Melbourne Institute Working Paper Series

Working Paper No. 2/07

Effects of Health on Wages of Australian Men

Lixin Cai

Effects of Health on Wages of Australian Men

Lixin Cai*
Melbourne Institute of Applied Economic and Social Research
The University of Melbourne

Melbourne Institute Working Paper No. 2/07

ISSN 1328-4991 (Print)

ISSN 1447-5863 (Online)

ISBN 978-0-7340-3232-4

February 2007

***** I would like to thank Bruce Bradbury, John Creedy, Guyonne Kalb and participants at the 9th Annual Labour Econometric Workshop in Adelaide for useful suggestions. Special thanks to Kostas Mavromaras for his very valuable comments. This research is supported by the University of Melbourne Early Career Researcher Grant Scheme. The paper uses the data in the confidentialised unit record file from the Australian Department of Families, Community Services and Indigenous Affairs' (FaCSIA) 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 author and should not be attributed to either FaCSIA or the Melbourne Institute.

Melbourne Institute of Applied Economic and Social Research
The University of Melbourne
Victoria 3010 Australia
***Telephone* (03) 8344 2100**
***Fax* (03) 8344 2111**
***Email* melb-inst@unimelb.edu.au**
***WWW Address* http://www.melbourneinstitute.com**

Abstract

As a form of human capital health like education determines individuals' productivity and thus wage rates. While there are numerous overseas studies that examine the effect of health on wages, research on this issue using Australian data is scarce. This paper uses the Household, Income and Labour Dynamics in Australia (HILDA) survey to investigate the effect of health on the wages of working-age Australian men. A simultaneous equation model of health and wages is estimated to account for endogeneity of health. The results confirm the finding in the literature that health has a significant and positive effect on wages, but the significant effect is found only when measurement error and endogeneity of health are accounted for. The reverse effect of wages on health is found insignificant, but there is evidence on the endogeneity of health arising from unobserved factors.

1. **Introduction**

As a form of human capital health like education determines individuals' productivity and thus wage rates. This study estimates the effect of health on the wages of working-age Australian men. Because wages may have a reverse effect on health and both health and wages can be affected by some unobserved common factors, health may be endogenous in wage determination models. To account for the potential endogeneity of health, I estimate a simultaneous equation model of health and wages.[1] The model is estimated using the Household, Income and Labour Dynamics in Australia (HILDA) survey, which is a longitudinal survey with four-wave data available currently. The study uses the third wave HILDA because this wave contains more detained information than the other waves on health conditions which are used to instrument the potentially endogenous health variable.

Human capital has long been recognized as a key driving force of productivity and thus economic growth. Traditionally human capital has been interpreted as education and skills. However, recently increasing attention has been given to health as a form of human capital.[2] For example, on the macro-level a growing number of studies have shown that population health has a significant and positive effect on economic development, particularly in low-income countries (Bloom and Mahal, 1997; Bloom and Sachs, 1998; Bloom and Canning, 2000; Bhargava et al., 2001; and Bloom, Canning and Sevilla, 2001; Rivera and Currais, 1999). The reason that population health is important in determining economic growth is that at the micro-level health like education affects individuals' productivity and labour supply.[3] For instance, numerous studies for both developed and developing countries have found that better health increases individuals' labour force participation and wages (e.g. Currie and

---

[1] A model that accounts for sample selection was also specified and estimated, but it turned out that sample selection bias was not an issue in the data used because the error terms in the selection (employment) equation and the wage equation were found insignificant. The description of the model that accounts for sample selection and the estimation results are presented in Appendix C.

[2] In his pioneering work on human capital, Becker (1964) draws an analogy between investment in health capital and in other forms of human capital such as education. This framework is further developed by Grossman (1972), where health is explicitly treated as an endogenous variable, together with individuals' consumption and labour supply. According to Grossman (1972), health like education increases the quality of individuals' private and market time, so that people with better health are able to do things more efficiently. In this sense health and the capacity of adequately performing job requirements are closely related. For example, poor health is likely to have an adverse effect on work performance.

[3] Health not only affects individuals' productive activity such as work, it also directly influences individual utility because better health means reduced pain and suffering (Haveman et al., 1994).

Madrian, 1999; Stern, 1989; Bound et al., 1999; Campolieti, 2002; Cai and Kalb, 2006; Lee, 1982; Haveman et al., 1994).[4]

Despite a large body of overseas literature on the effect of health on wages (see Currie and Madrian (1999) for an extensive review of the US studies), research using Australia data is scarce. There are only a few papers that examine either disability or alcohol consumption on earnings (Brazenor, 2002; Barrett, 2002; and Lye and Hirschberg, 2004); it seems there has been no study that estimates the effect of health *per se* on wages. This study provides empirical evidence using current Australian data. In addition, I use the full information maximum likelihood (FIML) method to estimate a simultaneous equation model to obtain more efficient estimation results than could obtained from a two-stage method and to infer the nature of the endogeneity of health.

The results of the paper confirm the common finding in the literature that health has a positive and significant effect on wages, but this significant effect is found only when endogeneity and measurement error of health are accounted for. The reverse effect from wages to health is found insignificant, but there is evidence on the endogeneity of health arising from unobserved factors.

The rest of the paper is arranged as follows. Section 2 briefly reviews the literature; Section 3 discusses the econometric model and estimation method; Section 4 describes the data and model specification; Section 5 presents the estimation results. The paper is concluded in Section 6.

2. The literature

The literature on the effect of health on labour market performance is large and has a long history. The labour market performance measures that have been examined include not only wages, but also earnings or incomes.[5] This study focuses on the effect of health on wages because wages are often used as a measure of individuals' productivity. Studies that use earnings as a measure of labour market performance do

---

[4] In addition to directly affecting wages, health may also affect labour force participation by changing individuals' preferences between leisure and work. For example, poor health may lead to people to value time out of labour markets more since the time needed to care for one's health increases with ill health.

[5] Here wages refer to the wage rate (i.e. earnings per hour), which is often used as a measure of individuals' productivity. Earnings are the product of wages and hours worked. Incomes include earnings, called employment incomes in labour economics, and non-employment incomes, such as interest or dividend earned from investments, and government transfers.

not distinguish the effect of health on productivity from that on labour supply. Because the interest of this study is in the effect of health on wages and a simultaneous equation model is used for such a purpose, this brief literature review only focuses on studies that employ a similar modelling strategy and use wages, as opposed to earnings or incomes, as the outcome variable. But this condition does not apply to Australian studies reviewed here because the current study is about Australia and there are only a few relevant Australian studies.[6]

The earliest paper examining the effect of health on wages in a simultaneous equation framework seems to be by Grossman and Benham (1974). A simultaneous equation system of (log) weekly wages, (log) annual weeks worked and health is specified in the econometric model to account for endogeneity of health; a two-stage method is used to estimate the system on a sample of white males over 18 years of age.[7] They find that the derived ill-health variable has a negative and significant effect on wages and wages have a positive effect on ill-health; the effect of wages on health is significant when persons over 64 years were included but insignificant when they were excluded. In particular, they find that accounting for endogeneity of health increases the estimated effect of health substantially.

Using the data from the National Longitudinal Survey of Men 45-59, Survey 1966, Lee (1982) estimates a simultaneous equation model of health and wages, with two discrete indicators for unobserved health capital. He finds that latent health has a positive and significant impact on (log) hourly wages. However, in contrast to Grossman and Benham (1974), Lee (1982) finds that the wage rate has a positive and significant effect on health and that controlling for the reverse effect of wages on health substantially reduces the estimated effect of health on wages.

Haveman et al. (1994) estimate a simultaneous equation model of (annual) hourly wages, annual hours worked and health status, using the Michigan Panel Study of Income Dynamics (PSID). However, while the variable hours worked is allowed to affect health in their model, the wage rate does not enter the health equation. Therefore, the reverse effect of wages on health is not examined in their study. In

[6] For other overseas studies particular in the US, see Currie and Madrian (1999).
[7] In their model wages and health are allowed to affect each other; wages and health are allowed to affect weeks worked; but weeks worked are not allowed to affect wages and health. The health variable is derived using the Principal Component Analysis of the variables the number of symptoms reported and self-reported health status.

addition, it is the one-year lagged health status that is allowed to affect wages.[8] The model is estimated on white men aged between 24 and 65 years using the GMM estimator. They find that lagged ill-health significantly reduces wages and that the effect is larger when endogeneity of health is accounted for.

There is no published Australian study that has estimated the effect of health on wages and only a few papers examine the impact of different health measures on earnings.[9] For example, using 1998 Australian Disability, Ageing and Carers Survey (DACS), Brazenor (2002) examines the impact of various disabilities, such as loss of sight, speech difficulty, nervous or emotional condition and disfigurement or deformity, on (log) weekly earnings and finds that disability has a negative impact on earnings and the effect varies with disability types. Barrett (2002) estimates the effect of alcohol consumption on the earnings of full-time male workers, using the 1989-90 Australian National Health Survey (NHS), and finds that moderate drinking leads to a significant earnings premium relative to abstention and heavy drinking. Lye and Hirschberg (2004) also use the 1989-90 NHS to examine the effect of alcohol consumption on earnings by emphasizing the interaction between smoking and drinking by estimating the effect separately for smokers and non-smokers. For smokers they find similar results to Barrett (2002), but no significant effect is found for non-smokers.

3. Statistical model and estimation method

3.1 The statistical model

Endogeneity of health is an issue that often concerns researchers. Endogeneity may arise from unobserved factors, such as preferences, which affect both health and wages, or may be caused by the reverse effect of wages on health. The direction of endogeneity bias depends on the nature of the endogienity. For example, if the unobserved factors affect both health and wages in the same direction, there would be an upward bias to the estimated effect of health if health was treated as exogenous.

The reverse effect of wages on health arises according to the well-known health production model of Grossman (1972), where it is argued that health capital can be maintained and improved through investments that depend on resources available,

[8] Both the wage rate and lagged health are allowed to affect hours worked.
[9] For studies that examine the effect of health on labour force status in Australia, see Cai and Kalb (2006) and Wilkins (2003).

including both economic resources and time. Higher wages imply that more economic input into health production is possible, implying a positive effect of wages on health. On the other hand, an increase in the return to health capital increases the opportunity cost of health investment, which leads to individuals being involved in market activity more heavily and less time input into health production, suggesting a negative effect of wages on health (Grossman and Benham, 1974). Therefore, the direction of simultaneity bias is ambiguous in theory and can only be determined empirically.

To account for the endogeneity of health in wage models, a simultaneous equation model is specified in this study. First, the wage rate is determined by equation (1),

(1) $w = \alpha_w h^* + x_w^{'} \beta_w + \varepsilon_w$,

where w is the (hourly log-) wages; h^* is latent health; x_w is a vector of variables that affect wages, such as education and work experience; and ε_w is a random error term summarizing all unobserved factors that affect wages. Equation (1) is a standard Mincer's (1974) model of wage determination, with health added as an additional explanatory variable. Latent health enters the wage determination equation because the underlying health capital is not directly measurable.

The health determination equation is,

(2) $h^* = \alpha_h w + x_h^{'} \beta_h + \varepsilon_h$,

where x_h is a vector of variables that affect health status, such as specific health conditions and heath risk behavior, and ε_h is the error term summarizing all unobserved factors that affect health.[10] w enters the health determination equation due to the potential reverse effect as discussed earlier.

The two error terms ε_w and ε_h in equations (1) and (2) are allowed to be freely correlated to account for the endogeneity of health arising from unobserved factors. Because endogeneity of health can arise from either the reverse effect (i.e. $\alpha_h \neq 0$) or the correlation of the two error terms, a true test on the exogeneity of health must

[10] It may be argued that what matters in determining individuals' health is (family) incomes rather than individuals' wages. But the focus of this study is on the effect of health on wages. The specification used in equation (2) provides the functional form that allows the endogeneity of health from both sources, as discussed in the text, to be controlled for. Whether it is incomes or wages that are more appropriate in determining health is out of the scope of the current study.

examine the significance of both the estimates for α_h and for the correlation between ε_w and ε_h . x_w and x_h are variables assumed to be exogenous, and therefore uncorrelated with the error terms. x_w and x_h may have some common variables, but for identification purposes each should contain at least one equation specific variable. Also note that the use of latent health h^*, instead of the five-level self-reported health status h available in the data, represents a way of purging measurement error from self-reported health status (Bound et al. 1999).[11]

Latent health h^* is not directly observable, instead we have the five-level self-reported health status. The following equation relates latent health to observed health status,

$$(3)\ \ h = \begin{cases} 4 \quad (= excellent) & if \quad m_3 < h^* < m_4 = +\infty \\ 3 \quad (= very\ good) & if \quad m_2 < h^* \le m_3 \\ 2 \quad (= good) & if \quad m_1 < h^* \le m_2 \\ 1 \quad (= fair) & if \quad m_0 < h^* \le m_1 \\ 0 \quad (= poor) & if \quad -\infty = m_{-1} < h^* \le m_0 \end{cases} ,$$

where the *m*s are the cut-off point parameters to be estimated.

Equations (1) to (3) consist of the simultaneous equation system of wage and health determination, which can be estimated using two methods: two-stage and full information maximum likelihood (FIML).

3.2 The two-stage method

Through substitution, the two equations (1) and (2) can be written in their reduced forms:

$$(4)\ \ w = \frac{1}{1-\alpha_w\alpha_h}[x_w^{'}\beta_w + x_h^{'}\beta_h\alpha_w + (\varepsilon_w + \alpha_w\varepsilon_h)] = x^{*'}\Pi_w + \varepsilon_w^*$$

$$(5)\ \ h^* = \frac{1}{1-\alpha_w\alpha_h}[x_w^{'}\beta_w\alpha_h + x_h^{'}\beta_h + (\alpha_h\varepsilon_w + \varepsilon_h)] = x^{*'}\Pi_h + \varepsilon_h^* \ ,$$

[11] To be specific, let $h^* = h^{**} + \mu$ and $h^{**} = \alpha_h w + x_h\beta_h + \eta$, where h^{**} is true health, μ reporting error, and η a random variable independent of w and x_h. By substitution equation (2) is obtained with $\varepsilon_h = \eta + \mu$ and the reporting error is removed from predicted health. For detailed discussion on this, see Bound et al. (1999).

where $x^* = (x_w \cup x_h)$, referring to all exogenous variables included in the system; Π_w and Π_h are the reduced form parameters; $\varepsilon_w^* = \frac{1}{1-\alpha_w\alpha_h}(\varepsilon_w + \alpha_w\varepsilon_h)$ and $\varepsilon_h^* = \frac{1}{1-\alpha_w\alpha_h}(\alpha_h\varepsilon_w + \varepsilon_h)$ are the reduced from error terms. The first stage is to estimate equations (4) and (5), using OLS and ordered probit respectively to obtain consistent estimates for the reduced form parameters, denoted as $\hat{\Pi}_w$ and $\hat{\Pi}_h$. With these estimates, predicted wages and health can be obtained using,

(6) $\hat{w} = x^{*'}\hat{\Pi}_w$,

(7) $\hat{h}^* = x^{*'}\hat{\Pi}_h$.

In the second stage predicted wages and health, $\hat{w}$ and $\hat{h}^*$, are substituted into equations (1) and (2) to estimate the structural equation parameters, again using OLS and ordered probit.

In addition to predicted latent health $\hat{h}^*$, using equation (3) and the estimates for the cut-off points from the first stage estimation of the health equation, health status or categories $\hat{h}$ can be predicted and used in the second stage wages equation as well. Estimation results using both predicted health variables are reported in the empirical result section.[12]

3.3 The FIML method

The two-stage method provides consistent estimates and is also easy to implement by using standard statistical packages such as STATA, SAS or LIMDEP. However, the estimates from the two-stage method are not efficient because information on the correlation between ε_w and ε_h is not used. To obtain efficient estimates, the full information maximum likelihood (FIML) method is required. In addition, from FIML

[12] While there are five levels of self-reported health status in the data, I use three levels for predicted health status because only a few observations are predicted to fall into the bottom and top health categories. That is, for predicted health status the lowest level includes both poor and fair health and the top level includes very good and excellent health. Another difference between using the two predicted health measures is that while the covariance matrix in the second stage can be derived analytically when predicted latent health is used, the covariance matrix cannot be computed analytically when predicted health categories are used. However, for both predicted health measures the bootstrap method is used to obtain the standard errors of the second stage estimates.

the nature of the correlation between the two error terms can be inferred and a true test on the exogeneity of health can be conducted.

The FIML method requires spelling out the joint probability of given wages and health status. Denote the probability of having a wage rate y and health status k of individual i as $P_i(y,k) = \Pr ob(w_i = y,\ h_i = k)$, the log-likelihood function of a sample with N individuals can be written as,

(8) $L = \sum_{i=1}^{N}\sum_{k=0}^{4} I(h_i = k)\ln[P_i(y,k)]$,

where $I(h_i = k)$ is an indicator function, equal to 1 if $h_i = k$ and zero otherwise. $P_i(y,k)$ is defined in Appendix B.

4. Data and model specification

4.1 The data

The empirical analysis is based on the Household, Income and Labour Dynamics in Australia (HILDA) Survey, the first Australian longitudinal survey of its kind. A detailed documentation of the survey can be found in Watson and Wooden (2002). This paper uses the third wave HILDA survey because this wave collected more detailed information on health conditions than the other waves. The first wave of the survey was conducted between August and December 2001, where 7682 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 about one year apart.[13]

The HILDA survey contains detailed information on each individual’s labour market activities and history. The hourly wage rate used in this study is derived from the pre-tax total current weekly earnings and hours worked. The survey also collects information on earnings and working hours of the previous financial year. But since health status is measured as at current, current wages instead of wages of the previous financial year are used in the model.

[13] 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 (Melbourne Institute of Applied Economic and Social Research, 2005).

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. Information relating to individuals' health was collected in both personal interviews and self-completion questionnaires. In all waves of personal interviews, individuals were asked whether they had a long-term health condition, impairment or disability that restricted everyday activities and had lasted or was likely to last for six months or more. A special feature of the third wave survey was that in the interview questionnaire not only were individuals asked whether a long-term health condition was present, they were also asked what the health conditions were and when these conditions were first developed. Furthermore, in the self-completion questionnaire persons were asked whether they were told by a doctor or nurse that they had health conditions such as arthritis, asthma, etc. Other waves did not collect such detailed information on specific health conditions.

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: "*Would you describe your health as excellent, very good, good, fair, or poor?*". This self-reported health status is used as the observed counterpart of latent health capital as specified in equation (3) and the specific health conditions form the set of instruments employed in the health equation.

In addition to the endogeneity problem, researchers are often worried that self-reported health is subject to measurement error because individuals' awareness of health problems may vary with education, employment, income and health insurance status, and hence different individuals with identical underlying health status may respond differently to the same health question. In addition, discretized responses to the health question may also introduce measurement error because people may use different underlying cut-off points in assessing and reporting their health status. Measurement error leads to a downward bias to the effect of health if not properly accounted for. As mentioned earlier, the latent health model used in this paper handles the measurement error problem as well.

Despite self-reported health being a subjective measure of health, there is a large body of literature showing that this measure is a good indicator of health in the sense that it is highly correlated with medically determined health status (Nagi, 1969; Maddox and Douglas, 1973; LaRue et al., 1979; Ferraro, 1980). Tausman and Rosen (1982) find that self-reported health is close to “objective” health. Gerdtham et al. (1999) show that a continuous health status measure constructed from categorical self-reported health by the method of Wagstaff and van Doorslaer (1994) is highly correlated with other continuous measures of health, such as the rating scale measure and the time trade-off measure. These studies lend credence to the usefulness of self-reported health.

4.2 The sample

The analysis focuses on males aged between 25 and 64 years (inclusive). People aged over 64 years were excluded because the minimum Age Pension age for males is 65 years in Australia. Also excluded were persons aged under 25 years because many of them had not completed studies. Some of those aged 25-64 years were excluded if they were undertaking full-time study at the time of survey. In addition, I excluded those individuals with a missing value on the variables included in the model. First, to have a feeling of the relationship between health, employment and wages of these people, Table 1 presents employment status and wages by self-reported health status.

Table 1: Employment status and wages by self-reported health status

	Proportion of workers			Wages of workers		Log-wages of workers		
Health status	Mean	St.d.	Obs.	Mean	St.d.	Mean	St.d.	Obs.
Poor	0.145	0.354	117	28.424	16.736	3.220	0.505	17
Fair	0.513	0.500	423	22.018	12.083	2.985	0.452	217
Good	0.757	0.429	1144	23.045	14.406	3.009	0.504	866
Very good	0.800	0.400	1177	25.452	37.648	3.052	0.559	942
Excellent	0.773	0.419	366	25.770	15.487	3.122	0.482	283
All	0.720	0.449	3227	24.295	26.415	3.039	0.521	2325

A clear pattern between health status and employment status appears from the table: the better the health, the higher is the proportion being employed. For example, among those who reported poor health only about 15 percent were working, while among those reporting very good health about 80 percent were working. However, the relationship between wages and health status is not so clear-cut. While among those who reported at least fair health, the wage rate appears to increase with health, those who reported poor health had the highest wage rate. This seems to be at odds with expectation. But the simple tabulation has not controlled for other observed and unobserved factors that may also affect wages. Furthermore, the wage differentials between different health levels are not statistically significant.

4.3 Model specification

Table A1 in Appendix A presents definitions of the variables included in the model and their means. The grouping of the variables also shows detailed model specification, which is summarised as follows. The variables included in both the wage and health equations are marital status, education (four dummies), indigenous status, country of birth (three dummies) and residence of state (six dummies). Additional exogenous variables included in the wage equation but excluded from the health equation are work experience and its squared, capital city, casual job, part-time job, the interaction between casual and part-time job, union status and firm size (six dummies). Additional exogenous variables included in the health equation but excluded from the wage equation are age and its square, specific health conditions (25 dummies) and health risk behaviours (three dummies).

In the wage equation, in addition to the variables representing human capital such as education, health and work experience, I include state dummy and capital city variables to control for the wage differentials resulting from differences in labour market conditions and living costs across different areas. Marital status is included in the wage equation because it is often found married men earn higher wages than their unmarried counterparts (e.g. Gray, 1997; Loh, 1996; Breush and Gray, 2004). Indigenous status and immigration status are included in the wage equation to reflect their disadvantages in the labour market. The part-time and casual job variables reflect characteristics of jobs and thus may affect wages (Booth and Wood, 2004). Because union members are often found to attract a wage premium (Lewis, 1986; Steward, 1995; Hildreth, 1999), a union membership status variable is included in the wage equation. Finally, it has been found that larger firms pay higher wages than smaller ones (Brawn and Medoff, 1989; Main and Reilly, 1993). Firm size is thus included in the wage equation to control for this effect.

In the health equation, in addition to the set of specific health condition variables and the health risk behaviour variables, I include some demographic and socio-economic variables. Age is included because according to Grossman's health production theory (1972), people's health capital depreciates with age and the variable age squared is used to capture the non-linear depreciation rate. 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 found to be positively correlated with being married. Therefore, the marital status variable is included in the health equation as well. Education is often thought an important factor that affects health because education may improve individuals' health-related knowledge and thus the efficiency of health production (Grossman and Kaestner, 1997). State dummies are included to control for differences in health policy and services across states because hospitals and communities health services are managed by state governments.

By including different variables in the two equations, the exclusion restriction required to identify the simultaneous equation model is satisfied. But the question remains as to whether the exclusion restrictions imposed for each equation are valid. That is whether it is reasonable to assume that the identifying instruments (e.g. union) only affect one endogenous variable (e.g. wages) but not the other (e.g. health). The identifying instruments used in the health equation are variables on specific health conditions and health risk behaviour, which may not be subject to questions. This is because any occurrence of health conditions or diseases represents an adverse shock to the underlying health capital, if one believes Grossman's (1972) theory on health capital. In addition, in the literature on the effect of health on labour force status, it is a popular strategy to use specific health conditions or diseases to instrument potentially endogenous self-reported health (e.g. Stern, 1989; Bound, 1991; Bound et al., 1999; Campolieti, 2002; Disney et al., 2006). There are studies that examine the effects of health risk behaviour, such as alcohol consumption or smoking, on wages or earnings, but the theory used to argue for the effect is often that health risk behaviour affects health and thus wage.

However, one may doubt whether the identifying instruments employed by the wage equation, such as capital city, part-time job, casual job, firm size and union status, should be excluded from the health equation, although there are good reasons for them to be included in the wage equation.[14] For example, one may argue that living in a capital city is good to health because there is better access to health services. But on the other hand, higher stress and bad environmental conditions such as air pollution associated with large cities may be harmful to one's health. Therefore, the exact direction of the effect of capital city on health cannot be predicted theoretically. Similarly, there is no reason to believe that union status and firm size should affect health in a particular way. As for the part-time job variable, it can be argued that this variable should directly affect health because part-time workers work shorter hours, but Haveman et al. (1994) find that working hours have no direct effect on health. Despite these arguments for the exclusion from the health equation of the variables that are only included in the wage equation, to confirm the validity of these identification instruments, I conduct and report sensitivity analysis in the next section.

5. Empirical result

5.1 The wage equation

The estimation results for the wage equation are presented in Table 2.[15] Columns (1) and (4) show the results using simple OLS where self-reported health is treated as an exogenous variable. The difference between columns (1) and (4) is that in column (1) self-reported health is used as a continuous variable, while in column (4) it is used as categorical variables. Recall that two potential problems may bias the estimation results when self-reported health is treated as an exogenous variable: measurement error and endogeneity.[16] Nonetheless, these model results serve as a basis to be compared with models that tackle one or all of the biases.

[14] It is not for the purpose of satisfying exclusion restrictions that work experience and age are included in different equations, rather it is because these two variables are highly correlated and thus not appropriate to be included in the same equation. Work experience is included in the wage equation because it reflects human capital accumulation to a larger extent than does the age variable. One the other hand, health normally deteriorates with age and thus age is more appropriate to be included in the health equation.

[15] The results of the first stage estimation are reported in Table A2 in Appendix A.

[16] Another potential problem is sample selection bias, but as shown in Appendix C, sample selection bias is not an issue in the data used.

Table 2: Estimation results for the wage equation

Variables	Health as a continuous variable						Health as categorical variables[b]			
	Simple OLS[a] (1)		Two stage OLS (2)		FIML (3)		Simple OLS[a] (4)		Two stage OLS (5)	
	Coef.	S.E.	Coef.	S.E.[d]	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.[d]
Health[c]	0.0093	0.0117	0.0675***	0.0257	0.0741***	0.0211				
Good health[c]							-0.0241	0.0348	0.1280	0.0845
Very good/excellent health[c]							-0.0003	0.0339	0.1787**	0.0854
Married	0.0958***	0.0239	0.0910***	0.0226	0.0903***	0.0275	0.0966***	0.0240	0.0921***	0.0215
Degree	0.3640***	0.0296	0.3364***	0.0312	0.3340***	0.0341	0.3646***	0.0295	0.3469***	0.0306
Other post_sch qual	0.1130***	0.0256	0.1043***	0.0250	0.1035***	0.0289	0.1136***	0.0256	0.1038***	0.0253
Year 12	0.0951***	0.0368	0.0789**	0.0397	0.0778**	0.0389	0.0957***	0.0367	0.0811**	0.0392
Experience /10	0.1478***	0.0408	0.1594***	0.0407	0.1584***	0.0452	0.1458***	0.0408	0.1545***	0.0407
Experience squared /100	-0.0218***	0.0080	-0.0229***	0.0084	-0.0224***	0.0084	-0.0214***	0.0080	-0.0222***	0.0084
Indigenous	-0.0543	0.1036	-0.0697	0.0642	-0.0720	0.1885	-0.0552	0.1036	-0.0622	0.0643
Born OS En	0.0197	0.0302	0.0153	0.0323	0.0145	0.0308	0.0204	0.0302	0.0166	0.0324
Born non-En	-0.0755**	0.0329	-0.0810**	0.0336	-0.0817**	0.0348	-0.0757**	0.0329	-0.0782**	0.0349
VIC	-0.0551**	0.0261	-0.0586**	0.0258	-0.0589**	0.0283	-0.0547**	0.0261	-0.0544**	0.0244
QLD	-0.0504*	0.0278	-0.0434	0.0280	-0.0430	0.0303	-0.0499*	0.0278	-0.0434	0.0268
SA	-0.1460***	0.0367	-0.1470***	0.0385	-0.1477***	0.0370	-0.1454***	0.0367	-0.1459***	0.0377
WA/NT	-0.0210	0.0339	-0.0149	0.0359	-0.0150	0.0343	-0.0218	0.0339	-0.0150	0.0368
TAS	-0.0636	0.0671	-0.0711	0.0509	-0.0706	0.0894	-0.0618	0.0671	-0.0694	0.0519
Capital city	0.1321***	0.0216	0.1297***	0.0220	0.1306***	0.0221	0.1323***	0.0216	0.1296***	0.0216
Part-time	0.2379***	0.0525	0.2568***	0.0937	0.2567***	0.0355	0.2342***	0.0527	0.2507***	0.0911
Casual	0.0393	0.0369	0.0402	0.0397	0.0482	0.0358	0.0383	0.0370	0.0386	0.0389
Part-time & casual	-0.1955**	0.0793	-0.1902*	0.1197	-0.2118***	0.0618	-0.1926**	0.0794	-0.1855*	0.1126
Firm size 20-99	0.1742***	0.0245	0.1682***	0.0250	0.1725***	0.0274	0.1742***	0.0245	0.1706***	0.0241
Firm size 100-199	0.2049***	0.0347	0.2077***	0.0327	0.2002***	0.0377	0.2048***	0.0347	0.2073***	0.0324
Firm size 200-499	0.2535***	0.0349	0.2530***	0.0296	0.2453***	0.0437	0.2544***	0.0349	0.2530***	0.0288
Firm size 500+	0.3265***	0.0345	0.3210***	0.0344	0.3222***	0.0380	0.3268***	0.0345	0.3251***	0.0322
Firm size unknown	0.0214	0.1266	0.0161	0.1096	0.0113	0.1598	0.0198	0.1266	0.0164	0.1056
Union member	0.0750***	0.0218	0.0765***	0.0205	0.0768***	0.0241	0.0756***	0.0219	0.0768***	0.0215
Constant	2.3857***	0.0642	2.4263***	0.0528	2.4120***	0.0693	2.4191***	0.0640	2.2491***	0.1003
Adjusted R^2	0.1989		0.2020				0.1988		0.2019	
Log-likelihood						-4178.77				
No. obs.	2325		2325		2325		2325		2325	

*** Significant at 1%, ** 5% and * 10%. Note: (a) Simple OLS estimation treats self-reported health as an exogenous variable. (b) Due to insufficient observations, the five health levels are reduced to three: poor and fair are combined into one category, very good and excellent into one group. (c) For the two-stage OLS health refers to predicted health from the first stage estimation; for FIML method it refers to unobserved latent health. (d) The second-stage standard errors are obtained using the bootstrap method with 1000 replications.

Results in columns (2) and (5) also use the OLS but with predicted health from the first stage estimation replacing observed self-reported health status. Again column (2) uses predicted health as a continuous variable, while column (5) uses predicted health status or categories. As discussed earlier, the predicted health measures purge of the measure error and endogeneity problems. The model in column (3) estimates the wage and health equations simultaneously using the FIML method discussed earlier.

Because the effect of health on wages is the main interest of this study, we first look at the estimate for the health variable. The result in column (1) indicates that better health is indeed associated with a higher wage, but the estimate is very small in magnitude and statistically insignificant in any conventional significance levels. The results in column (4) show that people with good, very good or excellent health have a lower wage than people with poor or fair health, but again, these estimates are insignificant. On the other hand, the estimates for the health variable in columns (2), (3) and (5), where measurement error and endogeneity have been accounted for, are all significant and have expected signs.[17] Literally the estimate on latent health shows that everything else equal, a one unit increase in latent health raises the wage rate by 7 percent, which is about 7 times of the estimate from the simple OLS model in column (1). The estimate in column (5) indicates that compared to persons with poor or fair health, people with very good or excellent health can earn a wage 18 percent higher.

In contrast to the difference in the estimates for the health variable between OLS and other models, the estimates for all other variables are very similar across different models, and all the significant variables have expected signs. For example, other things being equal, men with a partner are found to earn a wage premium of about 9 percent compared with unmarried men. The estimates for the three education dummy variables show that the higher the education level, the higher is the wage rate. Those with a bachelor degree or higher earn a wage rate over 30 percent higher than those who do not complete schooling and those with other post-school qualifications are expected to have a wage rate about 10 percent higher. All model results show that work experience is awarded but at a decreasing rate, again consistent with other studies in the literature. Compared with native born Australian, immigrants from English speaking countries are not found to make a difference. On the other hand,

[17] In column (5) the two estimates for good and very good/excellent health dummies are jointly significant at the 5 percent level.

immigrants from non-English countries earn a wage rate 8 percent lower than men born in Australian. The lower wage rate of immigrants from non-English speaking countries perhaps reflects their language difficulties or discrimination. As expected those who work in capital cities are found to have a higher wage rate, perhaps to compensate for higher living costs. Employees who work part-time or on a casual arrangement are estimated to have a higher wage rate, but only for the part-time worker variable the estimate is significant. The result that part-time workers earn a wage premium seems contrary to the findings from other countries (Hirsch, 2004), but a similar result is found in Booth and Wood (2004), which use the HILDA survey waves 1 and 2. The result also indicates that casual part-time workers have a lower wage rate than other casual workers. The estimates for the state dummies show that, other things being equal, workers in other states earn lower wages than those in NSW and ACT, but only the estimates for VIC and SA are significant. Employees in large firms and trade union members are also found to earn wage premiums.

5.2 Is health exogenous?

Although the focus of this study is on the effect of health on wages, the exogeneity of the health variable is also of interest. As discussed earlier, the endogeneity of health may occur due to either the reverse effect of wages on health or the correlation of unobserved factors. The first row in Table 3 presents the estimates for the reverse effect of wages on health from different models of health determination. The full estimation results for the health equation are not discussed here because they are not the main interest of the study, but they are reported in Table A3 in Appendix A. The model in column (1) is a simple ordered probit treating observed wages as exogenous; the model in column (2) is an ordered probit as well, but predicted wages instead of observed wages is used to account for endogeneity; the model in column (3) estimates the wage and health equations simultaneously, using the FIML method. From all models the wage variable is found to be insignificant, suggesting that the reverse effect from wages to health cannot be established from the data.

Table 3: Estimates for the wage variable and the covariance of the two error terms

	Simple ordered probit (1)		Two stage probit (2)		FIML (3)	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
(Log-) wages	-0.0102	0.0465	0.0832	0.1509	0.0936	0.1585
$Cov(\varepsilon_w, \varepsilon_h)$					-0.0995**	0.0404

Only from the FIML method can the correlation of the two error terms be assessed. The last row in Table 3 shows the estimate for the covariance between the two error terms. The covariance is significant at the 5 percent significance level. As discussed earlier, a true test on the exogeneity of health must examine both the estimates for the wage variable and the correlation of the two error terms. That is the null hypothesis on the exogeneity of health is,

$$H_0: \ \alpha_h = 0, \ \ and \ \ \delta_{w,h} = 0;$$
$$H_1: \ \alpha_h \neq 0, \ \ or \ \ \delta_{w,h} \neq 0.$$

Where $\delta_{w,h}$ refers to the covariance between ε_w and ε_h. The Wald-test statistics for testing the joint significance of the two parameters is 12.25, which is significant at the 1 percent significance level. Therefore, the exogeneity of health cannot be accepted at any conventional significance levels, even though the reverse effect of wages on health is insignificant.

The negative correlation between the unobserved determinants of health and wages suggests that the unobserved factors have opposite effects on health and wages. Consequently, if the negative correlation between the two equations were not controlled for, the direct effect of health on wages would be underestimated. However, it is not obvious what these unobserved factors are and one can only speculate. One possible example of such factors is a person's inherent attitude toward leisure, work and health. For instance, there are people who care very much about their health and thus do not put much effort into their work. For this type of persons, the unobserved variables have a positive effect on health but a negative effect on productivity and wages. Another example may be individuals' motivation. Highly motivated people (e.g. workaholic) work harder than others and earn higher wages, but they may also be more likely to put high pressure on themselves, which is detrimental to their health.

The significant and negative correlation between the two error terms and the insignificance of the reverse effect of wages on health suggest that the much smaller and insignificant estimate for the health variable when treating health as exogenous is driven by both measurement error of self-reported health and the negative correlation of the unobserved determinants of health and wages.

5.3 Sensitivity analysis

The results so far depend on the exclusion restrictions imposed for each equation to identify the model. The excluded variables used to identify the wage equation are specific health conditions and health risk behaviour and thus should not be subject to controversy as discussed earlier. However, the variables excluded from the health equation, i.e. capital city, part-time job, casual job, the interaction between part-time and casual job, firm size and union status, may cause concerns because some theoretical arguments can be made for why the exclusion restrictions may fail, as mentioned earlier. The consequence of wrongly excluding the variables from the health equation is that the estimate on the wage variable might be biased and the inference on the endogeneity of health might be wrongly based as well. The direction of the bias depends on the sign of the correlation between wages and those excluded variables from the health equation.

Despite these concerns, the identifying instruments employed in the wage equation are not significant in the first stage estimation of the health equation, indicating that they have no directly effect on health. To see whether these variables have a significant effect in the second stage health equation and how the results of the endogenous variables change when the exclusion restrictions imposed for the health equation vary, I experimented by adding some of the excluded variables to the health equation.[18] The results for the two endogenous variables from these experiments and the statistics for testing the significance of the additional variables added to the health

[18] The variables included only in the wage equation are essentially instrumental variables for the endogenous wage variable. In an instrumental variable framework, one may also ask whether these instruments are valid, particularly in terms of independence to the error terms of the health equation, although these variables are assumed exogenous to the system at the outset (see Section 3). Overidentifying restriction tests are often used to verify the validity of instrumental variables for linear models, but the health equation here is a nonlinear model. In addition, it is the latent value of health that enters the wage equation. Such models it is difficult to test the validity of the instrumental variables (Davidson and Mackinnon, 1993, p667). Nevertheless, I conducted overidentifying restriction tests for the variables capital city, part-time job, casual job, part-time & casual job, firm size and union status, using the ivreg2 command in STATA 9, by treating the discrete health variable as a continuous variable. The test results indicate that all these variables are valid instruments.

equation are reported in Table 4.[19] The results in columns (1), (2) and (3) in Table 4 were obtained when the capital city variable, the three job characteristic variables (i.e. part-time job, casual job, part-time & casual job), and the five firm size dummies variables were added to the health equation, respectively; the results in column (4) were obtained from adding all the above three sets of variables to the health equation. The results in column (5) were obtained when the union status variable was added to the health equation (but with the capital city, job characteristic and firm size variables being removed from the health equation). The FIML method was used in conducting these experiments. Comparing the results in Table 4 with those reported in Tables 2 and 3 (column (3)), it is clear that varying the set of identifying instruments in the wage equation does not changes the estimates for the two endogenous variables.

The test statistics in the last two rows of the table show that none of these additional variables is significant in the health equation, implying that these variables have no direct effect on health and thus can be excluded from the health equation. Therefore, the simultaneous equation models results reported in Tables 2 and 3 are reliable.

Table 4: Estimates for the endogenous variables when the set of identification instruments in the wage equation varies

	Adding the following variables to the health equation									
	Capital city (1)		Job characteristics[(a)] (2)		Firm size[(b)] (3)		(1)+(2)+(3) (4)		Union status (5)	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Health	0.0742***	0.0211	0.0739***	0.0212	0.0744***	0.0211	0.0746***	0.0212	0.0742***	0.0211
Wage	0.0918	0.1736	0.0941	0.1635	0.0920	0.2765	0.0942	0.5756	0.0939	0.0175
χ^2	0.0193		2.9294		3.7347		6.9642		0.0024	
df	1		3		5		9		1	

Note: (a) Refers to the variables part-time job, casual job, and part-time & casual job. (b) Refers to the 5 firm size dummy variables.

6. Conclusion

Using the third wave HILDA Survey data, which contains more detained information on health conditions than the other waves, this paper estimated a simultaneous equation model to examine the effect of health on the wages of working-age Australian men.

[19] The results for other variables can be obtained on request. They are very similar to those reported in the paper.

The results confirm the finding in the overseas literature that health has a significant effect on wages, but the significant effect is found only when the endogeneity and measurement error of health are accounted for. The estimated effect of health is substantial in magnitude. For example, results from the model using predicted health categories indicates that people with very good health can earn a wage 18 percent higher compared to those with poor or fair health. The reverse effect of wages on health is not found for the group of people analyzed and the endogeneity of health is found to arise mainly from unobserved factors that affect both wages and health but in the opposite direction.

It should be acknowledged that wages are only one aspect of individuals' economic wellbeing that can be affected by health. The effects on total economic wellbeing would be larger because a person's labour force participation and working hours may also be affected by health and economic wellbeing is derived from the combination of wages and labour supply. Therefore, examining the effect of health on working hours is naturally the next step for future research. In addition, for future research the panel nature of the HILDA data could be explored to better control for unobserved heterogeneity. Another dimension of future research is to examine the effects of health on the variance of wages or earnings, since a larger variance of wages is associated with greater uncertainty that has adverse effects on individuals' wellbeing (Currie and Madrian, 1999).

Appendix A: Variable definitions and additional estimation results

Table A1: Definition and mean value of variables

Variable name	Definition of variable	Workers	Non-workers	All
Variables in both wage and health equations				
Married	1 if married or de facto	0.7763	0.7029	0.7558
Degree	1 if has a bachelor or higher degree	0.2641	0.1419	0.2299
Other post_sch qual	1 if has other non-degree post-school qualification	0.4009	0.4069	0.4025
Year 12	1 if completed year 12	0.1062	0.0843	0.1001
Year 11/lower - *omitted*	1 if highest education completed is lower than year 12			
Indigenous	1 if indigenous or Torres Strait Islander	0.0090	0.0133	0.0102
Australian born - *omitted*	1 if born in Australia			
Born OS En	1 if born overseas English speaking country	0.1252	0.1175	0.1230
Born non-En	1 if born in non-English speaking country	0.1028	0.1220	0.1082
NSW/ACT -*omitted*	1 if lives in NSW or ACT			
VIC	1 if lives in VIC	0.2538	0.2384	0.2495
QLD	1 if lives in QLD	0.2047	0.2162	0.2079
SA	1 if lives in SA	0.0916	0.1086	0.0964
WA/NT	1 if lives in WA or NT	0.1123	0.1153	0.1131
TAS	1 if lives in TAS	0.0232	0.0477	0.0301
Variables in wage equation but not in health equation				
Experience	years in employment since first leaving full-time education	23.55	28.04	24.80
St.d.		10.60	11.95	11.17
Experience squared	Square of experience	666.80	928.57	739.97
Capital city[(a)]	1 if living in a capital city	0.6103	0.4645	0.5696
Part-time[(b)]	1 if a part-time worker	0.0761		0.0548
Casual[(b)]	1 if a casual employee	0.1178		0.0849
Part-time & casual[(b)]	1 if a part-time and casual worker	0.0391		0.0282
Firm size<20[(b)] - *omitted*	1 if number of employees < 20	0.4146		0.2987
Firm size 20-99[(b)]	1 if number of employees 20-99	0.2680		0.1931
Firm size 100-199[(b)]	1 if number of employees 100-199	0.1015		0.0731
Firm size 200-499[(b)]	1 if number of employees 200-499	0.1006		0.0725
Firm size 500+[(b)]	1 if number of employees 500 or more	0.1092		0.0787
Firm size unknown[(b)]	1 if number of employees unknown but >20	0.0060		0.0043
Union member[(b)]	1 if a trade union member	0.3196		0.2302
Variables in health equation but not in wage equation				
Age	Age in 2003	42.17	49.23	44.14
St.d.		9.82	10.81	10.59
Age squared	Square of age	1874.57	2539.97	2060.56
Sight	1 if has sight problems not corrected by glasses/lenses	0.0125	0.0377	0.0195
Hearing	1 if has hearing problems	0.0344	0.0976	0.0521
Speech	1 if has speech problems	0.0026	0.0055	0.0034
Blackout	1 if has blackout, fits or loss of consciousness problems	0.0013	0.0188	0.0062
Learning	1 if slow at learning or understanding things	0.0034	0.0155	0.0068
Arm_finger	1 if limited use of arms or fingers	0.0112	0.0643	0.0260
Gripping	1 if has difficulty gripping things	0.0047	0.0610	0.0205
Feet_leg	1 if limited use of feet or legs	0.0181	0.0876	0.0375

Nerve_emotion	1 if has nervous or emotional condition requiring treatment	0.0082	0.0776	0.0276
Physical	1 if has conditions restricting physical activity or physical work	0.0473	0.1962	0.0889
Deformity	1 if has any disfiguration or deformity	0.0034	0.0166	0.0071
Mental	1 if has any mental illness requiring help or supervision	0.0039	0.0344	0.0124
Breath	1 if has shortness of breath or difficulty breathing	0.0086	0.0632	0.0239
Chronic pain	1 if has chronic or recurring pain	0.0189	0.1208	0.0474
LT damage	1 if long term effects as a result of a head injury, stroke or other brain damage	0.0026	0.0322	0.0108
LT ailment	1 if a long-term condition or ailment which is still restrictive even though it is being treated or medication being taken for it	0.0305	0.1552	0.0654
Arthritis	1 if has arthritis	0.0963	0.2561	0.1410
Asthma	1 if has asthma	0.0968	0.0865	0.0939
Cancer	1 if has any type of cancer	0.0280	0.0599	0.0369
Bronchitis	1 if has bronchitis or emphysema	0.0125	0.0388	0.0198
Diabetes	1 if has diabetes	0.0305	0.0865	0.0462
Coronary	1 if has heart or coronary disease	0.0206	0.0798	0.0372
Hypertension	1 if has high blood pressure or hypertension	0.1153	0.2317	0.1478
Circulatory	1 if has other circulatory condition	0.0103	0.0565	0.0232
Other condition	1 if has any other conditions	0.0034	0.0044	0.0037
Smoker	1 if currently smoking or ever smoked	0.6116	0.7162	0.6408
Heavy drinker	1 if a heavy drinker, defined as drinking more than 6 standard drinks a day when drinking	0.0929	0.1142	0.0989
Lack physical activity	1 if lack of physical activity, defined as no physical activity at all or less than one per week	0.0778	0.1619	0.1013
Variables in employment status (selection) equation only				
Child 0-4	1 if has child(ren) aged 0 to 4	0.1819	0.1064	0.1608
Child 5-14	1 if has child(ren) aged 5 to 14	0.2903	0.1951	0.2637
Aged 55p	1 if aged 55 years or over	0.1075	0.3703	0.1810
No. of obs.		2325	902	3227

Note (a) The capital cities do not include Hobart and Darwin because they cannot be identified from the data. (b) These variables are only defined for workers and thus not included in the employment status equation.

Table A2: The first-stage estimation results

	Wage equation		Health equation	
Variables	Coef.	S.E.	Coef.	S.E.
Married	0.0902***	0.0241	0.0355	0.0563
Degree	0.3811***	0.0321	0.3531***	0.0752
Other post_sch qual	0.1078***	0.0258	0.0892	0.0601
Year 12	0.0948***	0.0370	0.1717**	0.0863
Experience	0.1321*	0.0794	0.0029	0.1859
Experience squared	-0.0002	0.0150	0.0084	0.0351
Indigenous	-0.0607	0.1042	0.2477	0.2445
Born OS En country	0.0262	0.0304	0.0851	0.0710
Born non-En country	-0.0578*	0.0338	0.1499*	0.0791
VIC	-0.0552**	0.0261	0.0294	0.0609
QLD	-0.0506*	0.0278	-0.1215*	0.0649
SA	-0.1418***	0.0368	0.0072	0.0859
WA/NT	-0.0233	0.0339	-0.1104	0.0792
TAS	-0.0657	0.0670	0.0687	0.1572
Capital city	0.1358***	0.0217	0.0232	0.0507
Part-time	0.2688***	0.0537	0.0029	0.1258
Casual	0.0539	0.0373	0.0926	0.0870
Part-time & casual	-0.1872**	0.0797	-0.2513	0.1862
Firm size 20-99	0.1716***	0.0246	0.0664	0.0573
Firm size 100-199	0.2016***	0.0347	-0.0642	0.0809
Firm size 200-499	0.2437***	0.0349	-0.0555	0.0813
Firm size 500+	0.3265***	0.0345	0.0483	0.0808
Firm size unknown	-0.0108	0.1272	-0.0209	0.2967
Union member	0.0734***	0.0220	0.0102	0.0513
Age	0.0240	0.0739	-0.1753	0.1729
Age squared	-0.0330*	0.0186	0.0286	0.0434
Sight	0.0540	0.0880	-0.3216	0.2047
Hearing	-0.0544	0.0550	-0.1855	0.1284
Speech	-0.1261	0.1981	-0.5799	0.4643
Blackout	0.1587	0.2797	-1.4851**	0.6546
Learning	-0.6420***	0.1885	-0.4694	0.4415
Arm_finger	0.0694	0.1121	-0.4328*	0.2608
Gripping	0.1400	0.1699	0.5265	0.3957
Feet_leg	-0.0109	0.0788	-0.6868***	0.1854
Nerve_emotion	-0.0912	0.1176	-0.3914	0.2748
Physical	-0.0007	0.0494	-0.6562***	0.1158
Deformity	-0.0037	0.1701	-0.1935	0.3986
Mental	0.0128	0.1668	-0.9180**	0.3905
Breath	0.1389	0.1087	-0.7140***	0.2563
Chronic pain	-0.1088	0.0784	-0.1903	0.1831
LT damage	-0.1572	0.2248	0.8881*	0.5256
LT ailment	-0.0035	0.0609	-0.7171***	0.1434
Arthritis	-0.0475	0.0356	-0.3396***	0.0831
Asthma	-0.0005	0.0337	-0.0823	0.0786
Cancer	-0.0653	0.0596	-0.0160	0.1389
Bronchitis	-0.0236	0.0905	0.0068	0.2113
Diabetes	-0.0654	0.0587	-0.6028***	0.1374
Coronary	0.0857	0.0714	-0.3829**	0.1666
Hypertension	0.0156	0.0322	-0.2834***	0.0750
Circulatory	-0.2536**	0.1000	-0.5738**	0.2342
Other condition	0.2048	0.1662	-0.2221	0.3874
Smoker	-0.0137	0.0206	-0.1631***	0.0483
Heavy drinker	0.0048	0.0347	-0.1481*	0.0809
Lack physical activity	-0.0522	0.0368	-0.4930***	0.0862
Constant	2.4097***	0.0655		
Cutoff_1			-3.2093	0.1919
Cutoff_2			-1.7156	0.1571
Cutoff_3			-0.3283	0.1540
Cutoff_4			1.0153	0.1551
Adjusted R^2	0.2087			
Log-likelihood			-2666.29	
Psudo-R^2			0.0808	
No. obs.	2325		2325	

*** Significant at 1% level, ** 5% level and * 10% level.

Table A3: Estimation results for the health equation of workers

	Simple ordered probit (1)		Two stage probit (2)		FIML (3)	
Variables	Coef.	S.E.	Coef.	S.E.[b]	Coef.	S.E.
Log-wage[a]	-0.0102	0.0465	0.0832	0.1509	0.0936	0.1585
Age (-25)/10	-0.1555*	0.0866	-0.1717*	0.0959	-0.1566*	0.0871
Age squared/100	0.0347	0.0230	0.0380	0.0251	0.0326	0.0228
Married	0.0382	0.0559	0.0310	0.0624	0.0304	0.0561
Degree	0.3455***	0.0702	0.3068***	0.0927	0.2993***	0.0973
Other post_sch qual	0.0884	0.0600	0.0772	0.0630	0.0737	0.0658
Year 12	0.1653*	0.0856	0.1542*	0.0904	0.1522*	0.0894
Indigenous	0.2585	0.2430	0.2548	0.2493	0.2536	0.2555
Born OS En	0.0886	0.0702	0.0849	0.0719	0.0853	0.0717
Born non-En	0.1486*	0.0773	0.1521*	0.0793	0.1546**	0.0777
VIC	0.0291	0.0603	0.0328	0.0633	0.0347	0.0619
QLD	-0.1187*	0.0647	-0.1124*	0.0666	-0.1083	0.0701
SA	0.0101	0.0856	0.0219	0.0988	0.0278	0.0866
WA/NT	-0.1010	0.0788	-0.0998	0.0820	-0.0967	0.0784
TAS	0.0476	0.1549	0.0599	0.1760	0.0597	0.1484
Sight	-0.3178	0.2042	-0.3188	0.2453	-0.2816	0.1795
Hearing	-0.1838	0.1280	-0.1802	0.1447	-0.2004	0.1271
Speech	-0.5429	0.4631	-0.5259	0.3435	-0.5743	0.8779
Blackout	-1.4066**	0.6533	-1.4369***	0.5440	-1.3323	1.5017
Learning	-0.4827	0.4411	-0.4210*	0.2476	-0.7124	1.1829
Arm_finger	-0.4300*	0.2594	-0.4337	0.3806	-0.3816*	0.1994
Gripping	0.5077	0.3928	0.4962	0.4595	0.5332	0.4767
Feet_leg	-0.7072***	0.1847	-0.7042***	0.1844	-0.6865***	0.2048
Nerve_emotion	-0.3927	0.2743	-0.3883	0.3037	-0.4077	0.2723
Physical	-0.6705***	0.1150	-0.6682***	0.1231	-0.6430***	0.1139
Deformity	-0.1920	0.3974	-0.1869	0.4962	-0.1768	0.4007
Mental	-0.9878***	0.3850	-0.9826**	0.4574	-0.9409**	0.4064
Breath	-0.7317***	0.2548	-0.7428**	0.3495	-0.6516***	0.2183
Chronic pain	-0.1905	0.1825	-0.1835	0.2014	-0.2311	0.1841
LT damage	0.8505	0.5245	0.8646**	0.3744	0.7589	1.7671
LT ailment	-0.7311***	0.1421	-0.7306***	0.1718	-0.7084***	0.1310
Arthritis	-0.3336***	0.0828	-0.3302***	0.0889	-0.3391***	0.0802
Asthma	-0.0783	0.0783	-0.0783	0.0834	-0.0773	0.0768
Cancer	-0.0239	0.1384	-0.0180	0.1569	-0.0507	0.1264
Bronchitis	-0.0075	0.2102	-0.0018	0.2625	-0.0087	0.1776
Diabetes	-0.5958***	0.1371	-0.5918***	0.1350	-0.6000***	0.1436
Coronary	-0.3811**	0.1660	-0.3906**	0.1774	-0.3293*	0.1721
Hypertension	-0.2820***	0.0746	-0.2823***	0.0772	-0.2617***	0.0771
Circulatory	-0.5513**	0.2340	-0.5286**	0.2237	-0.6356**	0.2831
Other condition	-0.1747	0.3853	-0.1918	0.4574	-0.0856	0.3812
Smoker	-0.1675***	0.0481	-0.1654***	0.0478	-0.1658***	0.0481
Heavy drinker	-0.1450*	0.0802	-0.1458*	0.0863	-0.1376*	0.0799
Lack physical activity	-0.4864***	0.0861	-0.4806***	0.0978	-0.4865***	0.0806
Cutoff_1	-3.2737	0.2006	-3.0197	0.4483	-2.9754	0.4721
Cutoff_2	-1.7789	0.1667	-1.5255	0.4274	-1.4859	0.4511
Cutoff_3	-0.3953	0.1634	-0.1418	0.4264	-0.1043	0.4459
Cutoff_4	0.9458	0.1645	1.1995	0.4275	1.2347	0.4405
Log-likelihood	-2670.21		-2670.08		-4178.77	
Psudo-R^2	0.0794		0.0795			
No. obs.	2325		2325		2325	

Note: (a) In the two-stage model wages refers to predicted wages from the first stage estimation.
(b) The second-stage standard errors are obtained using the bootstrap method with 1000 replications.

Appendix B: Probability definition

Suppose that ε_w and ε_h follow a joint normal distribution with the covariance matrix given by $\text{cov}(\varepsilon_w, \varepsilon_h) = \begin{pmatrix} \delta_w^2 & \delta_{w,h} \\ \delta_{w,h} & \delta_h^2 \end{pmatrix}$, then the covariance matrix of the error terms in the reduced form equations (1) and (2) becomes,

$$\text{(a1)}\ \text{cov}(\varepsilon_w^*, \varepsilon_h^*) = \begin{pmatrix} 1 & -\alpha_w \\ -\alpha_h & 1 \end{pmatrix}^{-1} \begin{pmatrix} \delta_w^2 & \delta_{w,h} \\ \delta_{w,h} & \delta_h^2 \end{pmatrix} \left[\begin{pmatrix} 1 & -\alpha_w \\ -\alpha_h & 1 \end{pmatrix}^{-1} \right]^T \equiv \begin{pmatrix} \delta_w^{*2} & \delta_{w,h}^* \\ \delta_{w,h}^* & \delta_h^{*2} \end{pmatrix}.$$

With equation (a1) and the normal distribution assumption on ε_w and ε_h, the probability of observing an individual *i* with a wage rate *y* and health status *k* can be written as,

$$\text{(a2)}\quad \Pr(w = y, h = k) = \left[\frac{1}{\delta_w^*} \phi\left(\frac{y - \Delta_w}{\delta_w^*}\right) \right] \times \left[\Phi\left(\frac{m_k - \Delta_h - \rho_{w,h}^* \dfrac{y - \Delta_w}{\delta_w^*}}{(1 - \rho_{w,h}^{*2})^{1/2}} \right) - \Phi\left(\frac{m_{k-1} - \Delta_h - \rho_{w,h}^* \dfrac{y - \Delta_w}{\delta_w^*}}{(1 - \rho_{w,h}^{*2})^{1/2}} \right) \right],$$

where $\Delta_w = \dfrac{1}{1 - \alpha_w \alpha_h} [x_w' \beta_w + x_h' \beta_h \times \alpha_w]$, $\Delta_h = \dfrac{1}{1 - \alpha_w \alpha_h} [x_w' \beta_w \times \alpha_h + x_h' \beta_h]$, and $\rho_{w,h}^* = \dfrac{\delta_{w,h}^*}{\delta_w^* \delta_h^*}$. $\phi(.)$ refers to the standard normal density function and $\Phi(.)$ the standard normal probability function.

The expression in the first bracket on the right-hand side of equation (a2) represents the probability of wages equal to y, while the expression in the second bracket refers to the probability of health status equal to k, conditioning on wages equal to *y*. The likelihood function of the sample can be formed using equation (a2) with *y* and *k* replaced by corresponding observed values of each individual. Because observed health status is a discrete variable, δ_h^2 has to be normalized to one for identification purposes.

Appendix C: Incorporating sample selection

C.1. The model and estimation strategy

To account for the fact that only individuals with positive wages can be analyzed using the simultaneous equation model specified in equations (1), (2) and (3), I augment the simultaneous equation model with a selection equation that postulates the determination of employment status,

(a3) $L^* = x_L^{'}\beta_L + \varepsilon_L$.

Where L^* is the latent value of being employed and x_L a vector of variables determining the value of employment. The error term ε_L is assumed to be correlated with both ε_w and ε_h . That is, unmeasured factors that affect wages, health and employment status are allowed to be freely correlated in the full system. To account for the possibility that potential wages and health status may also affect employment status, I include in x_L all variables that appear in both the health and wage equations except for those variables in the wage equation that describe job characteristics and thus are only defined for workers. In addition, to identify the selection equation, x_L includes some variables that are not included in the other two equations (see Table A1 in Appendix A). Again, the latent value of employment is not observable and the following equation relates unobserved latent value of employment to observed employment status,

(a4) $L = \begin{cases} 1 \quad (= employed) & if \quad L^* > 0 \\ 0 \quad (= not \quad employed) & if \quad L^* \le 0 \end{cases}$.

The FIML method is used to estimate the three equations simultaneously. With FIML another piece of information in the data can also be used to obtain more efficient results, i.e., health status of non-workers. For non-workers a similar health determination equation to equation (2) can be specified,

(a5) $h_{nw}^* = x_{h,nw}^{'}\beta_{h,nw} + \varepsilon_{h,nw}$,

where $x_{h,nw}$ includes the same set of variables as in x_h . The only difference between equations (2) and (a5) is that in the latter the wage rate is not included as a right-hand side variable because wages are not observed for non-workers.[20] To implement FIML, I assume all the error terms to follow a joint normal distribution with the covariance matrix,

$$\text{(a6)}\quad \operatorname{cov}(\varepsilon_w,\varepsilon_h,\varepsilon_L,\varepsilon_{h,nw}) = \begin{pmatrix} \delta_w^2 & \delta_{w,h} & \delta_{w,L} & . \\ \delta_{w,h} & \delta_h^2 & \delta_{h,L} & . \\ \delta_{w,L} & \delta_{h,L} & \delta_L^2 & \delta_{L,h_{nw}} \\ . & . & \delta_{L,h_{nw}} & \delta_{h,nw}^2 \end{pmatrix}.$$

Given the discrete nature of observed health status and employment status, δ_h , δ_L and $\delta_{h,nw}$ have to be normalized to ones and $\delta_{h,L}$ and $\delta_{h_{nw},L}$ consequently become correlation coefficients. From equation (a6),

$$\text{(a7)}\quad \begin{aligned} \operatorname{cov}(\varepsilon_w^*,\varepsilon_h^*,\varepsilon_L^*) &= \begin{pmatrix} 1 & -\alpha_w & 0 \\ -\alpha_h & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \operatorname{cov}(\varepsilon_w,\varepsilon_h,\varepsilon_L) \left[\begin{pmatrix} 1 & -\alpha_w & 0 \\ -\alpha_h & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \right]^T \\ &\equiv \begin{pmatrix} \delta_w^{*2} & \delta_{w,h}^* & \delta_{w,L}^* \\ \delta_{w,h}^* & \delta_h^{*2} & \delta_{h,L}^* \\ \delta_{w,L}^* & \delta_{h,L}^* & \delta_L^{*2} \end{pmatrix}. \end{aligned}$$

For a worker with a wage rate *y* and health status *k*, the contribution to the likelihood function can be expressed as,

$$\text{(a8)}\quad \begin{aligned} \Pr(w=y,h=k,L=1) &= \frac{1}{\delta_w^*}\phi\left(\frac{y-\Delta_w}{\delta_w^*}\right) \\ &\times\left[\Phi_2\left(\frac{m_k-\Delta_h-\rho_{w,h}^*\dfrac{y-\Delta_w}{\delta_w^*}}{(1-\rho_{w,h}^{*2})^{1/2}}, \frac{\Delta_L+\rho_{w,L}^*\dfrac{y-\Delta_w}{\delta_w^*}}{(1-\rho_{w,L}^{*2})^{1/2}}, -\tilde{\rho}_{h,L}^*\right)\right. \\ &\left.-\Phi_2\left(\frac{m_{k-1}-\Delta_h-\rho_{w,h}^*\dfrac{y-\Delta_w}{\delta_w^*}}{(1-\rho_{w,h}^{*2})^{1/2}}, \frac{\Delta_L+\rho_{w,L}^*\dfrac{y-\Delta_w}{\delta_w^*}}{(1-\rho_{w,L}^{*2})^{1/2}}, -\tilde{\rho}_{h,L}^*\right)\right], \end{aligned}$$

where $\Delta_L = x_L^{'}\beta_L$, $\tilde{\rho}_{h,L}^* = (\rho_{h,L}^* - \rho_{w,h}^*\rho_{w,L}^*)/[(1-\rho_{w,h}^{*2})(1-\rho_{w,L}^{*2})]^{1/2}$ is the correlation coefficient between ε_h^* and ε_L^* given ε_w^* (see Johnson and Kotz 1974, p86-87),

[20] The separate health equations for workers and non-workers can also be justified by findings in Kreider (1999) that non-workers' self-reported health is prone to reporting bias.

$\rho^*_{h,L}$, $\rho^*_{w,h}$ and $\rho^*_{w,L}$ are the correlation coefficients between ε^*_h and ε^*_L, ε^*_h and ε^*_w, and ε^*_w and ε^*_L, respectively. Φ_2 stands for the bivariate standard normal distribution.

For a non-worker with health status *k*, the contribution to the likelihood function is

(a9) $\Pr(h = k, L = 0) = \Phi_2(m_k - \Delta_{h_{nw}}, -\Delta_L, \tilde{\rho}_{h_{nw},L}) - \Phi_2(m_{k-1} - \Delta_{h_{nw}}, -\Delta_L, \tilde{\rho}_{h_{nw},L})$,

where $\Delta_{h_{nw}} = x'_{h,nw}\beta_{h,nw}$ and $\tilde{\rho}_{h_{nw},L} = \delta_{h_{nw},L}$.

The likelihood function of the whole sample is formed using equations (a8) and (a9) for corresponding individuals.

C.2 Estimation results of the augmented model

	Wage equation		Health of workers		Health of non-workers		Selection equation	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Health	0.0557**	0.0245						
Wage			0.0842	0.1634				
Age			-0.1559*	0.0912	-0.3308**	0.1414	-0.5433***	0.1897
Age squared			0.0321	0.0265	0.1098***	0.0329	-0.1057**	0.0484
Married	0.0863***	0.0278	0.0328	0.0579	-0.0642	0.0780	0.1980***	0.0702
Degree	0.3341***	0.0341	0.3075***	0.0987	0.2237*	0.1213	0.5800***	0.0875
Other post_sch qual	0.1023***	0.0292	0.0760	0.0669	0.0116	0.0830	0.1314**	0.0666
Year 12	0.0756*	0.0390	0.1556*	0.0922	0.0594	0.1348	0.2918***	0.1024
Experience	0.1479***	0.0467					0.6225***	0.1704
Experience squared	-0.0197**	0.0089					0.0216	0.0309
Indigenous	-0.0692	0.1878	0.2548	0.2581	0.0221	0.2622	0.0700	0.3253
Born OS En country	0.0168	0.0309	0.0856	0.0719	-0.1498	0.1142	0.0452	0.0842
Born non-En country	-0.0758**	0.0351	0.1528*	0.0781	-0.2930***	0.1061	-0.0080	0.0934
VIC	-0.0552*	0.0285	0.0327	0.0626	0.0530	0.0962	-0.0988	0.0749
QLD	-0.0429	0.0305	-0.1117	0.0706	0.0098	0.1018	-0.0846	0.0776
SA	-0.1423***	0.0375	0.0225	0.0872	-0.0353	0.1322	-0.2350**	0.1032
WA/NT	-0.0132	0.0344	-0.1008	0.0794	0.0458	0.1184	-0.1461	0.0927
TAS	-0.0618	0.0903	0.0531	0.1517	0.0930	0.1602	-0.1874	0.1656
Capital city	0.1246***	0.0227					0.2465***	0.0571
Part-time	0.2642***	0.0373						
Casual	0.0489	0.0362						
Part-time & casual	-0.2080***	0.0629						
Firm size 20-99	0.1717***	0.0278						
Firm size 100-199	0.2001***	0.0380						
Firm size 200-499	0.2449***	0.0440						
Firm size 500+	0.3226***	0.0383						
Firm size unknown	0.0173	0.1620						
Union member	0.0767***	0.0242						
Sight			-0.2928	0.1851	0.3984*	0.2060	-0.2866	0.1903
Hearing			-0.1972	0.1286	-0.2765*	0.1543	0.0162	0.1350
Speech			-0.5719	0.9621	-0.2427	1.2095	0.6119	0.9182
Blackout			-1.3690	1.6726	0.1282	0.4012	-0.7201	0.7728
Learning			-0.6821	1.3107	-0.1702	0.3322	0.1823	0.4602
Arm_finger			-0.3952*	0.2065	-0.3760	0.2588	0.2368	0.2696

Gripping			0.5261	0.5102	0.3258	0.2375	-0.7172**	0.3094
Feet_leg			-0.6951***	0.2127	-0.0948	0.1596	-0.0261	0.1535
Nerve_emotion			-0.4152	0.2839	-0.4299**	0.1823	-0.5925***	0.2058
Physical			-0.6549***	0.1227	-0.3156***	0.1145	-0.4165***	0.1055
Deformity			-0.1811	0.4172	0.6675	0.4523	0.2993	0.5149
Mental			-0.9685**	0.4335	0.0501	0.2288	-0.6933**	0.2842
Breath			-0.6747***	0.2241	-0.2945	0.2016	-0.4181*	0.2322
Chronic pain			-0.2255	0.1892	-0.2436	0.1559	-0.2427	0.1595
LT damage			0.8155	2.3401	0.2955	0.2790	-0.7677**	0.3467
LT ailment			-0.7192***	0.1412	-0.5672***	0.1298	-0.3598***	0.1152
Arthritis			-0.3408***	0.0827	-0.2425**	0.0966	-0.1496*	0.0863
Asthma			-0.0786	0.0781	-0.2905**	0.1289	0.1621	0.1038
Cancer			-0.0444	0.1296	-0.1105	0.1533	-0.1658	0.1556
Bronchitis			-0.0072	0.1804	-0.1308	0.2229	-0.0039	0.2569
Diabetes			-0.6041***	0.1465	-0.2235	0.1523	-0.1037	0.1352
Coronary			-0.3451**	0.1758	-0.5047***	0.1832	-0.1628	0.1679
Hypertension			-0.2681***	0.0781	-0.1432	0.0963	-0.0747	0.0822
Circulatory			-0.6175**	0.2945	-0.2117	0.1953	-0.2811	0.2527
Other condition			-0.1055	0.3872	-0.5921	0.7907	-0.2129	0.5422
Smoker			-0.1674***	0.0492	-0.0322	0.0801	-0.0912	0.0591
Heavy drinker			-0.1409*	0.0806	-0.1079	0.1285	-0.0306	0.0940
Lack physical activity			-0.4904***	0.0825	-0.2464**	0.1033	-0.0858	0.0968
Child 0-4							-0.1583*	0.0815
Child 5-14							-0.0084	0.0677
Aged 55p							-0.3326**	0.1320
Constant	2.4463***	0.0729					0.3952**	0.1637
Cutoff_1			-2.9955***	0.5081	-1.4900***	0.2759		
Cutoff_2			-1.5047***	0.4886	-0.3785	0.2353		
Cutoff_3			-0.1231	0.4827	0.5765***	0.2082		
Cutoff_4			1.2160**	0.4763	1.5375***	0.1887		
$\ln(\sqrt{Var(\varepsilon_w)})$							-0.7578***	0.0110
$Cov(\varepsilon_w,\varepsilon_h)$							-0.0797*	0.0440
$Cov(\varepsilon_w,\varepsilon_L)$							-0.0673	0.0550
$Cov(\varepsilon_h,\varepsilon_L)$							0.0333	0.2101
$Cov(\varepsilon_{h,nw},\varepsilon_L)$							-0.8294***	0.1259
Log-likelihood			-6746.98					
No. obs.			3227					

References

Barrett, G.F. (2002). "The Effect of Alcohol Consumption on Earnings", *Economic Record*, Vol. 78, p79-96.

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

Bhargava, A., Jamison, D., Lau, L.J., and Murray, C. (2001). "Modelling the Effects of Health on Economic Growth," *Journal of Health Economics*, Vol. 20, p423-440.

Bloom, D.E. and Canning, D. (2000). "The Health and Wealth of Nations," Science, Vol. 18, p1207-1209.

Bloom, D.E., Canning, D. and Sevilla, J. (2001). "Effect of Health on Economic Growth: Theory and Evidence," *NBER Working Paper No. 8587*.

Bloom, D.E. and Mahal, A. (1997). "Does the AIDS Epidemic Threaten Economic Growth in Africa", *Journal of Econometrics*, Vol. 77, p105-124.

Bloom, D.E. and Sachs, J.D. (1998). "Geography, Demography, and Economic Growth in Africa", *Brookings Papers on Economic Activities*, Vol. 2, p207-295.

Booth, A. and Wood, M. (2004). "Back-to-front Down-under? Part-time/Full-time Wage Differentials in Australia", Mimeo, Australian National University.

Bound, J. (1991). "Self-Reported Versus Objective Measures of Health in Retirement Models." *Journal of Human Resources*, Vol. 26, p106-138.

Bound, J., Schoenbaum, M., Stinebrickner, T.R. and Waidmann, T. (1999). "The Dynamic Effects of Health on the Labor Force Transitions of Older Workers", *Labour Economics*, Vol. 6, p179-202.

Brazenor, R. (2002). "Disabilities and Labour Market Earnings in Australia." *Australian Journal of Labour Economics*, Vol. 5, p319-334.

Breusch, T. and Gray, E. (2004). "Does Marriage Improve the Wages of Men and Women in Australia?", paper presented at 12th Biennial Conference of the Australian Population Association, September, Canberra, Australia.

Brown, C. and Medoff, J. (1989). "The Employer Size-Wage Effect", *Journal of Political Economy*, Vol. 97, p1027-1059.

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

Campolieti, M. (2002). "Disability and the Labor Force Participation of Older Men in Canada." *Labour Economics,* Vol. 9, p405-432.

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

Davidson, R. and Mackinnon, J, G. (1993). *Estimation and Inference in Econometrics*, 2nd edition, New York: Oxford University Press.

Disney, R., Emmerson, C, and Wakefield, M. (2006). “Ill Health and Retirement in Britain: A Panel Data-Based Analysis”, *Journal of Health Economics*, Vol. 25, p621-649.

Ferraro, K. F. (1980). "Self-Ratings of Health among the Old and Old-Old." *Journal of Health and Social Behaviour,* Vol. 21, p377-383.

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,* Vol. 18, p117-224.

Gray, J. (1997). “The Fall in Men’s Returns to Marriage: Declining Productivity Effect or Changing Selection”, *Journal of Human Resources,* Vol. 32 (3), p 485-504.

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

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

Grossman, M. and Kaestner, R. (1997). “Effects of Education on Health”, in *The Social Benefits of Education,* eds. by Behrman, J.R. and Stacey, N., University of Michigan Press, Ann Arbor, MI: p69-122.

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.

Hildreth, A. (1999). “What Has Happened to the Union Wage Differential in Britain in the 1990s”, *Oxford Bulletin of Economics and Statistics*, Vol. 61, p5-31.

Hirsch, B.T. (2004). “Why Do Part-time Workers Earn Less? The Role of Worker and Job Skills”, *IZA Discussion Paper No*. 1261.

Johnson, N.L. and Kotz, S. (1974). *Continuous Multivariate Distributions*, New York: Wiley, p86-87.

Kreider, B. (1999). "Latent Work Disability and Reporting Bias", *Journal of Human Resources*, Vol. 34, p734-769.

LaRue, A., Bank, L., Jarvik, L. and Hetland, M. (1979). "Health in Old Age: How Physicians' Ratings and Self-Ratings Compare", *Journal of Gerontology*, Vol. 34, p687-691.

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

Lewis, H. G. (1986). *Union Relative Wage Effects: A Survey*. University of Chicago Press, Chicago, Illinois, USA.

Loh, E. (1996). “Productivity Differences and the Marriage Wage Premium for White Males”, *Journal of Human Resources*, Vol. 31, p566-589.

Lye, J.N. and Hirschberg, J. (2004). “Alcohol Consumption, Smoking and Wages”, *Applied Economics*, Vol. 14, p1807-1817.

Maddox, G. and Douglas, E. (1973). "Self-Assessment of Health: A Longitudinal Study of Elderly Subject", *Journal of Health and Social Behaviour*, Vol. 14, p87-93.

Main, B.G.M. and Reilly, B. (1993). “The Employer Size-Wage Gap: Evidence from Britain”, *Economica*, Vol. (60), p125-42.

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

Mincer, J. (1974). *Schooling, Experience and Earnings*, New York: NBER Press.

Mossey, J. M. and Shapiro, E. (1982). "Self Rated Health: A Predictor of Mortality among the Elderly", *American Journal of Public Health,* Vol. 72, p800-808.

Nagi, S. Z. (1969). "Congruency in Medical and Self-Assessment of Disability", *Industrial Medicine and Surgery,* Vol. 38, p27-36.

Rivera, B. and Currais, L. (1999). "Economic Growth and Health: Direct Impact or Reverse Causation?" *Applied Economics Letters,* Vol. 6, p761-764.

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

Stewart, M. (1995). "Union Wage Differentials in an Era of Declining Unionisation", *Oxford Bulletin of Economics and Statistics*, Vol. 57, p143-66.

Tausman, P. and Rosen, S. (1982). "Healthiness, Education and Marital Status", in *Economic Aspects of Health*, ed. by Fuchs, V., Chicago, University of Chicago Press: p121-140.

Wagstaff, A., and van Doorslaer, E. (1994). "Measuring Inequalities in Health in the Presence of Multiple Category Morbidity Indicators", *Health Economics,* Vol. 3, p281-291.

Ware, J., Snow, K. and Kosinski, M. (2000). S*F-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 Economic and Social Research, University of Melbourne.

Wilkins, R. (2003). "Labour Market Outcomes and Welfare Dependence of Persons with Disabilities in Australia", Melbourne Institute Working Paper Series No. 2/03, The University of Melbourne

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.