Health Status and Labour Force Status of Older Working-Age Australian Men

Lixin Cai and Guyonne Kalb
Melbourne Institute of Applied Economic and Social Research
University of Melbourne

Keywords: Health status, labour force status, panel data, maximum simulated likelihood

JEL Classification: J14; J26

Abstract: This research uses the Household, Income and Labour Dynamics in Australia (HILDA) Survey to investigate the impacts of health on labour force status of older working-age Australian men. We estimate a model that exploits the longitudinal nature of the data and takes account of the correlation between the two error terms in the health and labour force status equations. The results show that controlling for unobserved heterogeneity and the correlation between the two equations is important. It is also found that any restriction on the correlation between the two equations appears to lead to underestimation of the direct health effects.

Acknowledgements: We would like to thank Jan van Ours, John Creedy, Mark Wooden, and workshop participants at the University of Melbourne and at Tilburg University for valuable comments and suggestions. This paper is based on research being conducted as part of the research program "The Dynamics of Economic and Social Change: An Analysis of the Household, Income and Labour Dynamics in Australia Survey". It is supported by an Australian Research Council Discovery Grant (DP0342970). The paper uses the data in the confidentialised unit record file from the Department of Family and Community Services' (FaCS) Household, Income and Labour Dynamics in Australia (HILDA) Survey, which is managed by the Melbourne Institute of Applied Economic and Social Research. The findings and views reported in the paper, however, are those of the authors and should not be attributed to either FaCS or the Melbourne Institute.

Corresponding author: Lixin Cai, Melbourne Institute of Applied Economic and Social Research, the University of Melbourne, VIC 3010 Australia. Tel: 61 (0)3 8344 2094. Fax: 61 (0)3 8344 2111. Email: Cail@unimelb.edu.au.

Health Status and Labour Force Status of Older Working-Age Australian Men

1. Introduction

The trend of declining labour force participation by older working-age men combined with an ageing population has led many industrialised nations to develop policies encouraging older male workers to remain in the labour force. A better understanding of how an individual's health influences the labour supply decision among this group of workers would facilitate the development of effective policies and results in a better estimate of the costs of health limitations to the economy (Chirikos, 1993; and Haveman *et al.*, 1992).

The effect of health on labour market activities of older male workers has been under extensive examination in the US and other industrialized countries (Chirikos (1993); Currie and Madrian (1999)). However, research on this issue in Australia is limited. Only two papers have examined the relationship between health and labour force status using Australian data. Wilkins (2004) examines the impact of disability on employment status using the 1998 Survey of Disability, Ageing and Carers (SDAC), a cross-sectional survey collected by the Australian Bureau of Statistics (ABS). The disability status in SDAC is derived from a combination of long-term health conditions and specific activity restrictions, and thus is a relatively narrow measure of health. Cai and Kalb (2004) use the first wave of the Household, Income and Labour Dynamics in Australia (HILDA) Survey to explore the relationship between health and labour force participation by employing a simultaneous equation model approach controlling for the potential endogeneity of health. However, the longitudinal nature of the data was not exploited by the study because only the first wave of the data was available when the study was conducted.

The current study contributes to the literature on this issue by using the longitudinal nature of the HILDA survey data. The three-wave HILDA data currently available allow a better control for unobserved heterogeneity, decomposing it into a time-variant and a time-invariant component. Therefore, more efficient estimates for the health effects on labour force status can be provided using these data than a cross-sectional analysis could. Since it is highly likely that unobserved factors affect both health and labour force status, we estimate a model that takes into account the correlation between the two error terms in the health and labour force participation equations. This controls for the potential endogeneity of the health variable resulting from unobserved heterogeneity. The results show that controlling for the unobserved heterogeneity and the correlation between the two equations is important. That is, the estimated variances of the unobserved heterogeneity terms are significantly different from zero in both equations and the two error terms are found to be correlated. It is also found that treating health as an exogenous variable leads to underestimation of its direct effect on labour force participation.

The paper is arranged as follows. Section 2 provides some theoretical background on the relationship between health and labour supply. The econometric model and estimation strategies are described in Section 3. Section 4 discusses the data and variables, followed by the estimation results in Section 5. Finally Section 6 concludes the paper.

2. Conceptual framework

There are many reasons why health may be an important factor in an individual's decision on labour supply. First, health is a determinant of productivity because an individual's capacity to fulfil a job's requirements is closely related to the person's health. Health is therefore often regarded as a form of human capital that is valued by both employers and employees (Becker, 1964; Grossman, 1972). As such, individuals with better health can command higher wages, which, other things equal, increases a person's labour force participation probability. Second, health may influence individuals' labour supply indirectly through shifting preferences between income and leisure. For example, poor health may lead individuals to value leisure time more, perhaps because the time needed to care for one's health increases with ill health or because the burden of work may increase with poor health. On the other hand, poor health may cause people to require more income due to costs associated with poor health. Third, because life expectancy is determined by health, changes in health may change the time horizon over which labour supply decisions are made. For example, poor health may make early withdrawal from the labour market more attractive (Chirikos, 1993).

Although health is predetermined partly as an endowment at birth, health is not exogenous over one's lifetime. Like other forms of human capital, people can make investments into health to make improvements or reduce the depreciation of the stock of health.[1] Because this investment requires resources, such as time for recreation or exercise and income for health care, health is endogenous in the sense that people have to make choices in the production of health. As such, people's health may also be affected by labour supply, especially a person's labour supply history, because past labour supply may determine the availability of resources that can be utilised to invest into health. There are also arguments that people's current health may be affected by current labour market activities. For example, boredom or general lack of activity in non-participation may lead to a deterioration of health (Stern, 1989; Sickles and Taubman, 1986). However, given the way in which the health stock evolves, it is unlikely that current labour market activity plays a significant role in current health status, although it may affect future health directly and indirectly.

The relationship between health and labour supply can be formalised in an economic model that is based on the assumption that individuals maximise an intertemporal utility function.[2] In such a model, health together with labour supply and consumption of other goods enter the utility function as endogenous choice variables. In principle, it is possible to derive demand functions for health, for leisure and for other consumption goods from this economic model. In particular, the model can be solved to yield a conditional labour supply function in which labour supply depends on the endogenous health variable and a set of other exogenous variables (Currie and Madrian, 1999). In this paper, we do not formally derive this model but simply specify the variables entering the labour force equation following the literature on labour supply.

3. Statistical model and estimation strategy

Based on the framework described in Section 2, an econometric model is derived for estimation. We adopt a recursive system where current health affects current labour force status but where there are no feedback effects from current labour force status to current health. We adopt such a model for the following reasons. First, as argued above, although the health stock can be endogenously augmented by investment and

[1] Unlike education, however, health as a form of human capital is subject to adverse shocks, such as the occurrence of illness or an accident, which may reduce health stock dramatically. Health investments can be used to improve health or to prevent adverse health shocks.
[2] An early version of this model is proposed by Grossman (1972). A more recent description of the model is given by Currie and Madrian (1999).

can therefore be affected by employment status, the change of health stock resulting from investments is expected to be so slow that it is unlikely that current employment will have much effect on current health. Workplace conditions may have a direct effect on health, but again the impact is likely to take time to become apparent. Second, to the extent that individuals make economic decisions rationally, labour force status should not directly modify the unobservable health stock (Sickles and Taubman, 1986). In fact, there has been no convincing supportive evidence showing that current labour force status has a significant effect on current health (Sickles and Taubman, 1986), especially for older male workers (Cai and Kalb, 2004).[3] However, employment and unemployment histories are included in the model and are allowed to influence both current labour force status and current health.

In addition, there is an important practical reason for the recursive modelling strategy. A full simultaneous equation model that also allows for the feedback effect would require exclusion restrictions to be imposed on each equation for the model to be identified. That is, each equation must include some exogenous variables that are excluded from the other equation and these variables should have a significant effect in the equation in which they are uniquely included. While there are obvious candidates for variables to be included in the health equation and excluded from the labour force status equation, this is not the case for the labour force status equation. In the model estimation, we have some variables in the labour force status equation that are not included in the health equation. However, these variables are not significant for older men and cannot be used as valid instruments for model identification. Furthermore, for logical consistency in simultaneous equation models with discrete dependent variables, at most one observed endogenous variable is allowed on the right-hand side in a two-equation system (Maddala, 1983).

Finally, it should be emphasised that the focus of our paper (as well as in the literature) is on the labour force status equation. The health equation mainly serves to control for

[3] Cai and Kalb's (2004) evidence is based on the first wave cross-sectional HILDA data. For the three wave panel data we experimented on estimating a simultaneous equation model using the same sample as in the current paper by treating both health and labour force as latent endogenous variables. For the disturbance in each equation we used the same error decomposition as in the recursive model. Using the two-stage maximum likelihood estimation method, which provides consistent estimator, we found that the coefficient estimate for the latent labour force variable in the second stage health equation was negative and insignificant at any conventional significance levels, while the estimate for the latent health variable is positive and strongly significant. These results provide empirically supportive evidence for the recursive model.

the endogeneity of the health variable resulting from unobserved heterogeneity. The system of equations is as follows:

(1) $$l_{it}^{*} = f(h_{it}) + x_{L,it}\beta_L + \varepsilon_{L,it}$$

(2) $$h_{it}^{*} = x_{h,it}\beta_h + \varepsilon_{h,it} \qquad (i = 1,\ldots,N;\ t = 1,\ldots,T).$$

where equation (1) is for labour force status and equation (2) for health determination; l_{it}^{*} and h_{it}^{*} are the latent dependent variables, representing the value of being in the labour force and health stock; l_{it} and h_{it} are their observed counterparts; $x_{h,it}$ and $x_{L,it}$ are the two sets of exogenous or predetermined variables; β_h and β_{L} are the structural coefficients corresponding to $x_{h,it}$ and $x_{L,it}$ respectively; $f(h_{it})$ is a function representing the effect of observed health on labour force status; and $\varepsilon_{h,it}$ and $\varepsilon_{L,it}$ are the structural disturbances.

The longitudinal nature of the data set is exploited by using a conventional error components specification for the two disturbances. Specifically, we decompose each of the two disturbances into two parts – a time-invariant component, $\mu_{j,i}$, and a time-variant component, $v_{j,it}$:

(3) $$\varepsilon_{j,it} = \mu_{j,i} + v_{j,it} \qquad (j = h,\ L).$$

Although $\mu_{j,i}$ can be estimated either as a fixed or as a random effect in a general framework of panel data models, the discrete nature of the observed dependent variables in our data necessitates the assumption of a random effect for the time-invariant components (see Hsiao (2003) for a detailed discussion of this issue). Clearly, the time-invariant component introduces a correlation among the disturbances for different time periods of the same equation:

(4) $$\operatorname{cov}(\varepsilon_{j,is}, \varepsilon_{j,lt}) = \begin{cases} \delta_{j(\mu)} + \delta_{j(v)} & for\ i = l,\ s = t \\ \delta_{j(\mu)} & for\ i = l,\ s \neq t, \\ 0 & for\ i \neq l \end{cases} \qquad \text{for } j = h,\ \mathrm{L};$$

where $\delta_{j(\mu)}$ is the variance of the time-invariant component and $\delta_{j(v)}$ is the variance of the time-variant component.

In addition, with regard to the correlation across the two equations, this may occur through different scenarios, each leading to a different variance-covariance structure for the disturbance terms. First, if the correlation only comes from the time-invariant component, that is, $\mathrm{cov}(v_{h,is}, v_{L,it}) = \delta_{hL(v)} = 0$ for all *s* and *t*, and $\mathrm{cov}(\mu_{h,i}, \mu_{L,i}) = \delta_{hL(\mu)} \neq 0$, the two equations are correlated both for the same time period and across different time periods,

(5) $\mathrm{cov}(\varepsilon_{h,is}, \varepsilon_{L,it}) = \delta_{hL(\mu)}$, for all *s* and *t*.

A second scenario is for the correlation to come from the same period time-variant component only, that is, $\mathrm{cov}(v_{h,is}, v_{L,it}) = \begin{cases} \delta_{hL(v)} \neq 0 & for\ s = t \\ 0 & for\ s \neq t \end{cases}$, and $\mathrm{cov}(\mu_{h,i}, \mu_{L,i}) = 0$.

In this case, the two equations are only correlated in the same time period and not correlated across time periods,

(6) $$\mathrm{cov}(\varepsilon_{h,is}, \varepsilon_{L,it}) = \begin{cases} \delta_{hL(v)} & for\ s = t \\ 0 & for\ s \neq t \end{cases}.$$

In the third alternative, the correlation comes from both components, that is $\mathrm{cov}(\mu_{h,i}, \mu_{L,i}) = \delta_{hL(\mu)} \neq 0$ and $\mathrm{cov}(v_{h,is}, v_{L,it}) = \begin{cases} \delta_{hL(v)} \neq 0 & for\ s = t \\ 0 & for\ s \neq t \end{cases}$.

Then,

(7) $$\mathrm{cov}(\varepsilon_{h,is}, \varepsilon_{L,it}) = \begin{cases} \delta_{hL(\mu)} + \delta_{hL(v)} & for\ s = t \\ \delta_{hL(\mu)} & for\ s \neq t\,. \end{cases}$$

Combining (4) with one of (5), (6) or (7) defines the variance-covariance matrix of the disturbance terms[4]. We further assume that the two components of each disturbance are independently normally distributed, which means that the two disturbances are jointly normally distributed. In the model, $\delta_{j(v)}$ with $j = h, L$ needs to be normalised to one for the model to be identified. Consequently, the covariance

[4] In principle, it is possible to allow the time-variant component to be correlated across different time periods of the same equation. For example, it could follow a first order autoregressive process. However, because we have only three wave data currently, not much gain is expected from this at the moment. It will be left for future research when more waves become available.

$\delta_{hL(\nu)}$ is equal to the correlation coefficient between the two time-variant error components.

The statistical model is similar to that specified by Sickles and Taubman (1986). However, Sickles and Taubman estimated the model based on the assumption that the correlation across the two equations is due to the same period time-variant component only; that is, the variance-covariance of the disturbance terms is defined by (4) and (6). The formula in (6) is based on the assumption that both health and labour force status are affected by shocks occurring in the same time period and the impact of the shocks would not last for more than one period. The assumption required in (5) may be more plausible than the one in (6) because the former implies that there may be some permanent unobservable personal attributes or preferences that affect both health and labour force status. Equation (7) takes both possibilities into account, leaving all options open, and thus is less restrictive. In our main specification of the model we use the variance-covariance structure resulting from (4) and (7), which is more general than that in Sickles and Taubman (1986).

The unobserved dependent variables underlying the model need to be linked to their observed discrete counterparts. In our data, five ordered health levels are available. Although unemployment can be identified from the data, we only distinguish two labour force states in the paper, participation and non-participation, with labour force participation including unemployment, because the proportion of unemployed is very small (only about 3 percent) in the sample.[5] The corresponding observed values of the dependent variables are:

$$(8) \qquad h_{it} = \begin{cases} 4 & (= excellent) & if \quad A_3 < h_{it}^* < A_4 = +\infty \\ 3 & (= very \quad good) & if \quad A_2 < h_{it}^* \leq A_3 \\ 2 & (= good) & if \quad A_1 < h_{it}^* \leq A_2 \\ 1 & (= fair) & if \quad A_0 < h_{it}^* \leq A_1 \\ 0 & (= poor) & if \quad -\infty = A_{-1} < h_{it}^* \leq A_0 \end{cases}$$

[5] An alternative division of labour force status is employed versus not employed. We prefer participation versus non-participation because whether an individual participates or not is a personal choice (that is, a labour supply decision), while whether a persons is employed or not also depends on labour demand. In other words, by using participation versus non-participation, we focus on individuals' labour supply decisions. However, we found that the results using the alternative division of employed versus not employed are in all respects very similar to the results in this paper.

(9) $$l_{it} = \begin{cases} 1 \quad (= in\ the\ \ labour\ \ force) \quad if \quad B_0 < l_{it}^* < B_1 = +\infty \\ 0 \quad (= not\ \ in\ the\ \ labour\ \ force)\ \ if \quad -\infty = B_{-1} < l_{it}^* \le B_0 \end{cases}.$$

Equations (1), (2), (8) and (9) constitute a system of equations. In addition, we specify

(10) $$f(h) = \alpha_1 \times excellent + \alpha_2 \times very\_good + \alpha_3 \times good + \alpha_4 \times fair.$$

The parameters to be estimated are the coefficients β_h and β_L in equations (1) and (2), α_1 to α_4 in equation (10); the cut-off points in equations (8) and (9), A_0 to A_3, and B_0 [6]; and the parameters in the variance-covariance matrix of the disturbance terms.

In principle, the parameters can be estimated using the maximum likelihood estimator (MLE) by assuming a joint density $f(\varepsilon_{h,i1},...,\varepsilon_{h,iT};\ \varepsilon_{l,i1},...,\varepsilon_{l,iT})$. However, a computational issue arises when $T \ge 2$ and an evaluation of at least a four-dimensional integral is required. In our case, with three wave data, an evaluation of six-dimensional integrals for observing each specific configuration of health-labour force state for the same individual over the three time periods would be required if traditional MLE were to be used. This renders such estimation infeasible.[7] However, recently developed simulation-based estimation techniques provide a feasible approach to overcome this problem (Stern, 1997; Lerman and Manski, 1981).

Simulation-based estimation procedures replace functions (usually definite integrals), which are computationally intractable when using numerical or analytical methods, with random approximations for these functions. There are generally two approaches to simulation-based estimators: direct simulation of the likelihood function or indirect likelihood simulation by using an expression for the score of the likelihood (Hyslop, 1999). In this paper we use the former approach, known as maximum simulated likelihood (MSL).[8] Following Hyslop (1999), let the log-likelihood function for the unknown parameter vector θ, given the random sample of observations (x_i, $i = 1,...,N$), be

[6] B_0 is normalised to zero after including a constant term in the labour force participation equation.
[7] See Hajivassiliou and Ruud (1994, p. 2399-2400) for an example of this.
[8] See Hyslop (1999) and Stern (1997) for a discussion on advantages and disadvantages of different simulation estimators.

(11) $$L_N(\theta) = \sum_{i=1}^{N} \ln(L(\theta; x_i)) .$$

Let $\{\xi_i\} = \{\xi_{i1}, ..., \xi_{iR}\}$ be a sequence of primitive simulators, independent of the parameters of the model and the data.[9] These are used in the following way:

(12) $$\tilde{L}(\theta; x_i, \xi_i) = (1/R)\sum_{r=1}^{R} \tilde{L}(\theta; x_i, \xi_{ri}),$$

where $\tilde{L}(\theta; x_i, \xi_{ir})$ is an unbiased simulator for $L(\theta; x_i)$ and R is the number of simulation replications. The maximum simulated likelihood estimator for θ is defined as:

(13) $$\hat{\theta}_{MSL} = \arg\max_{\theta} \sum_{i=1}^{N} \ln(\tilde{L}(\theta; x_i, \xi_i).$$

MSL estimation requires obtaining an unbiased simulator for the likelihood function. The simulator used here is the smooth recursive conditional (SRC) simulator, or Geweke-Hajivassiliou-Keane (GHK) simulator. This simulator is continuous in the parameters, strictly bounded by zero and one, unbiased and consistent in the number of replications R.[10] Although the GHK simulator is unbiased for the likelihood function, the resulting log (simulated) likelihood function will be biased with a finite number of replications due to the nonlinear logarithmic transformation. Therefore, the MSL estimator obtained will be inconsistent for finite R. However, MSL is consistent if the number of replications $R \to \infty$ as the sample size $N \to \infty$, and is asymptotically efficient if $R/\sqrt{N} \to \infty$ (Hajivassiliou and Ruud, 1994).

In addition, we use antithetic acceleration in simulating the random draws. These simulators use the original set of uniform random draws along with their reflections or mirror images to estimate the likelihood function.[11] Antithetic acceleration is a powerful variance reduction method (Geweke, 1988). Hajivassiliou (2000) presents Monte-Carlo evidence suggesting that the antithetically accelerated simulator for multivariate normal rectangle probabilities is superior to the standard GHK simulator.

[9] The asymptotic theory developed for the simulation-based estimators requires that the same fixed values of the primitive random draws be used at each iteration of the estimation.

[10] For a detailed discussion of the GHK simulator and its properties, see Hajivassiliou (1993), Börsch-Supan and Hajivassiliou (1993), Keane (1993, 1994) or Stern (1997).

[11] If z is a draw from the uniform distribution, $z_r = 1 - z$ is its reflection or mirror image.

In this paper 500 replications (that is, 250 random draws) are used in simulating the likelihood function. [12]

4. Data and variables

The data used for this paper come from the first three waves of the Household, Income and Labour Dynamics in Australia (HILDA) Survey. Details of this survey are documented in Watson and Wooden (2002). In the first wave, 7683 households representing 66 percent of all in-scope households were interviewed, generating a sample of 15,127 persons who were 15 years or older and eligible for interviews, of whom 13,969 were successfully interviewed. Subsequent interviews for later waves were conducted one year apart. In addition to the data collected through personal interviews, each person completing a personal interview was also given a self-completion questionnaire to be returned on completion by mail or handed back to the interviewer at a subsequent visit to the household. Attrition is a common problem with longitudinal survey data. The HILDA attrition rates for waves 2 and 3 were 13.2 percent and 9.6 percent respectively, which is not much higher than other longitudinal surveys. For the group of people studied in this paper, the attrition rates were lower than average (Melbourne Institute of Applied Economic and Social Research, 2005).

The HILDA survey contains detailed information on each individual's labour market activities and history. 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. Specific examples of these long-term conditions were shown on a card, examples of which are, limited use of fingers or arms, or problems with eyesight that could not be corrected with glasses or contact lenses. For those who had a long-term condition, impairment or disability, three follow-up questions were asked: whether the condition or disability limited the type of work or the amount of work; how much the condition or disability limited work; and whether the condition or disability first developed in the last 12 months.

[12] Relative to the sample size of 788 individuals, 500 replications is large. Initially, 300 replications were tried. Comparing the estimation results from 300 and 500 replications, there are some changes in the fourth digit after the decimal point for the estimated coefficients, but very few changes in the third digit.

In the self-completion questionnaire, the Short Form 36 health status questions (SF-36) 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-assessed health status question, with five levels scaled from poor to excellent. This self-assessed health status is used as the discrete observed counterpart of the latent health stock in the model.

There is a concern in the literature regarding the use of subjective health measures in estimating the effects of health on labour supply[13]; that is, self-assessed health may be used to justify an individual's labour force status. For example, those not in the labour force may report poor health to justify their non-participation or the receipt of disability-related benefits. The consequence of justification is that when using self-assessed health to estimate the effect of health on labour force status, the impacts could be overestimated.[14]

However, the evidence on the justification hypothesis in the literature is not conclusive. Studies that use objective health measures such as subsequent mortality tend to find a smaller effect than that estimated from using subjective health measures (Parsons, 1982; Anderson and Burkhauser, 1984, 1985), suggesting that the justification hypothesis may exist. However, Bound (1991) shows that the smaller effect could be due to measurement errors in objective health measures in the sense that they are not perfectly correlated with those aspects of health that impact on labour supply (that is, work capacity). Using a simultaneous equation model Stern (1989) does not find evidence supporting the justification hypothesis. Using a similar model to Stern (1989), Cai and Kalb (2004) find that the latent labour force variable has an insignificant effect on latent health for older working-age Australian men. Using an instrumental variables approach, Dwyer and Mitchell (1999) also find no evidence for the justification hypothesis.

The literature that directly evaluates the reporting bias of the health variables is much less extensive and the existing evidence is again mixed. While Bazzoli (1985) concludes that the importance of self-assessed health may be overstated in early-

[13] For example, see Lambrinos, 1981; Myers, 1982; Parsons, 1982; Bazzoli, 1985; Anderson and Burkhauser, 1984, 1985; Stern, 1989; Bound, 1991; Kreider, 1999; and Dwyer and Mitchell, 1999.

[14] Bound (1991) argues that the overestimation resulting from the justification hypothesis may be offset by the error-in-variable bias if the health variable is also measured with errors. Therefore, the direction of the overall bias on the estimate for the self-assessed health variable is not clear and the magnitude of the bias may be small.

retirement studies, Boaz and Muller (1990) find that the early retirees do not exaggerate their health problems to justify retirement status. By benchmarking the self-assessed health of workers, Kreider (1999) finds that non-workers overstate their health problems. However, the underlying assumption in Kreider's study that workers provide unbiased reports on their health is questionable (Myers, 1982; Stern, 1989). If subjective measures of health reflect leisure preferences, workers may downplay their health problems because they enjoy working (Dwyer and Mitchell, 1999; Benitez-Silva *et al.*, 2000). Benitez-Silva *et al.* (2000) use the definition of disability by the US Social Security Administration (SSA) as a social standard to assess the potential bias of the self-assessed disability status in the Health and Retirement Survey (HRS). They find no evidence of reporting bias in the sense that the disability benefit applicants in HRS do not exaggerate their health problems, compared with the standards used by the SSA.

Furthermore, the assumption underlying the justification hypothesis may be questionable. The hypothesis is based implicitly on the assumption that individuals first make a labour supply decision, and then report a health status level to justify that decision. However, it could also be the case that individuals first assess their health status and then formulate a strategy for modifying their labour market behaviour that accommodates their judgment (Chirikos, 1993). This argument is taken by Santiago and Muschkin (1996, p304) for supporting their use of self-assessed health: "We were concerned about tapping into the effects of the cognitive dimension of disability. If individuals perceive themselves to be in poor health or unable to work, these perceptions may be the primary factors underlying labour market behaviours".

In addition, there is a large literature showing that self-assessed measures of health are good indicators of health in the sense that they are highly correlated with medically determined health status (Nagi, 1969; Maddox and Douglas, 1973; LaRue *et al.*, 1979; Ferraro, 1980) and also good predictors of mortality and morbidity (Mossey and Shapiro, 1982; Idler and Kasl, 1995; McCallum, Shadbolt and Wang, 1994; Connelly *et al.*, 1989). Furthermore, Gerdtham *et al.* (1999) show that a continuous health status measure constructed from categorical self-assessed health by the method of Wagstaff and Van Doorslaer (1994) is highly correlated with other continuous measures of health.

Given the above reasons, the fact that most studies that have included a health measure to explain labour supply decisions have employed self-assessed health as an explanatory variable, and the fact that our data do not have an objective measure of general health, we use self-assessed health in our model.

Before describing the variables included in the model, we tabulate labour force status against self-assessed health status in Table 1 for the sample. We define our sample to be males aged between 51 and 64 (inclusive) years in all three waves. People over 64 years of age are excluded because they are eligible for the Age Pension and therefore face different incentives compared to those of working age. People over 64 years of age are largely out of the labour force.

A positive (negative) relationship between labour force participation (non-participation) and health status appears from the simple tabulation. Specifically, the proportion of older working-age Australian men who are not in the labour force decreases with health status for all three waves of data. That is, the better their health, the more likely an individual is to be in the labour force. For example, while over 81 percent of older working-age Australian men who reported poor health in the three waves of the survey were not in the labour force, only 15 percent who reported very good health were not in the labour force.

Table 1: Labour force status by self-assessed health status of older working-age Australian men (column percentages)

	Health status					
Labour force status[(a)]	Poor (0)	Fair (1)	Good (2)	Very good (3)	Excellent (4)	All
			Wave 1			
% not in labour force	71.43	54.03	18.39	11.99	16.09	24.87
% in labour force	28.57	45.97	81.61	88.01	83.91	75.13
Observations	49	124	261	267	87	788
			Wave 2			
% not in labour force	88.37	46.45	21.11	15.6	11.43	27.16
% in labour force	11.63	53.55	78.89	84.4	88.57	72.84
Observations	43	155	270	250	70	788
			Wave 3			
% not in labour force	83.33	55.1	22.03	17.96	12.9	29.95
% in labour force	16.67	44.9	77.97	82.04	87.1	70.05
Observations	48	147	286	245	62	788
			Waves 1 to 3			
% not in labour force	80.71	51.64	20.56	15.09	13.7	27.33
% in labour force	19.29	48.36	79.44	84.91	86.3	72.67
Observations	140	426	817	762	219	2,364

Note (a): In labour force includes employed and unemployed persons.

Although both labour force participation and health status are persistent variables, where the majority of respondents remains in the same category from one year to the next year, a substantial proportion (just under 50 per cent) change health status and a lower proportion (around 10 per cent) change labour force status.

Table 2 provides the definitions for all variables and shows in which equation(s) they are included. The identification condition for a recursive model like the one specified in this paper is that x_h includes some variables that are not included in x_L (Maddala, 1983, p. 122). As shown in Table 2, this condition is satisfied.

The variables included in the labour force status equation are standard in the literature of labour supply determination. However, for the health equation, some justification may be needed for the inclusion of some of the variables listed. Age is included because it is often observed that health deteriorates with age (Kenkel, 1995). Australian survey data show that the disability incidence rate increases with age (ABS, 1998) and the probability of entering the disability benefit program is higher among older people than among younger people (Cai and Gregory, 2003). In other studies, it has been noticed that health and marital status are closely related (for example, see Wilson and Oswald (2005) and references therein). Although there are different hypotheses about the mechanism through which this relationship is established, health is often observed to be positively correlated with being married. Education may improve health through enhanced awareness of health-related knowledge. In addition, it may serve to help control for the impact of parental socio-economic status (SES) on an individual's health since individual educational achievement is influenced by the parents' SES. Finally, education and age are key factors in the determination of an individual's wage. Since wage is only observed for labour force participants and (potential) wage is likely to influence health, we include education and age as instruments for wage in both the health and labour force equations. The variable *capital city* is included because individuals living in capital cities may have better access to health services, which would have a positive effect on health, although living in a capital city can be more stressful at the same time, which would affect health negatively.

To explain the difference in self-assessed health status between individuals, it would be ideal to have some specific and objective health indicators in the health equation, such as symptoms, types and severity of disability or health conditions. Unfortunately, such detailed objective measures are not available in the HILDA survey. However,

Table 2: Variable definition

Endogenous variables	
labour force status	0 non-participation, 1 participation
health	self-assessed health status, 0=poor, 1=fair, 2=good, 3=very good, 4=excellent
Variables appearing in both equations (x_h and x_L)	
Demographic and education	
age	age deviation from 50
married	1 if married or de facto
indigenous	1 if indigenous or Torres Strait Islander
degree	1 if has a bachelor or higher degree
other post-sch qual	1 if has other non-degree post-school qualifications
completed year 12	1 if highest education completed is year 12
year 11 or lower	1 if highest education completed is lower than 12
Job history, occupation and spouse's labour force status	
employ history	years in employment since first leaving full-time education
unemploy history	years in unemployment since first leaving full-time education
white collar 1	1 if last or current job as a manager, administrator or professional
white collar 2	1 if last or current job as a clerical, sales or service worker
blue collar	1 if last or current job as a tradesperson, labourer, production or transport worker or related worker
spouse in LF	1 if married and the spouse in the labour force
Income and whether living in a capital city	
own NL income	monthly income from investment (for example interest and dividend), private transfer and windfall income in previous financial year
spouse income	monthly spouse's income in previous financial year
capital city	1 if living in a capital city(a)
Additional variables appearing in the labour force participation equation (x_L)	
Health as categorical variable	
excellent health	1 if self-assessed health status = 4
very good health	1 if self-assessed health status = 3
good health	1 if self-assessed health status = 2
fair health	1 if self-assessed health status = 1
poor health	1 if self-assessed health status = 0
Demographic	
born overseas	1 if born overseas
born non-En country	1 if born in a non-English speaking foreign country
poor English	1 if spoken English is poor
child 0-14	1 if has child(ren) aged 0 to 14
married*child 0-14	interaction between married and child 0-14
Additional variables appearing in the health equation (x_h)	
smoker	1 if currently smoking or ever smoked
health condition	1 if has long-term health conditions
lack physical activity	1 if lack of physical activity, defined as no physical activity at all or less than once per week
heavy drinker	1 if a heavy drinker, defined as drinking more than 6 standard drinks a day when drinking
physical functioning	Index of physical functioning, ranging from 0 to 100.

Note (a), The capital cities do not include Hobart and Darwin because they cannot be identified from the data.

two summary indicators of health problems are available in the data and are included in the health equation. The first is the existence of long-term health conditions (*health condition*). This variable is included although it is self-assessed, following a suggestion by Bound, Schoenbaum and Waidmann (1995) that it is reasonable to treat self-reports of chronic health conditions as exogenous. Bound, Schoenbaum and Waidmann (1996) argue that survey questions that are more specific and concrete should be less subjective and therefore less susceptible to the rationalisation endogeneity problems. The variable *health condition* is used here in a similar way to Stern (1989). The difference is that Stern (1989) includes a list of specific long-term health conditions in his health equation, while we only have a summary indicator. However, HILDA respondents are shown a card listing specific examples of these conditions when answering this question. The second is one of the SF-36 indices, the index for physical functioning. Because this index is constructed based on individuals' answers to the questions about specific physical functioning limitations, such as climbing one flight of stairs, lifting or carrying groceries, or bending, kneeling or stooping, it can be treated as an exogenous variable. The index value ranges from 0 to 100, with 0 indicating that there is no physical functioning limitation (see Ware *et al.*, (2000) for the construction and interpretation of the index).

By including being a smoker, a heavy drinker and lack of physical activities variables in the health equation only, we assume that they affect labour force status only through their impacts on health. Smoking has been used as a rate of time preference indicator (Barsky *et al.*, 1997), but there is no strong evidence here to assume that smoking or drinking affects labour force status except indirectly through health (see Cai and Kalb, 2004).

The impact of unemployment on health has been discussed frequently (Wilson and Walker, 1993; Jin, Shah and Svoboda, 1995; Mathers and Schofield, 1998) and is included in our model as well. The inclusion of an employment history variable can also be justified in theory although its expected effect is ambiguous. On the one hand, employment may put stress on individuals or bad work conditions may be harmful to health; on the other hand, employment may make people happier and enhance self-confidence which could have a positive effect on health. In addition, wealth and income, which are important factors in health determination, depend to some extent on employment experience. Family income and wealth depend at least partly on the spouse's employment and furthermore the spouse's participation in the labour force may release pressure from the individual by reducing the risk of unemployment. Therefore, the spouse's labour force participation could have an impact on an individual's health status. Although we do not directly control for wages in the labour force status equation, two income variables, the individual's own non-labour income and the spouse's income (if married), are included to control for the impacts of non-labour income on labour supply.

In addition to the categorical health variables, a few other variables in the labour force status equation are not included in the health equation. These are being an immigrant, the ability of speaking English and the presence of children. Firstly, being an immigrant, especially from a non-English speaking background, and poor proficiency in speaking English may affect labour force status through language or cultural problems but it is unlikely that the effects on health can be identified if subgroups of countries of origin, with for example different levels of health care and different propensities for diseases, cannot be distinguished. For aggregate groups of immigrants, the overall effect may be small, given that the effects for individuals from different countries may offset each other. In addition, as the duration since the immigrants' arrival in Australia increases, their health status will become more similar to the health status of individuals born in Australia. Similarly, marital/de facto status interacted with the presence of children is likely to affect labour force status, but it is unlikely to affect an individual's health status in a specific direction.

Table 3 presents variable means for three waves of the sample used in the analysis. As expected, given the ageing of our sample, the proportion in the labour force decreases from wave 1 to wave 3, while the mean health deteriorates. The decrease in mean health is mainly due to the proportion reporting excellent or very good health decreasing, increasing the proportion in good or fair health, rather than increasing the proportion reporting poor health.

5. Estimation results

Tables 4 and 5 present the estimation results for the labour force status equation and health equation respectively. These results are based on the variance-covariance structure of (4) and (7), which imposes the least restrictions on the two disturbance terms among the scenarios specified earlier. This is therefore our preferred specification.

The first subsection presents the results on the labour force status equation, followed by the health equation results in subsection 5.2. The correlation coefficients are discussed in 5.3. In subsection 5.4, we compare the estimated health effects from the preferred specification with those from alternative specifications. For ease of interpretation, the table also reports the mean marginal effects (MME) based on marginal distributions.[15]

Table 3: Descriptive statistics of the sample

	Wave 1		Wave 2		Wave 3		Wave 1 to 3	
	Mean	*Std.Dev*	Mean	*Std.Dev*	Mean	*Std.Dev*	Mean	*Std.Dev*
labour force	0.75	*0.43*	0.73	*0.45*	0.70	*0.46*	0.73	*0.45*
health	2.28	*1.05*	2.19	*1.03*	2.16	*1.02*	2.21	*1.03*
excellent health	0.11	*0.31*	0.09	*0.28*	0.08	*0.27*	0.09	*0.29*
very good health	0.34	*0.47*	0.32	*0.47*	0.31	*0.46*	0.32	*0.47*
good health	0.33	*0.47*	0.34	*0.47*	0.36	*0.48*	0.35	*0.48*
fair health	0.16	*0.36*	0.20	*0.40*	0.19	*0.39*	0.18	*0.38*
poor health	0.06	*0.24*	0.05	*0.23*	0.06	*0.24*	0.06	*0.24*
age	55.7	*3.7*	56.6	*3.7*	57.6	*3.7*	56.6	*3.8*
married	0.81	*0.39*	0.82	*0.39*	0.82	*0.39*	0.82	*0.39*
child 0-14	0.13	*0.34*	0.11	*0.31*	0.10	*0.30*	0.11	*0.32*
married*child 0-14	0.11	*0.32*	0.09	*0.29*	0.08	*0.27*	0.10	*0.29*
born overseas	0.30	*0.46*	0.30	*0.46*	0.30	*0.46*	0.30	*0.46*
born non-En country	0.13	*0.34*	0.13	*0.34*	0.13	*0.34*	0.13	*0.34*
poor English	0.004	*0.062*	0.004	*0.062*	0.006	*0.079*	0.005	*0.068*
indigenous	0.01	*0.11*	0.01	*0.11*	0.01	*0.11*	0.01	*0.11*
degree	0.20	*0.40*	0.20	*0.40*	0.20	*0.40*	0.20	*0.40*
other post-sch qual	0.42	*0.49*	0.42	*0.49*	0.43	*0.49*	0.42	*0.49*
completed year 12	0.07	*0.25*	0.07	*0.25*	0.06	*0.25*	0.07	*0.25*
year 11 or lower	0.31	*0.46*	0.31	*0.46*	0.30	*0.46*	0.31	*0.46*
white collar 1	0.35	*0.48*	0.34	*0.48*	0.34	*0.48*	0.34	*0.48*
white collar 2	0.30	*0.46*	0.30	*0.46*	0.31	*0.46*	0.30	*0.46*
blue collar	0.35	*0.48*	0.36	*0.48*	0.36	*0.48*	0.36	*0.48*
smoker	0.63	*0.48*	0.65	*0.48*	0.66	*0.47*	0.65	*0.48*
health condition	0.35	*0.48*	0.33	*0.47*	0.41	*0.49*	0.36	*0.48*
heavy drinker	0.05	*0.21*	0.06	*0.24*	0.06	*0.24*	0.05	*0.23*
lack of physical activity	0.13	*0.34*	0.14	*0.34*	0.14	*0.35*	0.14	*0.34*
physical functioning	79.4	*24.5*	79.2	*23.8*	78.1	*24.7*	78.9	*24.3*
living in capital city	0.53	*0.50*	0.53	*0.50*	0.54	*0.50*	0.53	*0.50*
employ history	36.48	*6.16*	37.21	*6.22*	37.90	*6.29*	37.20	*6.25*
unemploy history	0.55	*1.94*	0.60	*2.01*	0.63	*2.06*	0.59	*2.00*
spouse in LF	0.53	*0.50*	0.48	*0.50*	0.47	*0.50*	0.49	*0.50*
spouse income	1620	*2244*	1820	*3100*	1870	*3603*	1770	*3035*
own NL income	344	*1707*	552	*2448*	755	*4082*	550	*2923*
number of observations	788		788		788		2364	

[15] The marginal effects are calculated for each observation, and then averaged to obtain the mean marginal effects (MME) over the sample. The standard errors of the MMEs are calculated using the Delta method (Greene, 1993).

Table 4: Estimation results for the labour force status equation

Variables	Coefficient estimates		Mean marginal effect (MME) estimates	
	Coef.	s.e.	MME (L=1)	s.e.
constant	-8.7319***	0.9506	0.7114[a]	
excellent health	6.5673***	0.5083	0.7382***	0.0467
very good health	5.2730***	0.4062	0.6626***	0.0439
good health	3.8198***	0.3420	0.5173***	0.0380
fair health	2.2788***	0.2981	0.3047***	0.0333
age/5	-1.7950***	0.2175	-0.1750***	0.0158
born overseas	0.0597	0.2966	0.0058	0.0288
born non_En country	0.2056	0.4098	0.0196	0.0389
indigenous	-0.3442	1.0832	-0.0346	0.1123
married	0.1665	0.2735	0.0164	0.0273
child 0-14	-0.2679	0.6117	-0.0266	0.0618
married*child 0-14	0.447	0.7312	0.0421	0.0663
spouse in LF	1.2885***	0.2090	0.1301***	0.0208
degree	1.5715***	0.4037	0.1471***	0.0333
other post-sch qual	0.3826	0.2558	0.0404	0.0269
completed year 12	0.7046	0.4857	0.0722	0.0472
employ history/10	1.8435***	0.2635	0.1791***	0.0206
unemploy history	0.1443**	0.0573	0.0140**	0.0055
poor Eng	-0.1465	2.6361	-0.0144	0.2634
white collar 2	0.3518	0.2395	0.0352	0.0235
blue collar	0.8339***	0.2221	0.0796***	0.0202
capital city	0.2832	0.2277	0.0277	0.0222
own NL income/1000	-0.0485*	0.0270	-0.0047*	0.0026
spouse income/1000	-0.018	0.0473	-0.0018	0.0046

Note: a) average predicted probability in the sample.
***significant at 1%, ** 5% and * 10%.

5.1. The labour force equation

Firstly, the health categorical variables are each strongly significant in the labour force participation equation. The omitted health category is poor health. The results show that compared with people with poor health, people having any other health status have a higher probability of participation. For instance, the MME estimates indicate that everything else equal, a person with excellent health will have a probability of participation that is about 74 percentage points higher than a person with poor health. The health effect is not linear. While a change from poor to fair health increases the probability of participation by about 30 percentage points, a change from very good to excellent health only raises the probability of participation by about 8 percentage points. A change in health status at the lower end of the health scale has a much larger effect on the probability of participation than a change at the upper end of the scale. This finding justifies our specification in which health enters the labour force participation equation as a set of dummy variables rather than as a continuous variable.

The variable age is scaled for estimation purposes by dividing it by 5. The sign of the age variable is as expected: it has a negative effect on participation. From the MME estimate, holding other things constant, an increase in age by 5 years reduces the probability of labour force participation by about 18 percentage points.[16]

The coefficient on the variable *spouse in LF* compares persons who are married to spouses in the labour force with those married to spouses not in the labour force. This variable is strongly significant with a positive sign, implying that compared with an older male married to a spouse not in the labour force, an older male married to a spouse in the labour force will have a labour force participation probability that is 13 percentage points higher. This result is consistent with the observation that there are few married couples where only the wife participates in the labour force. An explanation for this result may be that leisure time of older married couples is complementary in the sense that couples tend to spend leisure time together (Blau and Riphahn, 1999; Coile, 2004).

For the educational variables, the omitted category is the lowest level: those who did not complete year 12. The educational variables are jointly significant at the 5% significance level and have the expected signs. However, only the degree variable is individually significant. The MME estimate indicates that compared with a man who did not complete year 12, but is otherwise the same, an older male with a degree has a labour force participation probability that is 15 percentage points higher.

Employment history appears to be important to explain current participation. Theoretically, the sign of the employment history variable is ambiguous. On the one hand, the longer one has been employed, the more income one has earned (and saved) and the person might prefer to consume more leisure now. This is the income effect. On the other hand, the longer one is employed, the more experience one has and the higher the earnings that can be commanded. As a result, the person might be less likely to leave the labour force. This is the opportunity cost effect. The estimated sign

[16] In an alternative specification, we included a squared age variable in the model. In the labour force equation, the age variable was rendered weakly significant (negative) and the squared age variable was weakly significant (negative) as well. Both age variables were insignificant in the health equation. The estimates for the other coefficients were very similar to those presented here.

on this variable suggests that employment history has a positive impact on labour force participation. That is, the opportunity cost effect dominates the income effect. The results from the MME estimate indicate that a 10-year increase in past employment raises the probability of participation by about 18 percentage points. Theoretically, the unemployment history has an ambiguous impact as well, due to a similar line of argument. The estimated result indicates that the income effect dominates the opportunity cost effect. One additional year of unemployment in the past raises the probability of participation by 1.4 percentage points.

Occupation is also important in older men's labour force participation. For men, who are not employed at the time of the interview, the occupation in the last job is used. Only one man in the sample had never worked and was dropped from the analysis. The omitted occupational category is white-collar category 1, representing managers, administrators and professionals. While the occupational variables are jointly significant at the 5% level, only the blue-collar variable is individually significant. The MME estimate shows that compared with managers, administrators and professionals, older blue-collar male workers have a probability of participation that is 8 percentage points higher. One explanation is that blue-collar workers might not have been able to save enough for their retirement due to relatively low earnings in their types of jobs.[17]

For the two income variables, only the own non-labour income variable is weakly significant and has the expected sign. The MME estimate for this variable indicates that a $1000 increase in monthly own non-labour income reduces the probability of participation by less than 0.5 percentage point. This effect is very small from an economic point of view.

The variables, whether having children under 15 years of age, English speaking ability, whether living in a capital city and spouse's income, also have the expected signs, but these are insignificant.[18]

5.2. The health equation

[17] To test the argument of whether the positive estimate for the blue-collar variable could have been driven by more blue-collar workers being unemployed while in the labour force, we estimated a model where the labour force status was divided into employed and not employed. The estimated effect of the blue-collar variable remained positive and significant, suggesting that this was not a plausible explanation.

[18] Only 3 to 4 men per wave are classified as having poor English skills.

The estimation results for the health equation are presented in Table 5. In the literature on the effects of health on labour force status, the determination of health is not the focus. The health equation, if estimated, is to account for the endogeneity of self-assessed health resulting from unobserved variables (for example, Stern, 1989), or to construct a health index to be used in the labour force status equation (for example, Campolieti, 2002). Although in this paper, the main aim of estimating the health equation is to control for the endogeneity of the health variable as well, it is worthwhile to discuss briefly the estimation results for the health equation.

Except the variable *married* all significant variables have the expected sign.[19] Consistent with the theory that health depreciates with age (Grossman, 1999), we obtain a negative coefficient on the age variable, although it is insignificant.[20] However, the MME estimate for this variable is weakly significant for the probability of being in poor or very good health, with opposite signs. Also consistent with established findings that indigenous people have poorer health than non-indigenous Australians (Gray, Hunter and Taylor, 2003), the variable indicating indigenous status has a negative coefficient and is weakly significant. An older working-age man married to a spouse in the labour force has better health than an older man married to a spouse who is out of the labour force. This is perhaps because there is less pressure on a man if his spouse is also working than if his spouse is not. The educational variables have the expected sign. Again, these variables are jointly significant at the 5% level, with the degree variable being strongly significant, indicating that older males with a degree have better health. This may be due to better knowledge on healthy lifestyles and to better working conditions during one's career.

The employment history variable is weakly significant and has a positive sign, implying that past employment experience increases the health of older working-age Australian men. However, it is not clear whether this occurs because more years of past employment enable more investments in health, or a longer employment history simply reflects the good health of a person. Past unemployment experience seems to have an adverse effect on health, but is not accurately estimated.

[19] Note that because the variable *spouse in LF* is essentially an interaction variable between marital status and spouses' labour force status, the estimate for the variable *married* actually measures the difference between people who are married to a spouse not in the labour force and single persons.
[20] Investment in health may offset the adverse effect of age on health.

Table 5: Estimation results for the health equation

Variables	Coefficient estimates		Mean marginal effect (MME) estimates									
	Coef.	s.e.	MME (h=0)	s.e.	MME (h=1)	s.e.	MME (h=2)	s.e.	MME (h=3)	s.e.	MME (h=4)	s.e.
age/5	-0.1397	0.0903	0.0057*	0.0035	0.0127	0.0078	0.0081	0.0077	-0.0153*	0.0091	-0.0112	0.0095
indigenous	-0.6521*	0.3963	0.0332	0.0267	0.063	0.0405	0.0233	0.0232	-0.0784	0.051	-0.0411*	0.0241
married	-0.5386***	0.1443	0.0201**	0.0087	0.0461***	0.0138	0.0359**	0.0155	-0.0529**	0.0207	-0.0493**	0.0198
spouse in LF	0.2535**	0.1096	-0.0111*	0.0062	-0.0236**	0.0108	-0.0133	0.0094	0.0291**	0.0141	0.0189*	0.0102
degree	0.5606***	0.1807	-0.0202***	0.0072	-0.0510***	0.0153	-0.0375	0.0234	0.0601***	0.019	0.0487*	0.0273
other post-sch qual	0.0594	0.1346	-0.0026	0.0057	-0.0058	0.0129	-0.0031	0.0076	0.0071	0.016	0.0043	0.0101
completed year 12	0.1384	0.2186	-0.0058	0.0088	-0.0133	0.0206	-0.0075	0.0138	0.0163	0.0252	0.0102	0.0177
employ history/10	0.2387*	0.1231	-0.0098***	0.0029	-0.0217**	0.0092	-0.0138	0.0135	0.0261***	0.0092	0.0192	0.016
unemploy history	-0.0463	0.039	0.0019	0.0019	0.0042	0.0037	0.0027	0.0022	-0.0051	0.0047	-0.0037	0.003
health condition	-1.0081***	0.0827	0.0377**	0.0158	0.1150***	0.0194	0.0466	0.0434	-0.1352***	0.0181	-0.0642**	0.0257
smoker	-0.3952***	0.1064	0.0153**	0.0067	0.0362***	0.0111	0.0245*	0.0125	-0.0432***	0.0157	-0.0328**	0.0138
heavy drinker	-0.0591	0.1685	0.0025	0.0073	0.0054	0.0156	0.0033	0.0093	-0.0065	0.019	-0.0047	0.0131
lack physical activity	-0.4547***	0.102	0.0205**	0.0088	0.0446***	0.0116	0.0205	0.0155	-0.0540***	0.0152	-0.0316**	0.0129
physical functioning	0.2968***	0.0149	-0.0122***	0.0038	-0.0270***	0.0029	-0.0171*	0.0089	0.0325***	0.0063	0.0239***	0.0087
white collar 2	-0.2067**	0.1048	0.0079	0.0051	0.0189*	0.0102	0.0131	0.0082	-0.0225*	0.0132	-0.0173*	0.0098
blue collar	-0.3202***	0.1181	0.0127*	0.0068	0.0297**	0.0122	0.0191*	0.0107	-0.0358**	0.0163	-0.0258**	0.0119
capital city	0.0562	0.1064	-0.0023	0.0044	-0.0051	0.0097	-0.0032	0.0065	0.0062	0.0117	0.0045	0.0088
own NL income/1000	-0.0025	0.0212	0.0001	0.0009	0.0002	0.0019	0.0001	0.0012	-0.0003	0.0023	-0.0002	0.0017
spouse income/1000	0.0177	0.0196	-0.0008	0.0009	-0.0016	0.0018	-0.0009	0.0011	0.002	0.0022	0.0013	0.0015
cut_1 (A_0)	-1.7086***	0.4775										
cut_2 (A_1)	0.5228	0.4805										
cut_3 (A_2)	2.6322***	0.4751										
cut_4 (A_3)	4.8887***	0.471										
$\delta_{L(\mu)}$	3.9856***	0.6845										
$\delta_{h(\mu)}$	1.6320***	0.1502										
$\delta_{hL(\mu)}$	-0.9560***	0.2206										
$\delta_{hL(\nu)}$	-0.6854***	0.0826										
log-likelihood	-3021.05											
No. persons	788											

***significant at 1%, ** 5% and * 10%.

The five health-related variables, which are used to identify the labour force status equation, all have the expected signs and are strongly significant except for the heavy drinking variable. Importantly, the results show that unhealthy behaviour, such as smoking or lack of physical activity, has a negative effect on people's health.

The occupational variables also have the expected signs and are significant, indicating that compared with professionals, managers and administrators, other white and blue-collar workers tend to have poorer health. However, it is not clear whether this occurs because working conditions of low-level occupations are "bad" and harmful to health, the workers in low-level occupations invest less in health, or less healthy people are more likely to work in low-level occupations.

The variables *spouse income* and *living in a capital city* have the expected signs, but are insignificant. The variable *own non-labour income* has an unexpected sign, but again it is insignificant.

5.3. The correlation between equations

The last panel in Table 5 also presents the parameter estimates for the variance-covariance matrix. All four parameters are strongly significant, indicating that there are efficiency gains from the flexible specification of the error terms and the use of panel data. The variances of the two time-invariant terms are sizable. In the labour force status equation, the variance of the time-invariant term accounts for 80 percent of the total error variance. In the health equation, the time-invariant effect accounts for over 60 percent of the total variance. Both the covariance of the time-invariant unobserved heterogeneity and the correlation between the time-variant error components are negative and strongly significant. The negative and significant correlation of the time-variant error components is similar to that found by Sickles and Taubman (1986).

The negative correlation and covariance between the unobserved time-variant and time-invariant determinants of health and labour force status suggest that these unobserved variables have opposite effects on health and labour force status. Consequently, if the negative correlation between the two equations were not controlled for, the direct effect of health on labour force participation would be underestimated (see the next subsection). However, it is not immediately obvious what these unobserved factors are. One possible example of a time-invariant unobserved variable is people's inherent attitude towards leisure, work and health. For instance, there are people who put a high priority on a healthy lifestyle and thus prefer more leisure time so that they can look after their health properly. On the other end of the scale, there are people who are highly ambitious with regard to their career, putting a high priority on work. This type of person is more likely to stay in the labour force even at an older age. As a result, they may also be more likely to put high pressure on themselves, which could negatively affect their health.[21] This inherent, individual unobserved variable therefore causes a negative correlation. Possible examples of time-variant unobserved variables are a temporary change in individuals' preferences, such that people withdraw from the labour force to engage in activities beneficial to health (for example, travelling); or a plant closure or a reduction in firm size rendering individuals unemployed but potentially with a decent payout or an attractive early retirement scheme, which could make it attractive to exit the labour force, while at the same time benefiting health.

In addition to the model reported here, alternative specifications were estimated to explore the sensitivity of the correlation between the error terms to such a change. For example, we dropped the spouses' labour force participation variable from both equations in one specification and we dropped the health variable from the labour force equation in another specification. In the first specification, the covariance coefficients of the time-invariant error components and the time-variant error components remain negative and significant, although the magnitude reduces, which is as expected. In the second specification, the covariance coefficients of the time-invariant error components and the time-variant error components became both positive and significant. This suggests that including the direct effect of health absorbs all positive correlation between health and labour force participation and controls for all factors that cause a positive correlation between the two equations. What remains in the error terms causes a negative correlation between the two equations.

[21] Wealth might be another example, given that wealthy people are more likely to retire earlier and also have better health. In the second wave of the HILDA, questions on wealth were asked. We therefore examined the effect of including net wealth from the second wave as a time-invariant variable in the model (in addition to the other two financial variables). The covariance of the time-invariant error components between the two equations was still significantly negative, with a size similar to the one reported in the paper. The change in other parameters was also negligible, but the effect of wealth was positive and significant at the 10% level in the health equation. These results suggest that wealth could not be the only source of time-invariant unobserved heterogeneity.

5.4. Alternative specifications of the correlation between equations

In this subsection, we look at the changes in estimated effects of health on labour force status when restrictions are imposed on the correlation between the two equations. The results reported in Tables 4 and 5 are those resulting from the correlation specification in (7). Alternative specifications in (5) or (6) are equivalent to imposing the restriction $\delta_{hL(\nu)} = 0$ or $\delta_{hL(\mu)} = 0$ on the equations in (7). For completeness, we also estimate the model imposing both restrictions, $\delta_{hL(\nu)} = 0$ and $\delta_{hL(\mu)} = 0$. The latter is equivalent to estimating equations (1) and (2) separately using probit and random effect ordered probit models respectively, and treating health as an exogenous variable in the labour force participation equation.

Before presenting the estimates from alternative specifications, Table 6 first provides the likelihood ratio tests on the restrictions imposed by the three alternative specifications. From the p-values, it can be seen that all restricted models are rejected at conventional significance levels, in favour of the model without any restrictions.

Table 6: Likelihood Ratio (LR) test for restrictions

	$\delta_{hL(\mu)} = 0$	$\delta_{hL(\nu)} = 0$	$\delta_{hL(\nu)} = 0$ $\delta_{hL(\mu)} = 0$
χ^2	19.3	37.9	38.1
P-value	0.0000	0.0000	0.0000

Table 7 presents the MMEs of the health variables from different specifications. The coefficient estimates for the different alternative specifications are presented in the Appendix. Compared to the specification from equation (7) (results reproduced in the first column in Table 7), any restriction on the correlation between the two equations reduces the estimated MMEs of the health variables, especially for the top two health categories (confirming our discussion in section 5.3). For example, the MME of very good health when $\delta_{hL(\nu)} = 0$ is imposed is 26 percentage points less than the MME estimated without such a restriction; the MME of fair health is about 17 percentage points less. The specification assuming there is no correlation between the two equations produces the smallest direct health effects. This finding suggests that treating (self-assessed) health as an exogenous variable in the labour force participation equation leads to underestimation of the direct health effects rather than overestimation, as would be implied by the justification hypothesis.

The marginal effects in Table 7 are based on the assumption that one can change health exogenously, without a change in the associated correlated unobserved heterogeneity terms of health and labour force participation. That is, health changes due to a change in the exogenous factors in the health equation (of course some exogenous factors also affect labour force participation directly, in which case the net effect may be different from the direct effect of health). If the change in health level is due to one of the unobserved factors, the correlation between the two equations should be taken into account to calculate the net effect. Given the relatively large negative estimates for $\delta_{hL(\mu)}$ and $\delta_{hL(\nu)}$, this is expected to make a substantial difference. Table 8 presents the predicted probability of labour force participation conditional on the health status. This accounts for the correlation of the unobserved terms. The reported conditional probabilities are averages over the sample.

Table 7: MMEs of health effects in alternative specifications (poor health is the base group)

	$\delta_{hL(\nu)} \neq 0$ $\delta_{hL(\mu)} \neq 0$	$\delta_{hL(\nu)} \neq 0$ $\delta_{hL(\mu)} = 0$	$\delta_{hL(\nu)} = 0$ $\delta_{hL(\mu)} \neq 0$	$\delta_{hL(\nu)} = 0$ $\delta_{hL(\mu)} = 0$
excellent health	0.7382***	0.5414***	0.4298***	0.3646***
s.e.	0.0467	0.0538	0.0555	0.0461
very good health	0.6626***	0.4910***	0.4037***	0.3490***
s.e.	0.0439	0.0504	0.0495	0.0426
good health	0.5173***	0.3946***	0.3353***	0.2922***
s.e.	0.0380	0.0464	0.0462	0.0422
fair health	0.3047***	0.2516***	0.2227***	0.1959***
s.e.	0.0333	0.0413	0.0394	0.0384

***significant at 1%, ** 5% and * 10%.

Table 8: Predicted conditional probability of labour force participation

	$\delta_{hL(\nu)} \neq 0$ $\delta_{hL(\mu)} \neq 0$	$\delta_{hL(\nu)} \neq 0$ $\delta_{hL(\mu)} = 0$	$\delta_{hL(\nu)} = 0$ $\delta_{hL(\mu)} \neq 0$	$\delta_{hL(\nu)} = 0$ $\delta_{hL(\mu)} = 0$
Predicted conditional probability of participation				
Poor health	0.57	0.45	0.49	0.45
Fair health	0.70	0.65	0.67	0.65
Good health	0.74	0.74	0.75	0.75
Very good health	0.77	0.81	0.79	0.80
Excellent health	0.77	0.83	0.80	0.82
Compared to poor				
Fair health	0.13	0.20	0.18	0.20
Good health	0.17	0.29	0.26	0.29
Very good health	0.20	0.35	0.30	0.35
Excellent health	0.20	0.38	0.31	0.36

Comparing the predicted conditional probability of people at different health levels in the lower section of Table 8, it can be seen that the implied effect of health is lower than that in Table 7.[22] In addition, comparing the most restricted specification in the last column with the least restricted specification in the first column, the effect of being poor on labour force participation is smaller in the least restricted specification. The difference between the results in Tables 7 and 8 lies in the inclusion of time-variant and time-invariant unobserved terms that are negatively correlated across the two equations. Nevertheless, the effect of health on labour force participation remains substantial even when the indirect effect from unobserved variables is taken into account.

6. Conclusion

This paper examines the health effects on labour force status of older working-age Australian men. We use a random-effect approach to exploit the panel data nature of the HILDA survey and to control for unobserved heterogeneity. In addition, we account for the endogeneity of the health variable, resulting from unobserved heterogeneity, by introducing a health equation and allowing the error terms in the health equation and the labour force status equation to be correlated.

The results show that the model estimated in the paper may provide efficiency gains because a relatively flexible variance-covariance structure was allowed for the error terms in the two equations. All terms in the specification are found to be significant, indicating that restricting some of these terms to zero would have reduced the efficiency of the estimates.

The estimates confirm the finding in the literature that health has a significant effect on labour supply. Importantly, comparing the different specifications of the correlation between the two equations, we find that any restriction on the correlation leads to an underestimation of the direct health effects. Particularly, treating health as

[22] The predicted probability of labour force participation conditional on poor health in Table 8 is higher than that observed from the data (in Table 1) and the predicted probability of labour force participation conditional on excellent health is lower. This is due to the correlation of health with observed heterogeneity. In Table 8 the probabilities are calculated for the whole sample, whereas in Table 1 the proportions are computed for subgroups with different levels of health. Those with poor health are likely to have other characteristics, which have a negative effect on labour force participation, whereas those with excellent health have other characteristics, which positively affect labour force participation. This reinforces the lower and higher probabilities of labour force participation of those in poor and excellent health respectively.

an exogenous variable, that is, to assume that the two equations are uncorrelated, substantially underestimates the direct health effects, which lends support to rejection of the justification hypothesis.

From a policy point of view, knowing the magnitude of health effects is important, because it can be used to estimate the indirect costs of health problems through a reduction in labour supply. The results presented in this paper show that the indirect costs of health problems may be underestimated when using models that treat health as an exogenous variable in the labour force status equation, compared with models that account for the endogeneity of the health variable, such as the one used in this paper. The larger indirect costs of health problems estimated from a more efficient model may lend support to investment in policies aimed at improving health, especially older workers' health.

References

Australian Bureau of Statistics (ABS), (1998), *Disability, Ageing and Carers Survey*, ABS cat, 4430.0.

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

--- (1985). "The Retirement-Health Nexus: A New Measure of an Old Puzzle", *Journal of Human Resources* 20(3): 321-330.

Barsky, R.B., Juster, F.T., Kimball, M.S., and Shapiro, M.D. (1997), "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study", *Quarterly Journal of Economics*, May, 537-579.

Bazzoli, G.J. (1985). "The Early Retirement Decision: New Empirical Evidence on the Influence of Health", *Journal of Human Resources* 20: 214-234.

Becker, G.S. (1964). *Human Capital*. New York, Columbia University Press.

Benitez-Silva, H., Buchinsky, M., Chan, H.M., Cheidvasser, S. and Rust, J. (2000). "How Large is the Bias in Self-assessed Disability?" *NBER Working Paper* No. 7526.

Blau, D.M. and Riphahn, R.T. (1999). "Labor Force Transitions of Older Married Couples in Germany", *Labour Economics* 6: 229-251.

Boaz, R.F. and Muller, C.F. (1990). "The Validity of Health Limitations as a Reason for Deciding to Retire", *Health Services Research* 25(2): 361-386.

Börsch-Supan, A. and Hajivassiliou, V.A. (1993). "Smooth Unbiased Multivariate Probability Simulators for Maximum Likelihood Estimation of Limited Dependent Variable Models", *Journal of Econometrics* 58: 347-368.

Bound, J. (1991). "Self-assessed Versus Objective Measures of Health in Retirement Models", *Journal of Human Resources* 26(1): 106-138.

Bound, J., Schoenbaum, M. and Waidmann, T. (1995). "Race and Education Differences in Disability Status and Labor Force Attachment in the Health and Retirement Survey", *Journal of Human Resources* 30 (Supplement): S227-S267.

--- (1996). "Race Difference in Labor Force Attachment and Disability Status", *The Gerontologist* 36(3): 311-321.

Cai, L. and Gregory, R. G. (2003). "Inflows, Outflows and the Growth of the Disability Support Pension (DSP) Program", *Australian Social Policy 2002-03*: 121-143.

Cai, L. and Kalb, G. (2004). "Health Status and Labour Force Participation: Evidence from the HILDA Data", *Melbourne Institute Working Paper* No. 4/04, University of Melbourne, forthcoming in *Health Economics*.

Campolieti, M. (2002). "Disability and the Labor Force Participation of Older Men in Canada", *Labour Economics* 9: 405-432.

Chirikos, T.N. (1993). "The Relationship between Health and Labor Market Status", *Annual Reviews of Public Health* 14: 293-312.

Coile, C. (2004). "Retirement Incentives and Couples' Retirement Decisions", *Topic in Economic Analysis & Policy* 4(1): Article 17.

Connelly, J.E., Philbrick, J.T., Smith, R.G., Kaiser, D.L. and Wymer, A. (1989), "Health Perceptions of Primary Care Patients and the Influence on Health Care Utilisation", *Medical Care* 27: S99-S109.

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: 3310-3415.

Dwyer, D.S. and Mitchell, O.S. (1999). "Health Problems as Determinants of Retirement: Are Self-Rated Measures Endogenous?" *Journal of Health Economics* 18: 173-193.

Ferraro, K.F. (1980). "Self-Ratings of Health among the Old and Old-Old", *Journal of Health and Social Behavior* 21: 377-383.

Gerdtham, U.G., Johannesson, M., Lundberg, L., and Isacson, D. (1999). "A Note on Validating Wagstaff and van Doorslaer's Health Measure in the Analysis of Inequality in Health", *Journal of Health Economics* 18: 117-224.

Geweke, J. (1988). "Antithetic Acceleration of Monte Carlo Integration in Bayesian Inference", *Journal of Econometrics* 38, 73-89.

Gray, M.C., Hunter, B.H. and Taylor, J. (2003). "Health Expenditure, Income and Health Status among Indigenous and Other Australians", in *The Economics of Indigenous Health,* ed. by Anderson, I., University of Melbourne Press, Melbourne.

Greene, W. H. (1993). *Econometric Analysis* (2nd ed.). New York: Macmillan Publishing Company.

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

--- (1999). "The Human Capital Model of the Demand for Health", in *Handbook of Health Economics*, ed. by Newhouse, J. P. and Culyer, A. J., Amsterdam: North-Holland.

Hajivassiliou, V. (1993). "Simulation Estimation Methods for Limited Dependent Variable Models", in *Handbook of Statistics*, ed. by Maddala, G. S., Rao, C. R. and Vinod, H. D., Amsterdam and New York: North-Holland: 519-144.

Hajivassiliou, V. (2000). "Some Practical Issues in Maximum Simulated Likelihood", in *Simulation-Based Inference in Econometrics: Methods and Applications*, ed. by Mariano, R., Schuermann, T. and Weeks, M. J., Cambridge University Press: 71-99.

Hajivassiliou, V. and Ruud, P. (1994). "Classical Estimation Methods for LDV Models Using Simulation", in *Handbook of Econometrics*, Vol. 4, ed. by Engle, R.

and McFadden, D., Amsterdam, Elsevier Science Publishers.

Haveman, R.H., Wolfe, B., Buron, L. and Hill, S. (1992). Productivity Losses from Health/Disability in the United States from 1973-1983: An Earning Capacity Approach. Ann. Meet. Am. Econ. Assoc., New Orleans.

Hsiao, C. (2003). *Analysis of Panel Data*, 2nd Edition, Cambridge University Press.

Hyslop, D.R. (1999). "State Dependence, Serial Correlation and Heterogeneity in Intertemporal Labor Force Participation of Married Women", *Econometrica* 67(6): 1255-1294.

Idler, E.L. and Kasl, S.V. (1995). "Self-Ratings of Health: Do They Also Predict Change in Functional Ability?" *Journal of Gerontology* 50B: 344-353.

Jin, R.L., Shah, C.P. and Svoboda, T.J. (1995), "The Impact of Unemployment on Health: A Review of the Evidence", *Canadian Medical Association Journal* 153: 529-540.

Keane, M.P. (1993). "Simulation Estimation for Panel Data Models with Limited Dependent Variables", in *Handbook of Statistics,* Vol. 11, ed. by Maddala, G. S., Rao, C. R. and Vinod, H. D., Amsterdam and New York: North-Holland.

--- (1994). "A Computationally Practical Simulation Estimator for Panel Data", *Econometrica* 62: 95-116.

Kenkel, D.S. (1995). "Should You Eat Breakfast? Estimates from Health Production Functions", *Health Economics* 4: 15-29.

Kreider, B. (1999). "Latent Work Disability and Reporting Bias", *Journal of Human Resources* 34(4): 734-769.

Lambrinos, J. (1981). "Health: A Source of Bias in Labor Supply Models", *Review of Economics and Statistics* 63 (2): 206-212.

LaRue, A., Bank, L., Jarvic, J. and Hewtland, M. (1979). "Health in Old Age: How Physicians' Ratings and Self-Rating Compare", *Journal of Gerontology* 34: 687-691.

Lerman, S. and Manski, C. (1981). "On the Use of Simulated Frequencies to Approximate Choice Probabilities", in *Structural Analysis of Discrete Data with Econometric Applications*, ed. by Manski, C. and McFadden, D. Cambridge, MIT Press: 305-319.

Maddala, G. (1983). *Limited Dependent and Qualitative Variables in Econometrics*, New York: Cambridge University Press.

Maddox, G. and Douglas, E. (1973). "Self-Assessment of Health: A Longitudinal Study of Elderly Subject", *Journal of Health and Social Behaviour* 14: 87-93.

Mathers, C. and Schofield, D. (1998). "The Health Consequences of Unemployment: The Evidence", *Medical Journal of Australia* 168: 178-182.

McCallum, J. Shadbolt, B. and Wang, D. (1994). "Self-Rated Health and Survival: a 7-year Follow-up Study of Australian Elderly", *American Journal of Public Health* 847: 1100-1105.

Melbourne Institute of Applied Economic and Social Research, (2005). "Household, Income and Labour Dynamics in Australia (HILDA) Survey Annual Report 2004", University of Melbourne.

Mossey, J.M. and Shapiro, E. (1982). "Self Rated Health: A Predictor of Mortality among the Elderly", *American Journal of Public Health* 72: 800-808.

Myers, R. (1982). "Why Do People Retire from Work Early?" *Social Security Bulletin* September: 10-14.

Nagi, S.Z. (1969). "Congruency in Medical and Self-Assessment of Disability", *Industrial Medicine and Surgery* 38: 27-36.

Parsons, D.O. (1982). "The Male Labour Force Participation Decision: Health, Reported Health, and Economic Incentives", *Economica* 49(193): 81-91.

Santiago, A.M. and Muschkin, C.G. (1996). "Disentangling the Effects of Disability Status and Gender on the Labour Supply of Anglo, Black, and Latino Older Workers", *The Gerontologist* 36 (3): 299-310.

Sickles, R. and Taubman, P. (1986). "An Analysis of the Health and Retirement Status of the Elderly", *Econometrica* 54: 1339-1356.

Stern, S. (1989). "Measuring the Effect of Disability on Labour Force Participation", *Journal of Human Resources* 24(3): 361-395.

---- (1997). "Simulation Based Estimation", *Journal of Economic Literature* 35(4): 2006-2039.

Wagstaff, A., and van Doorslaer, E. (1994). "Measuring Inequalities in Health in the

Presence of Multiple Category Morbidity Indicators", *Health Economics* 3: 281-291.

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

Watson, N. and Wooden, M. (2002). "The Household, Income and Labour Dynamics in Australia (HILDA) Survey: Wave 1 Survey Methodology." HILDA Project Technical Paper Series No 1/02. Melbourne Institute of Applied Economics and Social Research, University of Melbourne.

Wilkins, R. (2004). "The Effects of Disability on Labour Force Status in Australia", *Australian Economic Review* 37 (4): 359-382.

Wilson, C.M. and Oswald, A.J. (2005). "How Does Marriage Affect Physical and Psychological Health? A Survey of the Longitudinal Evidence", *IZA Discussion Paper* No. 1619.

Wilson, S.H. and Walker, G.M. (1993). "Unemployment and Health: A Review", *Public Health* 107: 153-162.

Appendix

Table a1: Estimated coefficients for alternative specifications of the correlation

	$\delta_{hL(\nu)} \neq 0$ $\delta_{hL(\mu)} = 0$		$\delta_{hL(\nu)} = 0$ $\delta_{hL(\mu)} \neq 0$		$\delta_{hL(\nu)} = 0$ $\delta_{hL(\mu)} = 0$	
	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.
Labour force equation						
Constant	-9.3989***	1.0947	-9.4657***	1.1333	-9.6248***	1.1608
excellent health	5.1042***	0.5326	4.0963***	0.532	3.5917***	0.471
very good health	4.3435***	0.4464	3.7398***	0.4306	3.3821***	0.4018
good health	3.2526***	0.3866	2.9413***	0.3866	2.7007***	0.38
fair health	1.9910***	0.3451	1.8596***	0.3344	1.7221***	0.3406
age/5	-2.2584***	0.2409	-2.4017***	0.2579	-2.5448***	0.2606
born overseas	0.1693	0.3576	0.1051	0.3741	0.1437	0.3839
born non_En country	0.2234	0.4918	0.2185	0.5022	0.2246	0.5186
indigenous	-0.7409	1.2791	-0.8285	1.3364	-0.9662	1.3538
married	0.0842	0.3061	0.0695	0.3283	0.0603	0.3297
child 0-14	-0.1935	0.7182	-0.1077	0.7126	-0.0839	0.7373
married*child 0-14	0.4215	0.866	0.235	0.8677	0.2117	0.8915
spouse in LF	1.5293***	0.2268	1.6797***	0.2391	1.7312***	0.2338
degree	2.2247***	0.451	2.4377***	0.485	2.5983***	0.4703
other post-sch qual	0.5206*	0.2933	0.5702*	0.299	0.6048**	0.3065
completed year 12	0.9803*	0.5802	1.0782*	0.6073	1.1294*	0.6341
employ history/10	2.3581***	0.2919	2.5321***	0.3057	2.6900***	0.3051
unemploy history	0.1462**	0.0602	0.1463**	0.0608	0.1495**	0.0628
poor Eng	-0.0982	2.8878	-0.2735	2.7926	-0.2507	2.8949
white collar 2	0.3111	0.2711	0.2863	0.2832	0.2715	0.2855
blue collar	0.8452***	0.2515	0.7632***	0.2621	0.7764***	0.2623
capital city	0.3707	0.2591	0.4383	0.2723	0.4641*	0.2768
own NL income/1000	-0.0600**	0.0274	-0.0651**	0.0258	-0.0708**	0.0276
spouse income/1000	-0.0112	0.0562	-0.0091	0.06	-0.0064	0.0512
$\delta_{L(\mu)}$	4.9806***	0.9666	5.4618***	1.082	5.9359***	0.4701

Table a1: Continued

	$\delta_{hL(\nu)} \neq 0$ $\delta_{hL(\mu)} = 0$		$\delta_{hL(\nu)} = 0$ $\delta_{hL(\mu)} \neq 0$		$\delta_{hL(\nu)} = 0$ $\delta_{hL(\mu)} = 0$	
	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.
Health Equation						
age/5	-0.1434	0.0896	-0.1403	0.0902	-0.1498*	0.088
indigenous	-0.6809*	0.3839	-0.6585*	0.3945	-0.6727*	0.381
married	-0.5367***	0.1419	-0.5348***	0.1433	-0.5361***	0.141
spouse in LF	0.2306**	0.1086	0.2538**	0.1097	0.2474**	0.107
degree	0.6054***	0.1793	0.5517***	0.1803	0.5824***	0.1779
other post-sch qual	0.0665	0.1344	0.0548	0.1345	0.0595	0.1334
completed year 12	0.1489	0.2189	0.1293	0.2173	0.1393	0.2149
employ history/10	0.2573**	0.1219	0.2375*	0.1245	0.2548**	0.1186
unemploy history	-0.0444	0.0388	-0.0477	0.039	-0.0474	0.0386
health condition	-0.9747***	0.0844	-0.9943***	0.0856	-0.9716***	0.084
smoker	-0.3626***	0.1089	-0.3808***	0.1083	-0.3656***	0.1051
heavy drinker	-0.0205	0.1667	-0.0142	0.1705	-0.0038	0.1671
lack physical activity	-0.4712***	0.1044	-0.4819***	0.1058	-0.4811***	0.1045
physical functioning	0.3003***	0.0149	0.3048***	0.0151	0.3017***	0.0146
white collar 2	-0.1776*	0.1039	-0.2120**	0.1043	-0.2007*	0.103
blue collar	-0.2757**	0.1159	-0.3294***	0.1177	-0.3128***	0.1148
capital city	0.0538	0.1053	0.0568	0.1063	0.0585	0.1047
own NL income/1000	-0.0034	0.0209	-0.0021	0.0197	-0.0025	0.0174
spouse income/1000	0.0178	0.019	0.0174	0.0196	0.0176	0.0181
cut_1 (A_0)	-1.5622***	0.4726	-1.6455***	0.4812	-1.5998***	0.46
cut_2 (A_1)	0.6837	0.4754	0.5913	0.4853	0.6504	0.4606
cut_3 (A_2)	2.7882***	0.4704	2.6939***	0.4796	2.7585***	0.4557
cut_4 (A_3)	5.0330***	0.4677	4.9593***	0.4754	5.0267***	0.4531
$\delta_{h(\mu)}$	1.6143***	0.1483	1.6267***	0.1491	1.6530***	0.1161
$\delta_{hL(\mu)}$			-0.3898	0.2493		
$\delta_{hL(\nu)}$	-0.5179***	0.0921				
log-likelihood	-3030.70		-3040.00		-3040.1	

***significant at 1%, ** 5% and * 10%.