each individual are not sufficiently large for the Mundlak method to work.

independent of all previous health status and poverty status. This is a restrictive assumption and also 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 the data. The reduced form equations for initial health and poverty status can be expressed as,

(6) $H^*_{i0} = X_{H,i0}\beta_{H0} + \theta_H \mu_H + e_{H,i0}$, 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

(7) $P^*_{i0} = \alpha_{H0} H_{i0} + X_{P,i0}\beta_H + e_{P,i0}$, 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 (4). The conditional (on observed and unobserved heterogeneity) probability of observing the initial condition and a sequence of health and poverty status is

(8) $L^{'}_i(\mu_H, \mu_P) = \Pr(H_{i0} = h, P_{i0} = p \mid \mu_H, \mu_P) \prod_{t=1}^{T} \Pr(H_{it} = h, P_{it} = p \mid \mu_H, \mu_P)$.

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

(9) $L^{'}_i = \iint L^{'}_i(\mu_H, \mu_P) dG(\mu_H, \mu_P)$.

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

(10) $\tilde{L}_i = \frac{1}{R} \sum_{r=1}^{R} L^{'}_i(\mu^r_H, \mu^r_P)$.

Where μ^r_H and μ^r_P are the r^{th} random draws from the joint distribution of μ_H and μ_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 (10) over all the families in the sample.

3. Data, variables, and descriptive analysis

3.1. Data and variables

The empirical analysis is based on the HILDA Survey, 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 histories. 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 who were 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 following year's survey, 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 the analysis.

We use a balanced panel for our analysis. Our working sample includes 1,769 families headed by persons 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 included in the model, unless the missing value can be incorporated into the analysis.

---

[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.

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 the OECD-modified equivalence scale to calculate equivalised family disposable income.[5] A family is defined to be poor if its equivalised income is less than a half of the median in the sample. Using this definition 10.6 per cent families in the sample are classified as poor, with the poverty rate varying from 10.2 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 that 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.

While poverty is defined at the family level, health is measured at the individual level. But the joint model described in Section 2 has to be based on the same unit of analysis, either individuals or families. We use families as our unit of analysis, with health being represented by the reference person of a family (i.e. family head). The HILDA data do not define the reference person explicitly. In this study a reference person is defined as follows: for a couple family the reference person refers to the oldest male in the family; for other families, the reference person refers to the oldest person in the family.

3.2. Descriptive analysis

---

[5] The OECD-modified equivalence scale gives a weight of 1 to the first adult family member, 0.5 to the second and subsequent adults and a weight of 0.3 to each child in the family.

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. 10.6 per cent of the families in the pooled sample were in poverty; and on a yearly basis the poverty rate varies from 10.2 to 11.1 per cent.

Table 1: Poverty and health status by wave

Year/wave[a]	Ill-health	In poverty
2001	15.32	10.35
2002	16.98	10.29
2003	18.64	10.23
2004	17.78	11.09
2005	19.38	11.09
All years/waves	17.62	10.61

Note: The number of observations in each wave is 1,749.

Tables 2 and 3 present descriptive information on intertemporal persistence of health and poverty status, respectively. On a year-by-year basis both health status and poverty status exhibit 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 poverty in a year, 95.2 are expected to remain non-poor in the next year, and of those being poor in a year, 60.6 per cent are expected to remain poor in the next year.

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	92.61	7.39	5,794
	Ill-health	29.70	70.30	1,202
		Non-poverty	Poverty	
Poverty status at t	Non-poverty	95.18	4.82	6,262
	Poverty	39.37	60.63	734

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, 67.4 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 4.8 poor for four out of the five years. For health status, 76.7 per cent of the families in the sample were always in good health over the five years; 10.5 per cent were in ill-health for one year only; 4.8 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	67.41	76.73
1	11.21	10.46
2	5.37	4.80
3	4.57	2.52
4	4.75	2.00
5	6.69	3.49
No. of obs.	1,749	

As for the relationship between health and poverty, Table 4 indicates a clear positive association between ill-health and poverty. About 25 per cent of those in ill-health are 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	92.48	7.52	7,204
Ill-health	74.95	25.05	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 takes 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 variable 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 variable physical functioning index in Panel (c) serves as the instrument for health. This index is constructed from individuals' answers to questions about specific physical functioning limitations, such as climbing one flight of stairs, lifting or carrying groceries, or bending, kneeling or stooping. The index value ranges from 0 to 100, with 100 indicating that there is no physical functioning limitation (see Ware *et al.*, (2000) for the construction and interpretation of the index.) By using this variable as an instrument for health, it is assumed that this variable has no direct impact on income other than indirectly through the underlying health status. This is not an unreasonable assumption, given that the physical functioning index variable is constructed from specific physical health indicators.

In wave 5 respondents were asked questions about their personality character traits based on the 'Big 5' (Saucier 1994). In most studies these personality traits are unobservable and assumed to be making up part of the unobserved heterogeneity terms, but making them explicit is important given that health status is self reported. We use the derived variables for the 5 traits, which are scaled from 1 (least) to 7 (most).

Table 5: Summary statistics and model specification

	All	Good health	Ill health	Non-poverty	Poverty
(a). Endogenous variables					
Ill-health	0.1762	0.0000	1.0000	0.1478	0.4159
Poor	0.1061	0.0752	0.2505	0.0000	1.0000
(b). Variables included in both equations					
Aged 18-25	0.0226	0.0267	0.0039	0.0239	0.0119
Aged 26-35	0.1432	0.1589	0.0694	0.1517	0.0711
Aged36-45	0.2263	0.2382	0.1707	0.2308	0.1886
Aged 46-55	0.3380	0.3295	0.3777	0.3387	0.3319
Aged 56+	0.2699	0.2467	0.3783	0.2548	0.3966
Degree	0.2313	0.2485	0.1512	0.2488	0.0841
Other post-sch qualification	0.3870	0.3902	0.3718	0.3913	0.3502
Year 12	0.1060	0.1110	0.0824	0.1068	0.0991
Year 11 or below	0.2757	0.2503	0.3945	0.2530	0.4666
Australian born	0.7650	0.7705	0.7391	0.7650	0.7651
Overseas En-speak country	0.1344	0.1324	0.1434	0.1398	0.0884
Overseas Non-En country	0.1006	0.0970	0.1175	0.0952	0.1466
Couple with dependent children	0.2701	0.2862	0.1947	0.2805	0.1821
Couple without dependent children	0.3605	0.3604	0.3615	0.3778	0.2155
Sole parent	0.0333	0.0310	0.0441	0.0334	0.0323
Lone person	0.3361	0.3225	0.3997	0.3083	0.5700
Urban	0.6097	0.6258	0.5347	0.6295	0.4429
Inner region	0.2508	0.2418	0.2927	0.2406	0.3362
Outer region	0.1145	0.1056	0.1557	0.1053	0.1918
Remote area	0.0250	0.0268	0.0169	0.0246	0.0291
Extroversion	4.0026	4.0373	3.8407	4.0270	3.7979
Agreeableness	4.8939	4.8867	4.9276	4.9101	4.7572
Conscientiousness	4.8385	4.8892	4.6012	4.8719	4.5565
Emotional stability	4.8882	4.9238	4.7215	4.9187	4.6307
Openness to experience	4.0257	4.0213	4.0465	4.0623	3.7174
Personality info missing	0.0640	0.0661	0.0545	0.0585	0.1110
(c). Variables in health equation only					
Physical functioning index/100	0.8434	0.8972	0.5917	0.8623	0.6845
(d). Variables in initial condition					
Prop of life time in employment	0.8813	0.8977	0.7911	0.8975	0.7417
Standard deviation	*0.1896*	*0.1738*	*0.2413*	*0.1677*	*0.2871*
Father white collar	0.3036	0.3113	0.2612	0.3010	0.3260
Father other white collar	0.2207	0.2248	0.1978	0.2264	0.1713
Father blue collar	0.4437	0.4328	0.5037	0.4407	0.4696
Father occupation unknown	0.0320	0.0311	0.0373	0.0319	0.0331
Number of families (observations)	1749(8745)	(7204)	(1541)	(7817)	(928)

4. Model estimation results

4.1. Goodness-of-fit of the model

Before presenting the model estimation results, we first compare the model's predicted (joint and marginal) distributions of health and poverty with that observed directly from the data (Table 6). This provides us with a measure of how the estimated model fits the data. The figures in the table suggest that overall the model fits the data very 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 74.0 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 2.9 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 82.6 per cent, while the observed probability is 82.4 per cent. For poverty the model predicted and observed probabilities are 11.4 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.[6]

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

	Non-poverty	Poverty	Column sum
	Observed		
Good health	0.7618	0.0619	0.8238
Ill-health	0.1321	0.0441	0.1762
Row sum	0.8939	0.1061	1.0000
	Model predicted		
Good health	0.7404	0.0852	0.8256
Ill-health	0.1453	0.0291	0.1744
Row sum	0.8857	0.1143	1.0000
The Pearson Goodness-of-fit statistic			0.0160

[6] The Pearson goodness-of-fit statistic is computed as $GOF = \sum_{s=1}^{4} [(n_s - \hat{n}_s)/\hat{n}_s]$, where n_s and $\hat{n}_s$ are respectively the observed and predicted joint probability in the s^{th} cell.

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, e_H and e_P. 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 variable. For example, a positive sign on a variable in the health (poverty) equation implies that an increase in the variable, *ceteris paribus*, raises the probability of being in ill-health (poverty).

First, consider 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, implying that compared with a head of family that is not poor now, a head of family that is poor now is more likely to be in ill health in the following 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.

Turning to the intertemporal persistence of health and poverty, the lagged health variable is positive and significant in the health equation, suggesting that those who are in ill-health now are also more likely to be in ill-health in the following year. Similarly, the estimate on the lagged poverty variable indicates that those families who are poor now are more likely to be poor in the following 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, it is found that in general older household heads are more likely to be in ill-health than younger ones. These estimates indicate that compared with those aged 36-45, those aged 18-25 are less likely to be in ill-health, but those aged 46-55 are more likely to be in ill-health. Compared with Australian born people, immigrants from non-English speaking countries are more likely to be in ill-health. Those living in outer regional areas are more likely to be in ill-health compared with those who live in urban areas. The variable *physical functioning index* is significant and has the expected sign. Three of the five personality variables are also found to have a significant effect on health. None of the personality traits matter for income poverty.

Table 7: Coefficient estimates

	Health equation		Poverty equation	
	Coef.	S.e.	Coef.	S.e.
Ill-health			1.3711***	0.1546
Lagged ill-health	0.7746***	0.0825		
Lagged poor	0.5632***	0.1167	0.7689***	0.1126
Male	-0.0564	0.1212	-0.0706	0.1179
Aged 18-25	-0.8708**	0.3667	-0.5668	0.4490
Aged 26-35	-0.1619	0.1169	-0.2116	0.1379
Aged 46-55	0.2263**	0.0996	-0.0976	0.1027
Aged 56+	0.0597	0.1080	0.1959*	0.1102
Degree	-0.1369	0.1159	-0.5335***	0.1252
Other post-sch qualification	-0.1166	0.0901	-0.1942**	0.0983
Year 12	-0.1902	0.1486	-0.0892	0.1473
Overseas En-speak country	0.1010	0.1115	-0.1665	0.1265
Overseas Non-En country	0.3002**	0.1290	0.4106***	0.1333
Couple with dependent children	-0.1090	0.0943	0.1781*	0.1021
Sole parent	-0.3107	0.2489	-0.0532	0.2407
Lone person	0.0437	0.1014	0.5386***	0.1003
Inner region	0.0222	0.0885	0.4190***	0.0902
Outer region	0.2905***	0.1107	0.5016***	0.1077
Remote area	-0.0847	0.2760	0.2968	0.2905
Physical functioning index/100	-3.1783***	0.1755		
Extroversion	-0.1210***	0.0370	0.0392	0.0382
Agreeableness	0.0168	0.0467	0.0503	0.0498
Conscientiousness	-0.1483***	0.0423	-0.0310	0.0455
Emotional stability	-0.1343***	0.0420	0.0234	0.0423
Openness to experience	0.0110	0.0396	0.0037	0.0392
Personality info missing	-2.1862***	0.3558	0.8453**	0.3369
Constant	2.9694***	0.4199	-2.8128***	0.3813
Variance (μ_H)	0.7451***	0.1416		
Variance (μ_P)	0.7248***	0.1630		
Covariance (μ_H , μ_P)	-0.3218***	0.0662		
Correlation (ε_H , ε_P)	-0.5050***	0.1069		
Log-likelihood	-4324.79			
Number of individuals	1,749			

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

For the control variables in 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 a family headed by persons born in Australian. Lone person families are more likely to be in poverty compared with couple families without dependent children. Families living in non-urban areas are more likely to be in poverty compared with those living in urban areas.

The estimates for the covariance matrix of the error terms indicate that time invariant unobserved heterogeneity accounts for about 43 per cent of the total variance for the health equation and 42 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, and they are jointly significant as well, 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 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 MME estimates are reported in Table 8, together with the standard errors calculated using the Delta method.

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 about two percentage points higher. If a family is poor now, the probability of the family head being in ill-health in the following year is about 8 percentage points higher compared with a family that is not poor now.

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			0.0186*	0.0108		
Lagged ill-health	0.1257***	0.0193				
Lagged poor	0.0859***	0.0197	0.1335***	0.0254	0.1203	0.0132
Male	-0.0075	0.0163	-0.0092	0.0130	-0.0079	-0.0013
Aged 18-25	-0.0848***	0.0274	-0.0566**	0.0245	-0.0421	-0.0145
Aged 26-35	-0.0196	0.0140	-0.0220**	0.0111	-0.0187	-0.0033
Aged 46-55	0.0308**	0.0133	0.0021	0.0072	-0.0072	0.0051
Aged 56+	0.0077	0.0140	0.0159*	0.0086	0.0146	0.0013
Degree	-0.0184	0.0155	-0.0437***	0.0079	-0.0406	-0.0031
Other post-sch qualification	-0.0158	0.0123	-0.0148**	0.0058	-0.0122	-0.0026
Year 12	-0.0252	0.0191	-0.0125	0.0130	-0.0083	-0.0042
Overseas En-speak country	0.0133	0.0150	-0.0032	0.0110	-0.0154	0.0022
Overseas Non-En country	0.0420**	0.0192	0.0555***	0.0167	0.0484	0.0070
Couple without dependent children	-0.0141	0.0122	0.014	0.0097	0.0164	-0.0024
Sole parent	-0.038	0.0280	-0.0117	0.0229	-0.0053	-0.0064
Lone person	0.0059	0.0138	0.0425***	0.0089	0.0415	0.0010
Inner region	0.0029	0.0115	0.0380***	0.0087	0.0375	0.0005
Outer region	0.0407**	0.0163	0.0659***	0.0146	0.0592	0.0068
Remote area	-0.0107	0.0339	0.0335	0.0369	0.0353	-0.0018
Physical functioning index/100	-0.5610***	0.0357				
Extroversion	-0.0214***	0.0066	-0.0051**	0.0024	0.0015	-0.0066
Agreeableness	-0.003	0.0082	0.001	0.0029	0.0019	-0.0009
Conscientiousness	-0.0262***	0.0076	-0.0093***	0.0028	-0.0012	-0.0081
Emotional stability	-0.0237***	0.0075	-0.0064**	0.0027	0.0009	-0.0073
Openness to experience	0.0019	0.0070	0.0007	0.0024	0.0001	0.0006
Missing personality	-0.1667***	0.0142	0.0801	0.0550	0.1044	-0.0243

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

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

As for the intertemporal persistence of health, the probability of being in ill-health in the following year is about 12 percentage points higher for a family head that is in ill-health now than for a family head that is in good health. If a family is poor now, the probability that the family will be poor in the following year is about 13 percentage points higher compared with a family who is not poor now. The degree of intertemporal persistence in poverty appears to be similar to that in health. The effect of lagged poverty on current poverty is largely through the direct effect (12 per cent), as opposed to the indirect effect (1.3 per cent).

4.4. Alternative models

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 (i.e. random effects) and initial condition are still accounted for in the respective model. Panel (b) in Table 9 reports the MME estimates on the key variables from such a model, while Panel (a) in the table replicates the results in Table 8 for ease of comparison.

The simplest model that one can estimate is to ignore unobserved individual effects, the initial condition issue and the correlation between the equations altogether. This is equivalent to estimating a separate, simple dynamic probit model for health and for poverty respectively, by using the pooled data. The results from such a model are reported in panel (c) in Table 9.

Table 9: MME estimates of alternative models

	Health equation		Poverty equation	
	MME	S.e.	MME	S.e.
(a). Results from Table 8				
Ill-health			0.0186*	0.0108
lagged ill-health	0.1257***	0.0193		
lagged poor	0.0859***	0.0197	0.1335***	0.0254
(b). Uncorrelated dynamics with random effects & endogenous initial conditions				
Ill-health			0.0770***	0.0117
lagged ill-health	0.0872***	0.0148		
lagged poor	0.0487***	0.0160	0.0949***	0.0209
(c). Uncorrelated dynamics without random effects & exogenous initial condition				
Ill-health			0.0599***	0.0092
lagged ill-health	0.3895***	0.0175		
lagged poor	0.0276**	0.0115	0.4166***	0.0204

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

Compared with the MME estimate in Panel (a), the MME estimate in panel (b) on current health in the poverty equation is much larger; the difference is statistically significant at the 5 percent level. On the other hand, the MME estimates on the lagged poverty variable in both the health and poverty equations, and the lagged health variable in the health equation are smaller in Panel (b), where health is treated as exogenous, compared with the estimates in Panel (a), where health is treated as endogenous, and the differences of the estimates between the two models are statistically significant at the five per cent level.

Comparing Panel (c) with Panel (a), we also find that the estimate on current health in the poverty equation is much larger in Panel (c) than those in Panel (a), while the estimate on the lagged poverty variable is much smaller in Panel (c) than in Panel (a). The estimates on intertermporal persistence become much larger for both health and poverty in Panel (c) than in Panel (a). These comparisons suggest that incorrectly specified models may produce substantially biased results for the relationship between health and poverty.

4.5. Alternative poverty lines

The results reported in Tables 7 and 8 are obtained using the poverty line defined as a half of the median of the equivalised family income, where the OECD-modified equivalence scale is used for standardising family income. In Table a2 (Appendix) we report estimation results that use three different poverty lines: (a) 60 per cent of the median of the equivalised family income using the OECD-modified equivalence scale (as commonly used in Europe); (b) 50 per cent of the median of the equivalised family income using the Oxford equivalence scale; and (c) 50 per cent of the median of the equivalised family income using the square root equivalence scale.[7] The MME estimates vary among the different poverty lines, but the qualitative conclusion remains unchanged. That is, current health is found to have a significant effect on poverty, lagged poverty has a significant effect on health, and both health and poverty exhibit significant intertemporal persistence.

5. Conclusion

Using the HILDA Survey, this study examines the joint dynamics of health and poverty of Australian families. The joint modelling approach taken in this study, taking 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

[7] The Oxford scale assigns a value of one to the first adult family member, 0.7 to each additional adult and 0.5 to each child. The square root scale equivalises family income by dividing family income by the square root of the family size.

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	0.4679**	0.2070		
Ill-health			1.5574***	0.2628
Male	-0.0352	0.2333	0.2245	0.2447
Aged 18-25	-0.5434	0.3858	-0.4354	0.5175
Aged 26-35	-0.2933	0.2484	-0.0516	0.2376
Aged 46-55	0.0712	0.1849	-0.2731	0.1944
Aged 56+	0.2128	0.2109	-0.0818	0.2168
Degree	-0.3640	0.2329	-0.7557***	0.2489
Other post-sch qualification	-0.0177	0.1672	-0.4311**	0.1881
Year 12	-0.0564	0.2681	-0.4960*	0.2775
Overseas En-speak country	0.1783	0.2098	-0.0664	0.2310
Overseas Non-En country	-0.2817	0.2292	0.4068	0.2629
Couple without dependent children	-0.1817	0.1839	0.4918**	0.2106
Sole parent	-0.4640	0.3739	0.5851	0.3862
Lone person	0.0861	0.2068	0.7962***	0.2204
Inner region	0.1308	0.1692	0.5181***	0.1867
Outer region	0.4193*	0.2253	0.6866***	0.2316
Remote area	0.2034	0.5290	0.7727**	0.3674
Prop life-time emp	-0.0111***	0.0033	-0.0079**	0.0039
Father other white collar	0.3108	0.1895	-0.2106	0.2062
Father blue collar	0.1002	0.1644	-0.1714	0.1656
Father occupation unknown	0.4209	0.3917	-0.3404	0.5291
Physical functioning index	-3.9280***	0.3957		
Extroversion	-0.1442**	0.0731	0.0394	0.0758
Agreeableness	-0.0365	0.0827	0.0574	0.0857
Conscientiousness	-0.3214***	0.0790	-0.0092	0.0794
Emotional stability	-0.1171	0.0792	0.0120	0.0808
Openness to experience	0.1470**	0.0713	0.0158	0.0775
Missing personality	-1.9765***	0.6549	1.1824*	0.7184
Constant	4.7145***	0.8368	-2.4453***	0.8951
Coe. on random effects	1.2418***	0.1957	1.3258***	0.2720

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

Table a2: MME estimates of alternative poverty lines

	Health equation		Poverty equation	
	MME	S.e.	MME	S.e.
(a). OECD-modified scale, 60% median				
Ill-health			0.0325**	0.0149
lagged ill-health	0.1259***	0.0190		
lagged poor	0.0338**	0.0147	0.1445***	0.0227
(b). Square root scale, 50% median				
Ill-health			0.0207**	0.0115
lagged ill-health	0.1226***	0.0148		
lagged poor	0.0729***	0.0187	0.1292***	0.0229
(c). Oxford scale, 50% median				
Ill-health			0.0231***	0.0123
lagged ill-health	0.1156***	0.0184		
lagged poor	0.0344**	0.0165	0.0797***	0.0194

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

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 (1975) *First Main Report: Poverty in Australia*, Canberra: Australian Government Publishing Service.

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.

Grossman, M. (1972) "On the Concept of Health Capital and the Demand for Health." *Journal of Political Economy*, Vol. 80, p223-255.

Grossman and Benham, L. (1974) “Health, Hours and Wages”, in Perlman, M., ed., *The Economics of Health and Medical Care*, London: Macmillan, p205-233.

Haveman, R., Wolfe, B., Kreider, B. and Stone, M. (1994) "Market Work, Wages, and Men's Health." *Journal of Health Economics*, Vol. 13, p163-182.

Heckman, J. (1981). “The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process” in C.F. Manski and D.L. McFadden (eds.), Structural Analysis of Discreet Data with Econometric Applications, London: MIT Press.

Lee, L. (1982) "Health and Wages: A Simultaneous Equation Model with Multiple Discrete Indicators", *International Economic Review*, Vol. 23, p199-221.

Marmot, M. and Wilkinson, R. G. (1999) “Social determinants of health” (Oxford: Oxford University Press).

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.

Mundlak, Y. (1978), “On the Pooling of Time Series and Cross Section Data”, Econometrica 46: 69-85.

Saucier, G. (1994) 'Mini-markers: A Brief Version of Goldberg's Unipolar Big-five Markers', *Journal of Personality Assessment*, Vol. 63, p506-516.

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.

Stern, S. (1989) "Measuring the Effect of Disability on Labour Force Participation", *Journal of Human Resources,* Vol. 24, p361-395.

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

Ware, J., Snow, K., Kosinski, M. and Gandek, B. (2000). SF-36 Health Survey: Manual and Interpretation Guide, Lincoln, RI, Quality Metric Incorporated.

Wilkinson R. and Marmot M. (1998) *Social Determinants of Health: The Solid Facts*. Copenhagen: World Health Organization (www.who.dk/healthy-cities).

Wooden, M. Freidin, S., and Watson, N. (2002) “The Household, Income and Labour Dynamics in Australia (HILDA) Survey: Wave 1”, *Australian Economic Review* Vol.35(3)**,** p339-348.