1351.0.55.049

Research Paper

Examining Association Between Self-Assessed Health Status and Labour Force Participation Using Pooled NHS Data

New Issue

Research Paper

Examining Association Between Self-Assessed Health Status and Labour Force Participation Using Pooled NHS Data

Teressa A. Belachew and Anil Kumar

Analytical Services Branch

AUSTRALIAN BUREAU OF STATISTICS

EMBARGO: 11.30 AM (CANBERRA TIME) FRI 28 FEB 2014

ABS Catalogue no. 1351.0.55.049

Produced by the Australian Bureau of Statistics

INQUIRIES

The ABS welcomes comments on the research presented in this paper.
For further information, please contact Mr Ruel Abello, Analytical Services Branch, on Canberra (02) 6252 6307 or email <analytical.services@abs.gov.au>.

EXAMINING ASSOCIATION BETWEEN SELF-ASSESSED HEALTH STATUS AND LABOUR FORCE PARTICIPATION USING POOLED NHS DATA

Teressa A. Belachew and Anil Kumar
Analytical Services Branch

EXECUTIVE SUMMARY

This research examined the relationship between self-assessed health status and labour force participation using pooled unit-record data from five National Health Surveys (NHS) conducted by the Australian Bureau of Statistics between 1989/90 and 2007/08. Descriptive analysis of labour force participation, health and other selected demographic and socioeconomic factors was conducted. A decomposition of age, period and cohort effects was undertaken to examine their separate effects on labour force participation. A logistic regression model was used to examine the association between participation in the labour force and health status controlling for other relevant demographic/ socioeconomic variables including age, period and cohorts. Two self-reported health indicators, namely, self-assessed general health status and the presence of selected long term health conditions, were alternatively used to represent the health status variable in the model.

During the study period of 1989/90 to 2007/08, overall labour force participation in Australia increased since 2001 mainly on account of an increase in female participation. In terms of age, participation began to decline for both males and females by age 55, with a much larger decline for women. By age 55–64, around two-thirds of males were still in the labour force compared to only a third of females.

With regards to health, a large majority of the Australian population (85%) enjoyed good to excellent self-assessed general health over the study period. The remaining proportion of the population had fair or poor self-assessed health status. There was an upward trend in some of the long-term health conditions, such as arthritis, asthma, diabetes and heart disease, while cancer stayed somewhat steady during the period under study.

Results from the logistic models showed that health status was an important factor associated with participation in the labour force, and this relationship was found to be robust to the alternative measures of the health variable used in the analysis. People with fair or poor self-assessed health status were less likely to participate in the labour force compared to those with good or better health. Overall, the association between health status and labour force participation appeared to be stronger for females than for males. The probability of participation of an 'average' female with fair or poor health was 0.162 lower than that of an 'average' female with good or better health while the probability for an 'average' male with fair or poor health was 0.099 lower than the same male with good or better health. There were also strong negative relationships between major chronic diseases – such as arthritis, asthma, cancer, diabetes and heart disease – and both males' and females' participation in the labour force.

In terms of other variables in the models, age played an important role in individuals' participation in the labour force. Marital status, non-school qualifications, proficiency in spoken English, and Indigenous status were other variables that were found to have significant influence on males' and females' likelihood of participation in the labour force. Additionally for females, presence of dependent children and location had an influence on their participation in the labour force. Period appeared to have largely a positive effect on females' participation in the labour force, while it had a negative effect on that of males'. After controlling for other variables, some cohort effects were also observed for both males and females, with the youngest cohorts showing lower participation in the labour force compared to their oldest counterparts.

CONTENTS

EXAMINING ASSOCIATION BETWEEN SELF-ASSESSED HEALTH STATUS AND LABOUR FORCE PARTICIPATION USING POOLED NHS DATA

Teressa A. Belachew and Anil Kumar
Analytical Services Branch

ABSTRACT

This paper examined the association between health status and participation in the labour force using pooled unit-record data created from the ABS's five consecutive cross-sectional National Health Surveys. Descriptive analysis of labour force participation, health status and other selected variables was conducted. A simple age–period–cohort decomposition model was used to examine the relative influence of these three factors on labour force participation. A logistic regression model was used to examine the association between participation in the labour force and health status, controlling for other relevant demographic/socioeconomic variables including age, period and cohorts. Two self-reported health indicators, namely self-assessed general health status and the presence of selected long-term health conditions, were alternatively used to represent the health status variable in the model. The empirical results suggested a statistically significant negative association between health status and labour force participation, and this relationship was found to be robust to the alternative measures of the health variable used in the analysis. Based on changes in predicted probabilities for both males and females, those with fair or poor self-assessed health status were less likely to participate in the labour force compared with those with good or better health. Likewise, there were also a strong negative relationships between major chronic diseases, such as arthritis, diabetes, heart disease, cancer and asthma, and labour force participation for both males and females, with the relative importance of these diseases varying for the sexes. The association between health indicators and labour force participation appeared to be stronger for females than for males.

1. INTRODUCTION

One of the issues of interest in the intergenerational report is the opportunity cost of sickness to workforce participation and productivity (Australian Government Treasury, 2010). It is perceived that overall poor health and/or prevalence of long-term health conditions, collectively known as national health priority areas, such as arthritis, asthma, cancer, diabetes, cardiovascular disease, and hypertensive disease, are increasing over time. These are believed to have an adverse effect on labour force participation. It has also been noted that people who are not in the labour force or unemployed have worse health status than those who are in the labour force or employed due possibly to higher prevalence of health risk factors among the former group than the latter. On account of these general perceptions, there has been a strong interest in the actual empirical links between self-assessed health status and/or individual long-term health conditions, that is, major chronic diseases, and participation in the labour force. This makes the investigation of the association between health status and participation in labour force a worthwhile research.

The relationship between health status and participation in the labour force has been analysed by different researchers (Cai and Cong, 2009; Cai and Kalb, 2006; Kalwij and Vermeulen, 2008). The effect of chronic diseases on participation in the labour force has also been studied (Cai and Cong, 2009). Poor health is believed to affect a person's capacity to work productively (Stronks *et al.*, 1997 and Bartley and Owen, 1996 – both cited in Jose *et al.*, 2004). Chronic health conditions have been found to diminish physical and mental capabilities, leading to disruption in normal work functioning (Jose *et al.*, 2004; Chirikos, 1993; Mathers, 1994 and Bound *et al.*, 1998).

A variety of datasets have been used in analysing the association between health and participation in the labour force. For instance, Bound *et al.* (1998) examined the dynamic relationship between health status and labour force behaviour among older working-age adults in the United States using longitudinal data. Cai and Kalb (2004) and Cai and Cong (2009) explored the effect of health on labour force participation in the Australian context using the Household, Income and Labour Dynamics in Australia (HILDA) survey data. Other analyses that examined the relationship between health and labour market outcomes were based on cross-sectional surveys (Mathers and Schofield, 1998; Bartley, 1994; and Wilson and Walker, 1993). Jose *et al.* (2004) examined the association between non-participation in the labour force and health using unit-record pooled data from three ABS's repeated National Health Surveys (1989/90, 1995 and 2001). These and similar studies (Kumar and Chessman, 2009; Kumar *et al.*, 2009) have favourably argued for the pooling of relevant datasets for more enriched analysis.

The purpose of this study was to examine the association between health status and participation in the labour force after controlling for diverse variables and thereby predict probabilities of participation in the labour force and their changes, that is, marginal effects, which could potentially arise from changes in health status and other key variables. The study used pooled unit record data from the ABS's five cross-sectional National Health Surveys (NHS) conducted between 1989/90 and 2007/08. The pooling of the five cross-sectional surveys allowed us to have increased sample size and to study age, period and cohort effects.

The main research questions this study aimed to address were: What relationship exists between indicators of self-assessed health status and participation in the labour force? Does poor self-assessed health preclude participation in the labour force? The core contribution of this work lies in using the above large and nationally representative dataset, a relatively large number of explanatory variables and alternative specifications of the health variable to check the robustness of empirical results. The study thus addressed both methodological and empirical issues involved in the investigation of these phenomena.

The rest of the paper is organised as follows. Section 2 presents the conceptual framework of studying labour market behaviour and the role of demographic, socioeconomic and health characteristics. Section 3 describes the data used in this study and the construction of pseudo-panel data. Section 4 describes the analytical techniques used in the study. Section 5 presents results from descriptive analysis. Section 6 looks at the decomposition and estimation of age, period and cohort effects. Section 7 presents results from logistic regression models. Section 8 presents conclusions of the study.

2. CONCEPTUAL FRAMEWORK OF LABOUR MARKET BEHAVIOUR

Based on the theoretical framework of labour market economics, an individual's behaviour towards work and the extent of work is predicted from the standard labour-leisure choice model (Hunter and Gray, 2001) and tends to be asymmetric across demographic groups (Hotchkiss and Robertson, 2012). A person's labour supply decisions involve a trade-off between time spent at home on market substitution activities, leisure, and paid work (Benjamin *et al.*, 2002; Gray and Hunter, 2002; Prowse, 2009). Evidently this decision is a highly complex one and involves many factors. For instance, an individual's decision to supply labour is considered in terms of his/her household/family and interactions that occur within it (Gray and Hunter, 2002). Within the above framework, researchers have used a wide variety of explanatory variables in labour force participation models.

Some of the common determinants of individual's behaviour towards participation in the labour force included: demographic, social and psychological variables (Grogan and Koka, 2010; Prowse, 2009). The presence of dependent children in a household plays an important role particularly in females' labour force participation (Contreras *et al.*, 2012; Oshio *et al.*, 2011; Grogan and Koka, 2010) and number of adults (Grogan and Koka, 2010; Oshio *et al.*, 2011). Observable personal characteristics, such as age, gender, marital status, and educational levels are expected to play important roles in individuals' behaviour towards work (Grogan and Koka, 2010; Oshio *et al.*, 2011).

Wage rates, taxation and government transfers are shown to be important determinants of labour force participation (Prowse, 2009). Job opportunities and consequently participation in the labour force are also expected to vary by institutional and contextual variables, such as location. Being located in a state or section of state or region where there are limited job opportunities tends to reduce participation in the labour force. Ethnic and religious backgrounds are also among theoretically relevant variables. For instance, Indigenous Australians tend to experience significant labour market disadvantage and consequently lower labour force participation rate (Gray and Hunter, 2002). Oshio *et al.* (2011) found cohort effects on labour force participation. The standard labour/leisure choice model suggests that labour market conditions in a given period have influence on individual's labour force participation (Hotchkiss and Roberson, 2006). The health status of an individual is also an important factor in labour supply decision as changes in health may affect individual's preference between work and leisure (Cai and Cong, 2009; Jones *et al.*, 2011). Health status may also affect the time horizon over which the supply of labour is made (Cai and Cong, 2009 citing Chirikos, 1993).

Given the focus of this study on the association between health and labour force participation, an appropriate measure of health is required. Health status generally refers to a person's state of wellbeing, and its meaning can vary according to individual or community expectations and context (ABS, 2001). It is a multifaceted concept and includes physical, mental and social wellbeing (AIHW, 2010; WHO, 1946). In the literature, there does not appear to be any consensus as to what is an ideal measure of health outcome[1] and in practice a number of alternative measures have been used, such as self-reported health status, presence of chronic health conditions, life expectancy, and some combination of health indicators (AIHW, 2010).

For the purpose of this study, we used self-reported health status as it has been used as a global measure of health in various empirical researches (Simon *et al.*, 2005; Crossley and Kennedy, 2002). This measure has been found to predict well the onset of disability and subsequent mortality and it is considered by many to be a useful measure of adult health status (Wagstaff and Van Doorslaer, 1994).

We used two alternative measures of health, namely self-reported general health status and presence of major chronic health conditions, both of which are available in the NHS.[2] For the major chronic health conditions, two variant measures of health outcome were examined. The first one was the presence of each of the five major chronic diseases, namely, arthritis, asthma, cancer, diabetes, and heart disease, as reported at the discretion of respondents. These health conditions were identified as the National Health Priority Areas (ABS, 2010 and 2008) as they have been associated with a higher burden of disease and accounted for a high financial burden[3] in Australia (ABS, 2010 and 2008). The second measure of the presence of chronic health conditions was the prevalence of one or more of these health conditions, expressed as a binary outcome, which took a value of 1 if a person had one or more of these conditions and 0 otherwise. These alternative measures of health status were used in order to test the robustness of the relationship between health status and labour force participation and see whether this relationship was sensitive to the health measure chosen.

1 This is especially so given the fact that no single instrument can measure all possible outcomes of interest in health at a group or person level.

2 In the NHS, self-reported health status has been captured by the question: "In general would you say your health is: excellent, very good, good, fair or poor?" while questions relating to the prevalence of major disease conditions is based on disease conditions classified under the ICD9 and ICD10 Disease classifications.

3 For instance, in 2004/05, the health expenditure on seven major disease groups accounted for $25.5 billion (ABS, 2010) and $22.3 billion in 2000/01 (ABS, 2008).

3. DATA

This study used pooled micro data from five National Health Surveys (NHSs). These surveys were conducted by the Australian Bureau of Statistics (ABS) in 1989/90, 1995, 2001, 2004/05 and 2007/08.[4] The NHS provides information on a range of health related issues and demographic and socioeconomic characteristics of the Australian population aged 0 and over. Health-related data items in each of the NHS's included individuals' self-assessed health status, prevalence of self-reported long-term health conditions, use of health services and facilities and prevalence of prominent health risk factors, such as smoking, alcohol consumption, exercise, diet and obesity. Demographic and socioeconomic information on respondents in the surveys included age, sex, marital status, presence of dependent children, indigenous status, country of birth, non-school qualifications, proficiency in spoken English, labour force status and personal and household income. The data also contained information on location of respondents, such as state, capital city/regional areas.

Prior to pooling data on selected variables from the surveys, an assessment of the surveys was made in order to check their comparability and consistency. Given the repeated nature of the NHSs, they were found to have more or less similar survey design, scope, coverage, sampling unit, reporting method, mode of survey and weighting method. Questionnaire wordings for most variables of interest were also found to be generally similar across the surveys. Where there were some differences with respect to some variables, efforts were made to align their definitions and/or categories as close as possible across the surveys prior to pooling the data.[5] For example, if the categories of variables were different across the surveys, the categories were collapsed to a minimum number to make them consistent and comparable across the surveys.[6] This process made pooling the data across the surveys feasible. Although not all variables of interest were available in all the surveys, every effort was made to ensure the inclusion of all variables of interest on which data were available in compiling the pooled dataset.

---

4 Health surveys conducted by the ABS in 1977/78 and 1983 also collected similar information but they are not part of the NHS series (ABS, 2006). These earlier surveys were not pooled together because of data comparability problems. Earlier NHS's were conducted every six years but commencing with the 2001 survey, the NHS was conducted every three years.

5 For instance, in the case of long-term health conditions, effort was made to align the definition of heart disease in 2001, 2004/05 and 2007/08 to that of the earlier two surveys (1989/90 and 1995), which had a narrower definition, to ensure comparability.

6 This was the case for example for health status, labour force status and educational qualification variables where there were fewer categories or slightly expanded categories in some surveys, particularly between the earlier two surveys and the latter three surveys.

The combined sample size of the pooled dataset was 181,626.[7] Appendix A describes the steps involved in the construction of the pooled dataset. Appendix B discusses survey and data comparability across the surveys. Appendix C provides a list of the main variables available in the pooled dataset.

The pooled repeated cross-sectional health surveys allowed us to create cohorts using the pseudo-panel or pseudo-longitudinal structure of the dataset. This was done by first grouping individuals by cohort. A cohort is defined as a group of individuals whose members share similar experiences or have similar characteristics which remain the same over all age and time periods. While cohorts can be defined in a number of ways,[8] year of birth has frequently been used to define them and the same was adopted here. A birth cohort is obtained by subtracting age from period. Once cohorts are defined by birth year, then these groups can be considered to be represented by individuals in older age groups in subsequent surveys. For example, persons aged 18–20 years in 1989/90 NHS can be assumed to be represented by those aged 24–26 years in the 1995 NHS, by those aged 30–32 years in the 2001 NHS, by those aged 33–35 years in the 2004/05 NHS and by those aged 36–38 years in the 2007/08 NHS. Other cohorts in the pooled dataset can be tracked in a similar fashion.

Appendix D presents the pseudo-panel structure of the pooled dataset for labour force participation rate with age along the rows, period along the columns and cohorts represented along the diagonals.[9] Although the creation of cohorts does not allow us to follow individuals over time, it allows us to follow a group of individuals who share a common characteristic over the life cycle and observe how the group differs from other groups. Each cohort is assumed to be a representative of the population in that cohort at each period over time.

This study was restricted to the population aged 18–64 years (hereafter referred to as study population) because we wanted to focus on the influence of health status on the formal working age population. Future work may consider those aged 18–74 years.

---

7 The respective sample sizes of each of the five surveys in the pooled dataset were 54241, 53828, 26863, 25906 and 20788.

8 For instance, cohorts can be defined in terms of birth year e.g. all those who were born between 1970 and 1975, migrants arriving in the country during 1980/90, disease cohort, education cohort, etc.

9 Since the NHSs have not been collected at even intervals, with the first three surveys collected at six year intervals and the last two surveys at three year intervals, we do not have equal width or intervals for age and period and hence we do not have a normal straight line diagonal. But the cohorts formed at the diagonals do satisfy the condition $C = P–A$, that is., given period (P) and age (A) the birth cohort (C) is obtained by subtracting age from period.

4. ANALYTICAL TECHNIQUES

In order to describe household and individual demographic, socioeconomic and health characteristics and locational variables, we used univariate and bivariate descriptive analytical techniques. T-test and chi-square test were used to compare people in the labour force (ILF) and those not in the labour force (NILF) and to compare females and males.

We also used the 'age–period–cohort' (APC) accounting model as proposed by Yang *et al.* (2008) to analyse and decompose age, period and cohort effects on labour force participation rates and to gain insights into the relative importance of these three dimensions. Age effect relates to changes that occur as people age.[10] Period effect relates to the effect of conditions prevailing at a given period or time point.[11] Cohort effect relates to the effect of specific or unique group characteristics or shared experiences on the outcome of interest.[12] APC effects are particularly of importance when we observe some outcome or event over time. In fact any phenomenon that has a time dimension has APC effects (McKenzie, 2005).

The basic APC decomposition accounting model can be expressed as in Equation (1) (Yang *et al.,* 2008).

$$g\left(M_{ijk}\right) = \mu + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ijk} \tag{1}$$

where M_{ijk} denotes the outcome of interest for the particular APC group or cell, μ denotes the intercept, α_i denotes the coefficient for the i-th age group, β_j denotes the coefficient for the j-th time period, γ_k denotes the coefficient for the k-th cohort, ε_{ijk} denotes a random error term and $g(\cdot)$ is the link function relating M_{ijk} to the effects.

Equation (1) is a class of generalised linear models and can take alternative functional forms, such as linear, log-linear or logistic (Yang *et al.* (2008). Labour force participation rate in each APC cell ranges between 0% and 100%. The corresponding probability of participation in each cell ranges between 0 and 1; the appropriate link function would then be a logit model that can be expressed as follows:

$$g\left(M_{ijk}\right) = \theta_{ijk} = \ln\left[\frac{m_{ijk}}{1 - m_{ijk}}\right] = \mu + \alpha_i + \beta_j + \gamma_k \tag{2}$$

10 For example, labour force participation being low when people are young and in education, rising thereafter and peaking during prime age years, before declining as people retire.

11 For instance, individuals experiencing lower unemployment during times of economic boom and higher unemployment during times of recession.

12 For example, older cohorts may have lower education levels and hence lower participation in the labour force compared to younger cohorts who are more educated.

where θ_{ijk} is the natural log of the odds of participation in the labour force and $m_{ijk} = M_{ijk} / 100$ is the probability of participation in the labour force for a given APC cell; the other parameters are as defined above. In the estimation of Equation (2), the APC parameters were each constrained such that their effects add up to zero[13] (Yang *et al*., 2008).

To examine the association between labour force participation and health at the individual level and incorporating other relevant demographic and socioeconomic variables, in addition to the age, period and cohort variables, we used a multiple logistic regression expressed as a logit model (Equation (3)):

$$\operatorname{logit}(P_h) = \ln\left[\frac{P_h}{1-P_h}\right] = \varphi + \eta_1 X_{h1} + \ldots + \eta_L X_{hL} \tag{3}$$

where $\ln[.]$ is the natural logarithm, P_h is the probability that person *h* was in the labour force, φ is the intercept term, η's are *L* regression parameters, and the X_h's are a set of *L* explanatory variables representing individual *h*'s observed characteristics.[14]

An issue that has been well recognized in the literature in examining the relationship between many socioeconomic variables, like health status and participation in the labour force, is the endogenous nature of the variables. The relationship between these two variables is not one-directional as health status can influence participation in the labour force and vice versa.[15] While we are aware of the potential endogeneity problem in the estimation process, the core objective of this work was limited to examining the association between the two variables rather than looking into their causal relationship.[16]

13 That is $\sum_i \alpha_i = \sum_j \beta_j = \sum_k \gamma_k = 0$.

14 It may be noted that Equations (2) and (3) share the same functional form (logistic / logit). Equation (2) was used to estimate the log- odds of participation at the APC cell or group level with the APC variables as independent variables. Equation (3) was used to estimate the log-odds of participation at individual level with APC and other socio-demographic variables as explanatory variables.

15 On one hand, a person's health may impact on the decision to participate in the labour force or not. Incapacity due to ill-health could be one of the reasons for not being in the labour force (Prowse, 2009). On the other hand, a more physically demanding, stressful work and/or working conditions that expose working individuals to hazardous substances could adversely affect health. Those individuals who participate in the labour force or are working could be more vulnerable to workplace illness and injuries. Labour force or employment status, such as not being in the labour force or being unemployed could expose people to health risk factors and as a consequence lead to poor health. Extended periods of unemployment and the potential resultant financial and/or psychological stress may also contribute to poor health (ABS, 2011).

16 Investigating the causal relationship between the two variables using instrumental variables method and/or true panel data might be an area of further work.

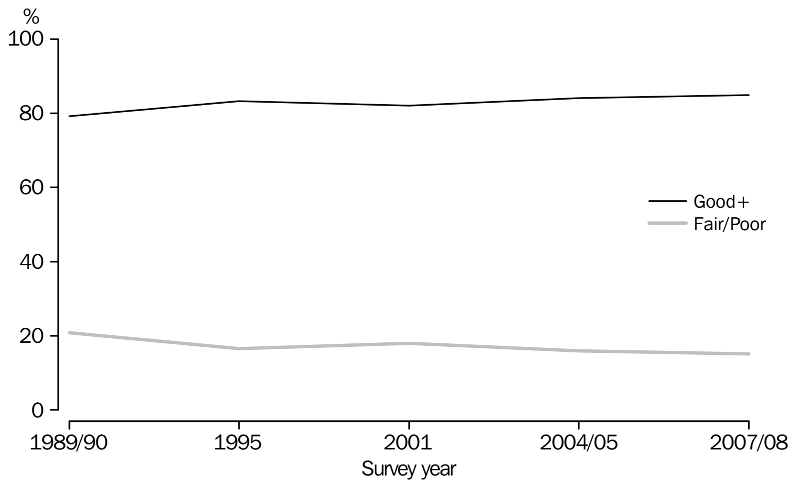
5. DESCRIPTIVE ANALYSIS

This section presents descriptive statistics and trends in health status, prevalence of major chronic diseases and relationship between labour force participation and selected explanatory variables. Appendix E presents summary statistics of selected variables in the pooled dataset for the population aged 18–64 years.

5.1 Health status

As noted in Section 2 above, health status was represented by self-reported health indicators, namely self-assessed general health and the presence of selected long-term health conditions. Self-assessed general health status is an important variable for which respondents reported whether their health was excellent, very good, good, fair or poor. These were regrouped into two broad categories, namely good or better and fair or poor health.[17] Figure 5.1 shows the trend in self-assessed general health status using these two categories. In 1989/90, 83.2% of Australians aged 18–64 years rated their general health as good or better and 16.8% as fair or poor health. In 2007/08, the proportion of the people aged 18–64 years who assessed their health as good or better increased to 87.6% while 12.4% reported fair or poor health. Over the study period of 1989/90 to 2007/08, on the average around 85% of the population aged 18–64 years reported good or better health status and 15% assessed their health as fair or poor.

5.1 Trend in self-assessed health status (%)*



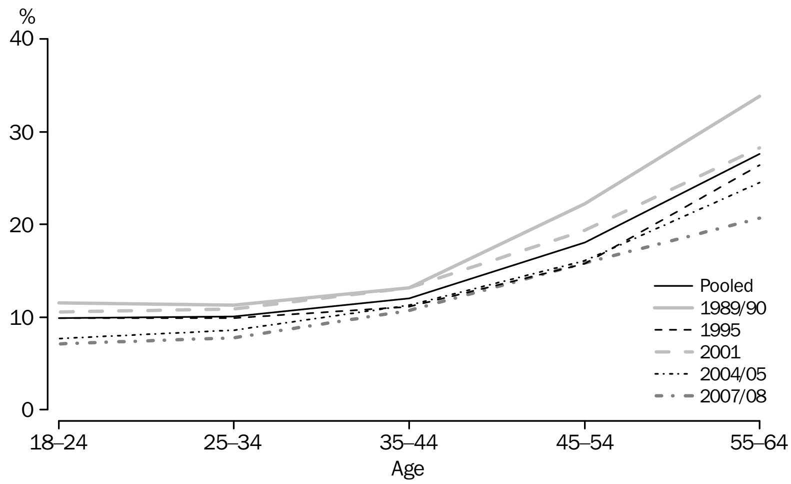
*Based on pooled data

17 It may be noted that in NHS 1989/90, self-assessed health status was reported as a four category variable (excellent, good, fair, poor) while it was reported as a five category variable (excellent, very good, good, fair, poor) in the subsequent surveys. Collapsing the variable in the later surveys from 5 categories to 4 categories of the 1989/90 survey in order to make all the surveys comparable was not possible due to some incompatibility between the definitions of the categories between the 1989/90 survey and the later four surveys. Hence we opted for the two categories so as to make the variable comparable across all the five surveys.

It is evident that through time there has been a slight increase in the proportion of those who assessed their health as good or better and a slightly declining trend in the fair or poor health between 1989/90 and 2007/08.

The proportion of the population that reported fair or poor self-assessed health status significantly increased with age (figure 5.2). The proportion rose steadily particularly after age of 25–34 years. However, closer examination of the proportions of the people that reported fair or poor self-assessed health status by age has revealed a general fall in the proportions for all age groups over time.

5.2 Proportion of population aged 18–64 years that reported fair or poor self-assessed health status, by age (%)*

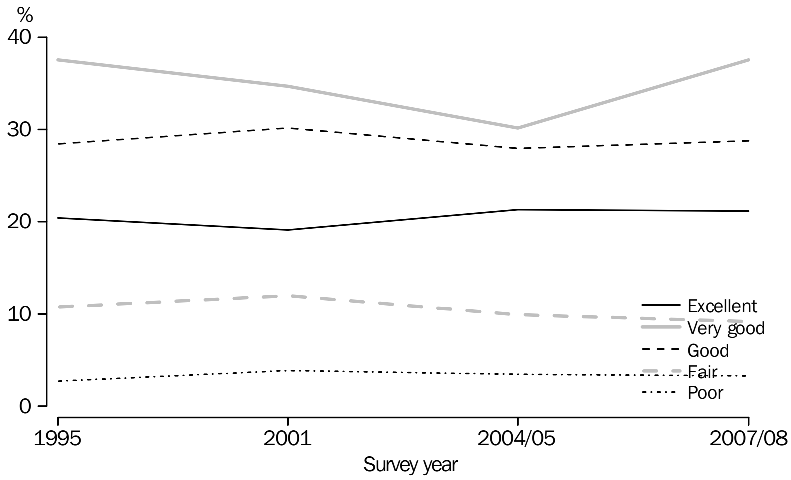


* Based on pooled data

Detailed account of the health status revealed that on the average around 20.4% of people reported excellent health, 36.9% very good health, 28.8% good health, 10.7% fair health and 3.2% reported poor health during the 1995 to 2007/08 period (figure 5.3).

With regard to the prevalence of each of the major long-term health conditions among the population aged 18–64 years, about 12.6% reported arthritis, 9% reported asthma, 1.4% reported cancer, 2% reported diabetes and 2.0% reported heart disease over the 1989/90 to 2007/08 period.

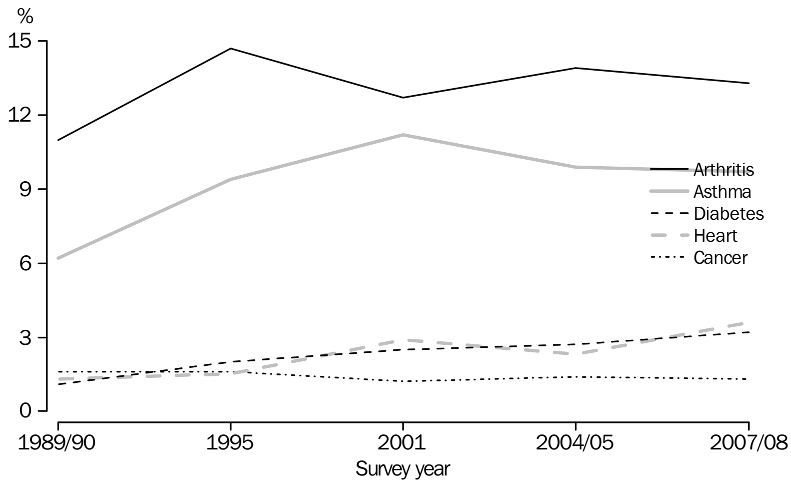
5.3 Trend in self-assessed health status, 1995 to 2007/08 (%)*



* Based on pooled data

Prevalence rates for these major health conditions over time are shown in figure 5.4. These rates have been age standardised. As these disease conditions are likely to be affected by the ageing of the population over time, age standardisation of the crude rates enables us to make valid comparisons across surveys by eliminating the effect of changing age structure over time.[18] With the exception of cancer, the other four disease conditions appeared to exhibit an upward trend over the study period, with some volatility noted for arthritis and asthma.

5.4 Trend in age-standardised prevalence rates for major long-term health conditions for persons aged 18–64 years (%)*

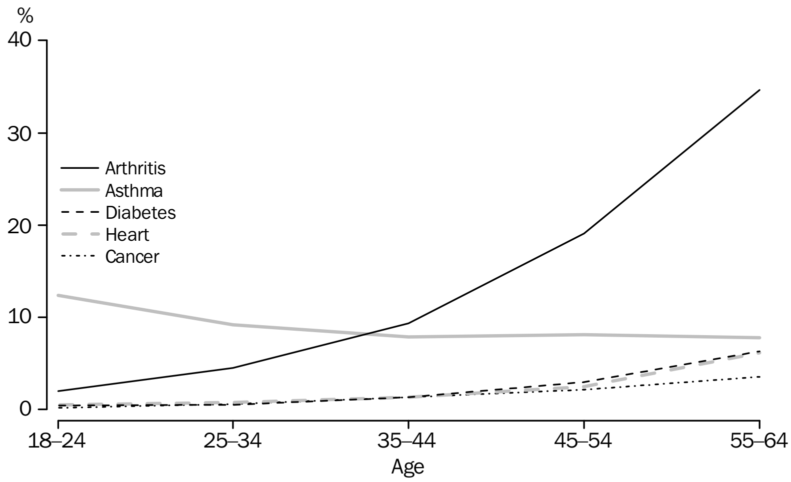


* Based on pooled data

18 Here we standardised the crude prevalence rates for the other years to the age distribution of 2001 using the direct method of age standardisation. The choice of 2001 as the standard or reference year was in line with the current practice by the ABS (ABS, 2004).

The proportion of people with long-term health conditions has been increasing with age (figure 5.5). For instance, the proportion of people with arthritis sharply rose as age increased. This was particularly true after the age of 35–44 years. Diabetes and heart disease were characterised with a gentle increase with age, while asthma exhibited a gentle decline with age.

5.5 Proportion of population aged 18–64 years with major chronic health conditions, by age (%)*

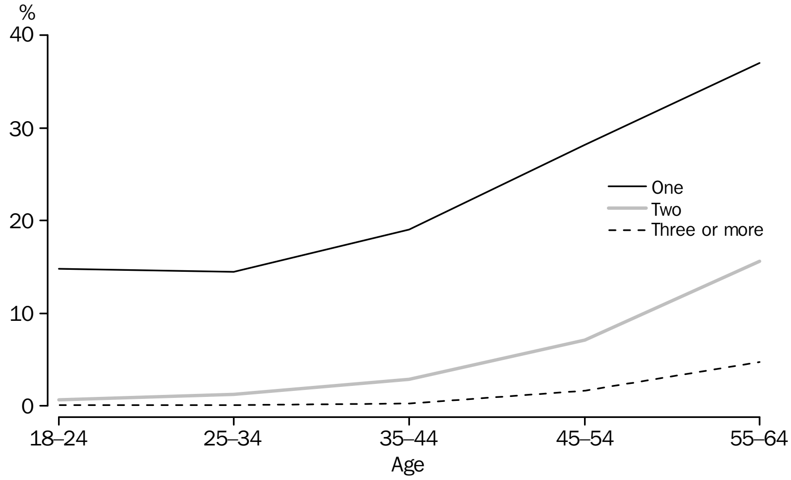


* Based on pooled data

Attempt was also made to assess the number of major long-term health conditions individuals had in relation to the five conditions considered in the paper. Over the study period, about 71.7% had no condition, 22.1% had one condition, and 5% had two and 1.2% had three or more chronic health conditions. Overall, about 28.3% of the people aged 18–64 years had one or more conditions. This proportion was double that of the proportion of the study population that reported fair or poor self-assessed general health status. This suggests that not all long-term health conditions have led to adverse self-assessed general health status. However, there were statistically significant associations between the general health status and all of the five major health conditions.

The proportions of the study population that had one or more long-term conditions were increasing with age (figure 5.6). The increase was the sharpest for those with one condition followed by those with two conditions. The proportion of those with three or more conditions gently rose after age of 35–44 years. With reference to trend over time, the proportion of people that had one condition rose between 1989/90 and 2004/05 and then stabilised (figure 5.7). The proportions of people with two, three or more long-term health conditions exhibited an upward trend over the study period.

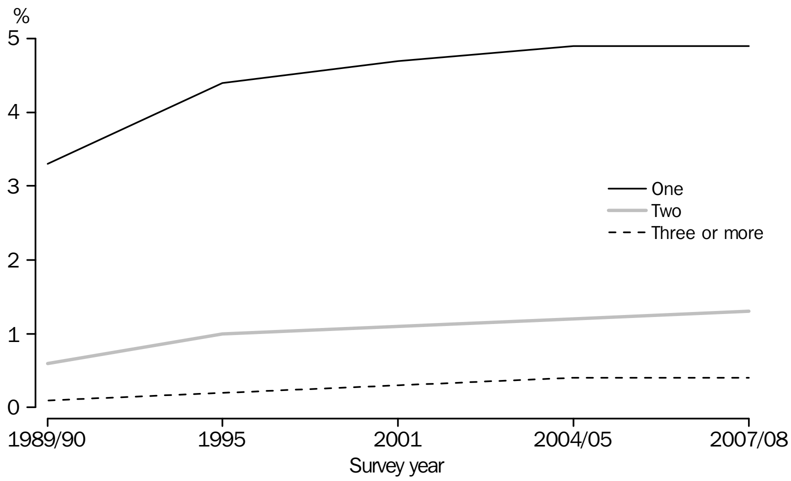
5.6 Proportion of population aged 18–64 years with one or more major health conditions, by age (%)*



* Based on pooled data

People in the labour force were significantly more likely to report good or better self-assessed health status compared with those not in the labour force. Likewise employed people were significantly more likely to report good or better general health compared with those unemployed. Household heads with dependent children were significantly more likely to report good or better self-assessed health compared with those without dependent children. After controlling for age, this remained the case for people in the age bracket of 18–54 years. This situation reversed for those aged 55–64 years.

5.7 Trend in proportion of population aged 18–64 years with one or more major health conditions, 1989/90 to 2007/08 (%)*



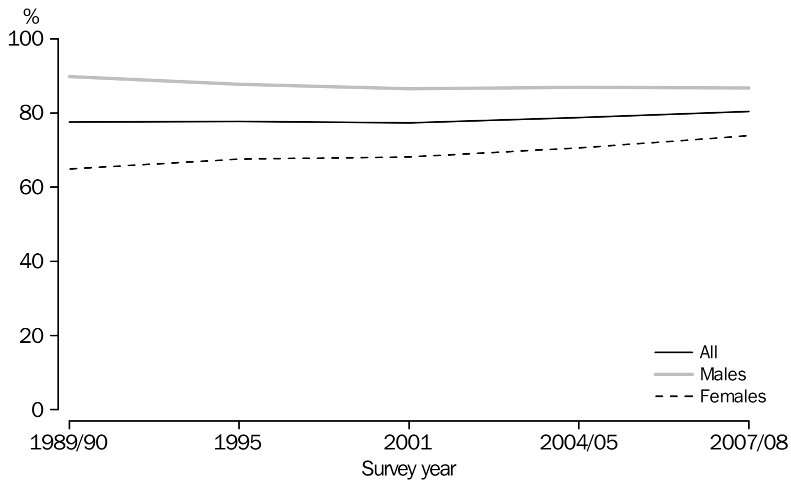
* Based on pooled data

Males, married people, those from low socioeconomic advantaged areas, those with lower levels of non-school qualifications, those having difficulty with spoken English and those born overseas were more likely to report fair or poor self-assessed health status compared with their counterparts. There were also statistically significant differences among the states/territories with respect to the proportions of people who reported self-assessed health status during the period under consideration. The states/territories could be divided into four groups based on the proportions of fair/poor health status: Australian Capital Territory – 12%, Victoria – 14%, New South Wales/Queensland/Northern Territory – 15% and South Australia/Tasmania – 16%.

5.2 Participation in the Labour Force

Trend in labour force participation over the study period is given in figure 5.8. Overall labour force participation appears to have increased since 2001 mainly on account of an increase in females' participation. Males' participation slightly declined between 1989/90 and 2001 and stabilized since then. Females' participation continued to increase, rising from 65% in 1989/90 to 74% in 2007/08. Despite the increase in females' participation, the rate still remained below that of males' participation. Average participation rates for males and females were 88% and 68%, respectively, during the study period.

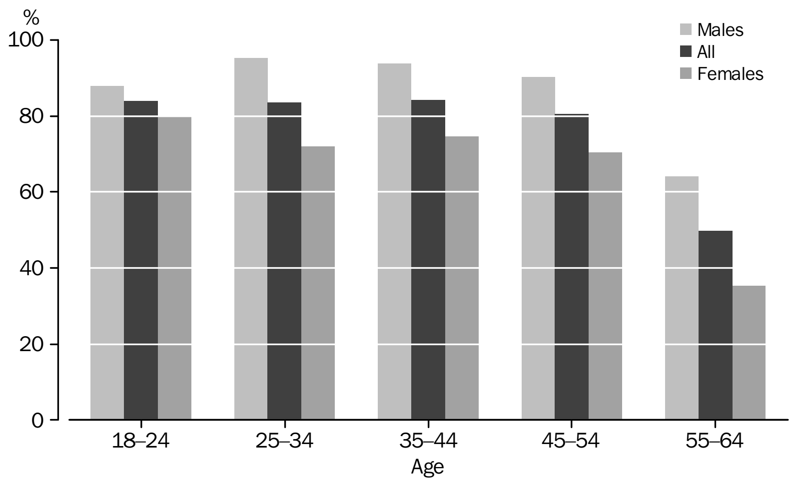
5.8 Labour force participation, by survey year, 1989/90 to 2007/08 (%)*



* Based on pooled data

The relationship between participation in the labour force and age appeared to be strong (figure 5.9). Participation was high during the prime ages (25–44 years), especially for males, before declining from the age of 45 years onwards. For females, participation was the highest at the age of 18–24 years before declining between the age of 25–34 years, largely reflecting the child bearing years. It increased slightly at the age of 35–44 years and then, like males, declined after the age of 45–54 years, but at a much faster rate. By the age of 55–64 years, around two-thirds of males were still in the labour force compared to only a third of females.

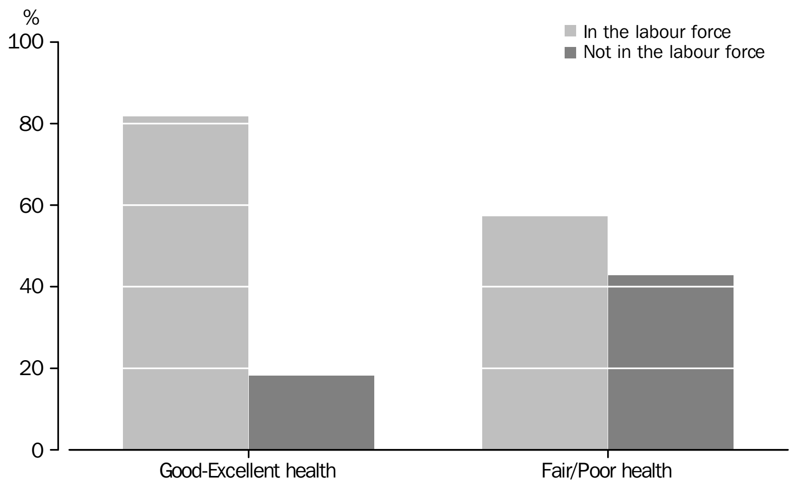
5.9 Labour force participation, by age (%)*



* Based on pooled data

People with good or better self-assessed health status were significantly more likely to be in the labour force compared with those with fair or poor self-assessed health status as shown in figure 5.10.

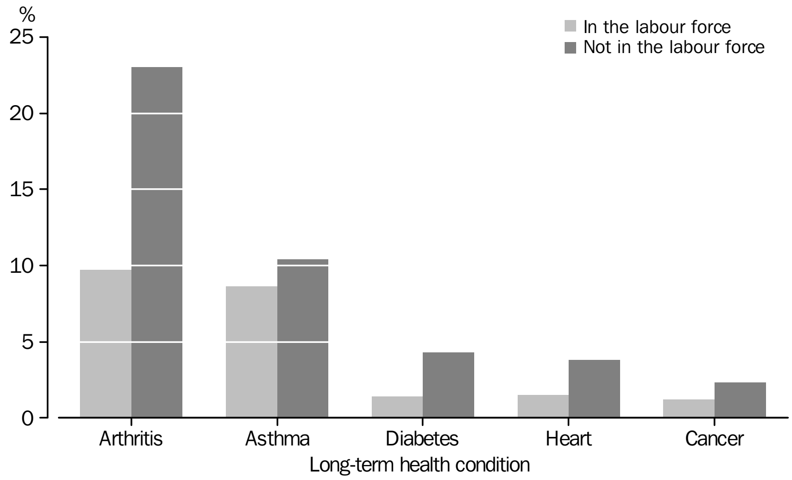
5.10 Labour force participation, by self-assessed health status (%)*



* Based on pooled data

Figure 5.11 displays the proportions of people with long-term health conditions against the labour force status. It is evident that the proportions of the people with each of the major conditions out of the labour force were significantly higher than those of in the labour force. This was particularly pronounced in the case of arthritis, diabetes and heart disease.

5.11 Labour force status, by major long-term health conditions (%)*



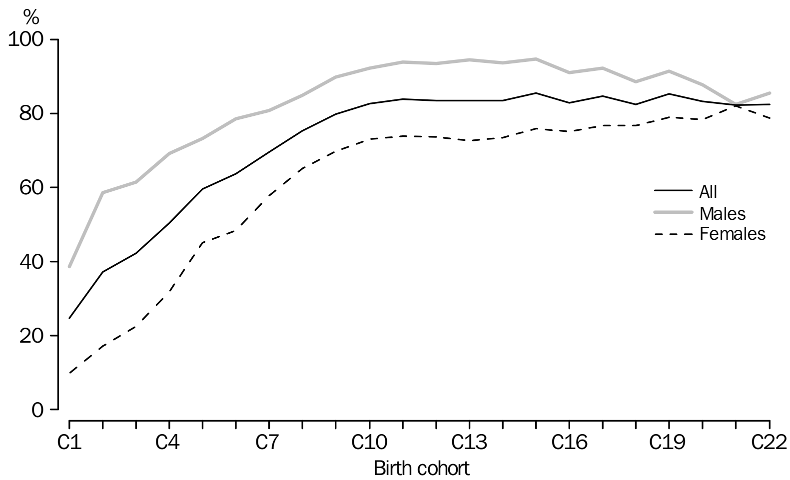
* Based on pooled data

Other variables that were found to be significantly associated with labour force participation included presence of dependent children in household, marital status, country of birth, Indigenous status, non-school qualifications, proficiency in spoken English and location. These relationships are illustrated by graphs presented in Appendix F.

6. DECOMPOSING AGE, PERIOD AND COHORT EFFECTS

In the previous section, we saw some strong relationships between labour force participation and age and period. We can similarly show labour force participation by cohort (figure 6.1). Cohorts here were defined in terms of a three-year birth cohorts which gave a total of 22 cohorts, with the oldest cohort (C1) born during 1925–26 and the youngest cohort (C22) born during 1987–89.[19]

6.1 Labour force participation, by cohort (%)*



* Based on pooled data

However, diagrams that show the relationship between age, period or cohort and labour force participation show only crude or unadjusted age, period or cohort effects, as each does not control for the other two effects that may also exist. For example, the relationship between period and labour force participation may ignore the effects of age and cohort that may also be affecting participation. The large difference in participation rate between the older and younger cohorts as shown in Figure 6.1 could simply be reflecting the age differences between the younger and older cohorts, rather than any inherent differences in the cohorts. Decomposing these APC effects and separating them can help to identify which of these effects, if any, are present and which effects are more dominant than the others in explaining participation in the labour force.

The basic APC decomposition model presented in Equation (2) in Section 4 allows us to estimate the net effects of each of the three variables, i.e. estimating the effect of one while controlling for the other two effects. However, the inclusion of all three APC variables in the model poses a problem in model estimation. This is because the APC variables are not independent of each other but any one variable is a linear

19 Given that the first three surveys were six years apart and the last two surveys were three years apart, we could define age in terms of three-year age groups and define cohorts accordingly in terms of three-year birth cohorts as this enabled us to trace the defined cohorts over the five surveys.

combination of the other two. For instance, given A and P, we can determine the birth cohort (C), since C = P – A.[20] This gives rise to the problem of perfect collinearity or the 'identification' problem in which case it is not possible to simultaneously estimate the true effects of APC, unless some additional constraints are imposed on the parameters of some of these variables.

Several methods have been proposed to resolve this problem. For the purpose of this study, we used an approach that is based on the intrinsic estimator method proposed by Yang *et al.* (2008) and used by Kumar *et al.* (2009). A more detailed discussion of APC decomposition and proposed solutions, including the intrinsic estimator method, can be found in Appendix G.

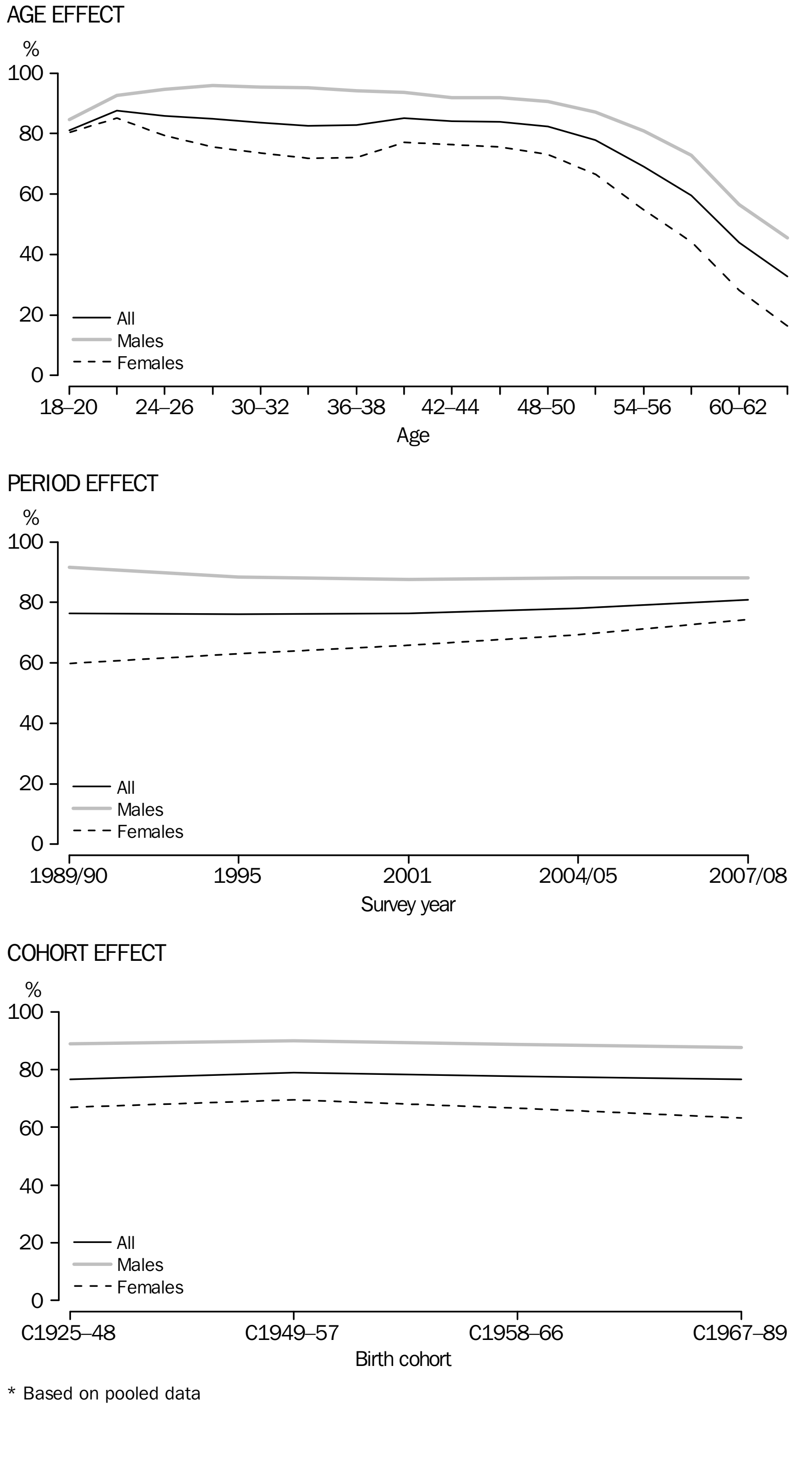
We initially applied the APC decomposition model using the intrinsic estimator method to the APC classification consisting of 16 three-year age groups, 5 single-year survey periods and 22 three-year birth cohorts. While reliable estimates for the age and period coefficients could be obtained, it appeared that the three-year birth cohorts used here were too narrow to detect any significant differences in labour force participation between different cohorts. The very old and the very young cohorts in our data were observed at one point in time only, which was during the period of their exit from and entry into the labour force, respectively, when participation was low. This coupled with the very small sample sizes for these cohorts made it difficult to undertake valid comparisons across cohorts and make any inferences about possible cohort effects. A broader definition of cohorts, with a longer period separating them apart, could possibly help capture any cohort effect on labour force participation, if it existed.

With this objective in mind, we redefined our cohorts more broadly to reflect four distinct periods or generations. These four cohorts were: those born during 1925–1948, 1949–1957, 1958–1966 and 1967–1989. These four groups were chosen on the basis that members in each of these groups could share some similar characteristics and experiences. Furthermore, the use of these four cohorts made them more or less equally represented in the sample. Although this grouping did not exactly match the distinct pre-baby boomer, baby-boomer and post baby-boomer periods, they could be said to roughly represent these generations. Given that the baby-boomer group itself was quite large in the age group that we were looking at, that is, 18–64 years old, this group was further split into two subgroups – those born immediately after WWII (1949–1957) and those born in the 1960s period (1958–1966). The last group (1967–1989) could be seen to represent the post-baby boomer generation.

20 For example in any period if we know a person's age then we can calculate his cohort group or birth year and in any period if we know a person's birth year then we can compute his age.

Figure 6.2 presents estimates of APC[21] effects with the above broader cohort groupings.

6.2 Age, period and cohort effects with broader cohort groupings (%)*



* Based on pooled data

21 Note that with this broader grouping of cohorts, the collinearity problem is no longer an issue. With the design matrix being full rank now normal matrix inversion can be used to derive the model coefficients rather than the need to use the intrinsic estimator method. However, implicit in the broadening of the cohort definition is that we are applying the 'coefficients constraints' solution to solve the identification problem. Here we are making several of the three-year cohorts the same by collapsing them together.

For exposition, we have combined the average underlying log-odds participation (captured by the μ parameter) separately with the individual age (α_i), period (β_j) and cohort (γ_k) effects and applied the log transformation to recover fitted labour force participation rates.[22]

The results showed strong age effect on labour force participation for both males and females. For females, it followed the expected trend, increasing between the ages of 18–23 years, then declining during the child-bearing years, then rising during early forties before declining after the age of 55 years. For males, there was a large increase in labour force participation from the age of 18 years to early thirties. The participation rate stabilised between the mid-thirties to late forties, and after the age of 55 years, it began to fall significantly like in the case of females. For all age groups, males' participation rate remained consistently higher than that of females.

There also appeared to be a strong period effect on labour force participation, especially for females. Their participation showed a steady rise during the study period. For males, it appeared to decline between 1989/90 to 2001 and it then stabilised. In all periods, males' participation rate still remained above that of females.

With respect to cohorts, not much difference in labour force participation across successive generations of males could be seen even with this broader definition of cohorts. For females, some decline in labour force participation for the younger generation was observed compared to the earlier generations, but it was difficult to say whether this was significant and what could possibly explain the decline. From the above analysis, it appeared that age and period effects were the more dominant drivers of labour force participation during the period under study, while cohort effects were less important.

It may be noted that the APC analysis conducted above looked only at the effect of these three variables on labour force participation. As labour force participation could also be influenced by other factors, the inclusion of these factors in the model could change the relative magnitude of the effects of the APC variables. In the next section, we incorporate other factors, in addition to APC, within a broader modelling framework in order to examine their relative influence on labour force participation. This would also allow us to assess the statistical significance of the APC variables which was not possible in the graphical analysis presented above. However, the analysis undertaken above in terms of alternative specifications of cohorts, graphical results and the identification of the relative importance of each of the APC variables on labour force participation has helped to inform and guide the subsequent modelling work undertaken in the next section.

---

22 For instance, with the age effect, the fitted participation rate would be given by $\frac{\exp(\mu + \alpha_i)}{1 + \exp(\mu + \alpha_i)} \times 100$.

7. RESULTS FROM MODEL-BASED ANALYSIS

This section presents results from estimation of logistic regression models for factors associated with labour force participation for persons aged 18–64 years. In such an analysis, our core interest was investigating whether health status has an association with labour force participation after controlling for age, period, cohort, and other relevant demographic and socioeconomic variables.

The dependent variable in the model was a binary variable which took a value of 1 if the person was in the labour force and 0 if not in the labour force. The explanatory variables were those identified earlier in Section 2 based on theoretical and empirical works and for which relevant data were available in our dataset. The variables included in the model were the health status, marital status, presence of dependent children, non-school qualifications , country of birth, proficiency in spoken English, Indigenous status and location; whether an individual was living in capital city or in a regional area.

We also included age, period and cohort variables in our model to test whether they had statistically significant effects on labour force participation that were peculiar to the age of a person, the period or year in which the survey was conducted[23] and the cohort group to which the person belonged to. The age of a person was entered in the model as a quadratic variable (i.e. age plus age squared).[24] The period variable was defined as dummy variables representing each of the five survey years. The broadly defined four cohorts, as used in the previous section, were included in the model as dummy variables. Entering the APC variables in the model in this way removed the collinearity problem, and gave a better model fit. This study considered the main effects of the above explanatory variables to avoid complications in the interpretation of the results that might arise from the inclusion of too many interaction terms.

Preliminary analysis suggested that males and females in the sample were significantly different in terms of various demographic and socioeconomic factors including labour force participation. On account of this, we estimated separate logistic regression models for males and females to reflect correct relationships.

---

23 In addition to accounting for the labour market conditions or business cycle, the period variable could also be seen to capture any significant differences in survey design over time, if they existed.

24 We tried several specifications for the age variable in the model (categorical, continuous) and found that age as a quadratic variable gave a better model fit than the other alternatives.

7.1 Logistic regression model estimates for association between self-assessed health status and labour force participation

	Males			*Females*		
Parameter	*Estimate*		*Marginal effect*	*Estimate*		*Marginal effect*
Intercept	–0.9339	***		–0.7964	***	
Has good–excellent health	0.0000					
Has poor/fair health	–1.4535	***	–0.099	–0.8594	***	–0.162
Has no children	0.0000					
Has children	0.0534		0.002	–1.2747	***	–0.237
Age (years)	0.2272	***	–0.002	0.1451	***	–0.007
Age squared	–0.0035	***		–0.00247	***	
Not married	0.0000					
Married	0.7700	***	0.033	–0.2757	***	–0.044
Non-Indigenous origin	0.0000					
Indigenous origin	–0.5360	***	–0.028	–0.6708	***	–0.128
Good proficiency in spoken English	0.0000					
Poor proficiency in spoken English	–0.8070	***	–0.047	–1.0041	***	–0.204
Overseas born	0.0000					
Australian born	0.3413	***	0.015	0.2925	***	0.049
Lives outside capital city	0.0000					
Lives in capital city	0.0376		0.002	0.1699	***	0.028
Has no non-school qualification	0.0000					
Has certificate/diploma education	0.5027	***	0.020	0.6036	***	0.090
Has degree and above education	0.7250	***	0.024	1.0889	***	0.139
Period 1989/90	0.0000					
Period 1995	–0.3885	***	–0.018	–0.0639	*	–0.010
Period 2001	–0.4093	***	–0.019	0.5336	***	0.077
Period 2004/05	–0.3969	***	–0.018	0.5525	***	0.079
Period 2007/08	–0.3067	***	–0.014	0.7886	***	0.106
Cohort born 1925–48	0.0000					
Cohort born 1949–57	0.0356		0.001	0.0660		0.010
Cohort born 1958–66	–0.1712	*	–0.007	–0.0687		–0.011
Cohort born 1967–89	–0.2721	**	–0.012	–0.3041	**	–0.050
n	44,985			47,815		
Likelihood Ratio (p Value)	6,922		(<0.0001)	8,456		(<0.0001)
Deviance Value/DF (p Value)	0.97		(0.98)	1.28		(<0.0001)
Max-rescaled R^2	0.27			0.23		
% Concordant	82.5			75.2		

*, ** and *** indicate the coefficient is statistically significant at 10%, 5% and 1%, respectively.

The logistic regression models were estimated using cohort-based weights[25] and the estimation of the standard errors took into account the underlying survey design.[26] The estimation results are given in table 7.1 covering the parameter estimates for the explanatory variables, their statistical significance and marginal effects.

The model diagnostics are presented at the bottom of the table. Most test statistics – likelihood ratio, max-rescaled R^2 (above 0.2[27]) and percent concordance[28] (above 75%) – suggest that the model estimates provided a reasonably good fit to the data for both males and females. The deviance test, which measures the degree of variability in the outcome variable, showed some evidence of over-dispersion for females but not for males. Efforts to correct the over-dispersion by applying the deviance or Pearson methods of estimation of dispersion parameters resulted in little change to the standard errors and statistical significance of the parameter estimates, and thus our statistical inferences.[29] [30]

25 Cohort weights are derived for each sex by dividing the average population (over five years) for each cohort by the total sample size (over five years) for that cohort and then rescaling these weights back to sum to total sample size. Using cohort-based weights ensures that each cohort has the same unchanged weight irrespective of which year or period that cohort is from.

26 The *NHS* is based on a stratified cluster survey. As such, information on stratification and clustering had to be used to obtain correct estimates of the standard errors. PROC SURVEYLOGISTIC in SAS was used to estimate the model parameters as this allows for incorporation of survey design (that is, strata and clusters) in the estimation process. While PROC LOGISTIC gave identical estimates for the coefficients, it produced different and incorrect estimates of the standard errors, as it assumes a simple random sampling survey design. Standard errors determine whether a particular variable is significant or not under alternative methods of estimation.

27 In social science research and where the objective is to study relationships among variables rather than do predictions, values above 0.2, as a rule-of-thumb, are considered acceptable.

28 The percent concordance measures the proportion of correct outcomes predicted by the model compared to the observed outcome.

29 This was because the dispersion parameters were close to 1. Scholars in the area suggest that over-dispersion is possible if the deviance is at least twice the degrees of freedom (Lindsey, 1999).

30 We also considered the Hosmer-Lemeshow goodness-of-fit (HL-GOF) test, but this test statistic produced results opposite to the other tests discussed in the text, indicating that the models were not fitting the data well. Even when we examined alternative aspects of the models (e.g. treated age as a categorical variable, added a reasonable number of interaction terms and used alternative link functions), the significance of the HL-GOF test did not change. It may be noted that there have been criticisms of the HL-GOF test as the test results have been found to be sensitive to the number of groups used to compute the test statistic (standard ten groups versus other grouping) and large sample size, in which case even small discrepancies between a model's predicted and observed counts become significant (Allison, 2013; Feudtner *et al.*, 2009; Hosmer *et al.*, 1997; Karmer and Zimmerman, 2007). Some researchers note that the significance of HL-GOF test of a model does not necessarily mean that the model is not useful or suspect (Karmer and Zimmerman, 2007).

With the exception of age, all other variables in the model are categorical. The model coefficients and marginal effects[31] of each of the variables are expressed relative to their respective reference categories, which are indicated in the table by a zero coefficient for each of the variables. For instance, the marginal effect of health status is the difference between the probability of labour force participation of a person with fair or poor health and the probability of a person with good or better health, holding all other variables in the model at their mean values.

Scrutinising the individual explanatory variables, we observe that a large number of them appeared to be statistically significant and assumed the expected signs for both males and females. There were a lot of similar effects of the variables between males and females, but there were also some significant differences between the sexes in the signs and statistical significance of the coefficients of the variables and in the magnitudes of the marginal effects of some variables.

The self-assessed health variable had a negative coefficient for both males and females. This indicates that those individuals with poor or fair self-assessed health were less likely to participate in the labour force compared to those with good or better health. The negative association between self-assessed general health status and labour force participation appeared to be stronger for females than for males as indicated by their respective marginal effects. For an 'average' female with fair to poor health, her probability of participation in the labour force was 0.162 lower than that of an 'average' female with good to excellent health. The corresponding marginal effect for an 'average' male was 0.099.

The certificate /diploma and degree or above levels of non-school qualifications had positive and statistically significant coefficients. These suggest that those people with non-school qualification were more likely to be in the labour force in comparison to people with no non-school qualification. The likelihood of participation increased with the level of qualification for females as those with Bachelor degree or above qualification had a higher likelihood of participation in the labour force than those with certificate/diploma and no qualification. For an average female with Bachelor degree or above qualification, her probability of participation in the labour force was 0.139 higher than that of an average female with no non-school qualification. The corresponding figure for a female with certificate/diploma qualification was 0.09. While education also had a statistically significant association with males' participation in the labour force, it was less strong than in the case of females' suggesting that

---

31 In the logistic regression model the marginal effect measures the change in the probability of an event of interest, e.g. labour force participation, resulting from a change in a certain explanatory variable, while keeping all the other covariates constant. For continuous variables the marginal effect measures the change in the probability of the event resulting from a small or one unit change in the particular variable from its mean value, while for categorical variables it measures the change in the probability of the event as the variable changes from 0 to 1, holding, in both cases, all other variables at their mean values. It may be noted that keeping the values of all the other variables in the model at their mean values, while changing only the variable of interest, implies that the marginal effect is being computed for an "average" person or individual.

education appeared to have much more influence on females' likelihood of participation in the labour force than that of males.

Poor proficiency in spoken English had a significant negative effect for both males and females suggesting that people with difficulty in spoken English were less likely to participate in the labour force compared to those who were proficient in the language. The effect was much stronger for females than males, with the probability of participation in the labour force for an average female with poor proficiency in spoken English being 0.204 lower than that of an average female with good proficiency. The corresponding marginal effect for males was just 0.047.

Being of Indigenous origin significantly reduced the likelihood of participation in the labour force for both males and females compared to those from non-indigenous background, with the effect being much stronger for females than males. An average Indigenous female had 0.128 lower and an average Indigenous male had 0.028 lower probability of participation in the labour force compared to an average non-indigenous male and female, respectively.

Marital status had a significant effect on both sexes' labour force participation, but it had an opposite effect for males and females. Married males were more likely to participate in the labour force compared to unmarried males, while married females were less likely to participate in the labour force compared to unmarried females. The probability of participation in the labour force for an average married male was 0.033 higher and for an average female 0.044 lower than their average non-married counterparts, respectively.

Age appeared to have a statistically significant influence on labour force participation for both males and females. The coefficient for the age and age squared variable in the model suggested that there was a curvilinear relationship between age and labour force participation for both the sexes, with participation initially increasing with age up to a certain age before beginning to decline. The marginal effect for age, which here was computed as a change in labour force participation resulting from a year increase in age from the mean age[32], showed that for an average male this reduced participation in the labour force by 0.002 and for an average female by 0.007.

The presence of dependent children had a strong negative effect on labour force participation for females, while it was not statistically significant for males, although it assumed the expected positive sign. An average female with a dependent child had a 0.237 lower probability of participation in labour force compared to an average female with no dependent children. The negative effect of this variable for females suggests that child caring responsibility for women was an important consideration in their decision to participate in the labour force.

32 The mean age for males and females in the sample was 37.6 and 37.7 years respectively.

Country of birth had a significant positive influence on both males and females participation in the labour, with those born in Australia more likely to be in the labour force compared to those born overseas. The location variable had a positive and statistically significant effect on females' participation in the labour force, but it was not significant for males, although it had the expected positive sign. This result suggests that females who lived in capital cities were significantly more likely to be in the labour force compared to those in regional areas.

The variable representing survey year, that is period, had negative coefficients for males compared with the reference period of 1989/90. This suggests that time had a negative effect on males' likelihood of participation in the labour force. For females, the period variable had a positive effect, with the exception of 1995, compared with the reference period. This suggests that time had favourably influenced females' likelihood of participation in the labour force, after controlling for other variables. These results were consistent with the results from APC analysis discussed in the previous section.

Cohort effects which appeared largely non-existent or inconclusive in the APC decomposition analysis showed that there was some evidence of cohort effects on labour force participation, after controlling for other available variables. The younger cohorts for both the sexes showed lower likelihood of participation in the labour force compared to their oldest counterparts. Increased educational opportunities and expectations for the younger generation to get more education compared to the older generations at the same stage of their life cycle could possibly explain the lower likelihood of participation in the labour force of the younger cohort compared to the older cohorts.

In order to further examine the association between individual long-term health conditions and labour force participation, we re-estimated Equation (3) by replacing the self-assessed general health status with the prevalence of five chronic diseases, while we kept all the other predictors in the previous model unchanged. We also re-estimated Equation (3) by replacing the general health status with another health outcome variable defined as having at least one of the five major health conditions or not. Full results of these two sets of models are presented in Appendixes H.1 and H.2, respectively. The likelihood ratio tests' χ^2 and p-values and other measures of association of these estimated models suggest that there were evidences that at least one of the independent variables contributed to the prediction of females' and males' participation in the labour force.

Partial model results for each of the five chronic diseases and their comorbidity variable are given in table 7.2 as Model 2 and Model 3, respectively, along with comparative estimates for the self-assessed general health status variable from the previous model (Model 1).

7.2 Partial logistic model results of labour force participation for alternative measures of health outcomes

	Males			*Females*		
	Estimate		*Marginal effect*	*Estimate*		*Marginal effect*
Model 1						
Has poor/fair health	–1.4535	***	–0.099	–0.8594	***	–0.162
Model 2						
Has arthritis	–0.7239	***	–0.044	–0.3672	***	–0.065
Has asthma	–0.2416	***	–0.012	–0.1418	***	–0.024
Has cancer	–0.4679	***	–0.027	–0.2466	***	–0.043
Has diabetes	–0.5956	***	–0.036	–0.6545	***	–0.126
Has heart disease	–0.7748	***	–0.051	–0.3629	***	–0.065
Model 3						
Has at least one of the five disease conditions	–0.6372	***	–0.035	–0.2944	***	–0.050

*** indicates the coefficient is statistically significant at 1%.

The five individual diseases assumed statistically significant negative coefficients in both the males' and females' model estimates suggesting that individuals with each of these conditions were less likely to participate in the labour force. These empirical results conform the findings of Cai and Cong (2009). Looking into their marginal effects suggest that the presence of heart disease, arthritis, diabetes, cancer and asthma, in that order, had statistically significant impact on males' likelihood of participation in the labour force. For females, diabetes was found to have a much stronger influence on their likelihood of participation in the labour force, followed by arthritis, heart disease, cancer and asthma, in that order. For each of the chronic diseases, the marginal effects were higher for females than for males.

The third alternative measure of health outcome was a binary variable which took a value of 1 if one or more of the five disease conditions were prevalent and 0 otherwise. Results from these model estimates were similar to results from earlier model estimates. The probability of an average male with at least one of the major long-term health conditions declined by 0.035 compared with an average male with no condition. The probability of an average female with at least one of the chronic diseases declined by 0.05 compared with an average female with no condition. The empirical results from the above three model estimates suggest that there was a strong negative association between poor health and participation in the labour force, which was robust to health outcome measures used in this study.

The above model results were discussed in terms of model coefficients and marginal effects. The marginal effects were discussed in terms of the effect of each variable on labour force participation for an average person. The probabilities and the marginal effects can also be computed for a person with any set of characteristics. Here we demonstrate the effect of a change in a person's health status on his/her probability of labour force participation for a 40 year old person with all other characteristics set to the base case or the reference person in Model 1, that is., values for all other variables are set to zero.[33] The predicted probability of participation in the labour force of such a reference male with good to excellent health was 0.93, which declined to 0.76 for a male with exactly the same characteristics, but with fair to poor health. For a reference female with good to excellent health, her probability of participation in the labour force was 0.74, which declined to 0.55 for the same female with fair to poor health. The slightly larger decline in the probability for females (–0.19) relative to males (–0.17) indicates that the marginal effect of health status was greater for females than males.[34] These results confirm earlier results in that self-assessed health status had a significant effect on the likelihood of females and males participation in the labour force.

33 As such the base or reference person for either male or female in this case refers to a person with the following characteristics: has no children, non-married, non-indigenous, has good spoken English, overseas-born, has no non-school qualification, observed in survey period 1989 and belonging to the cohort group born during 1925–48.

34 It may be noted that the marginal effects shown here are different from the marginal effects shown earlier because the marginal effects here are computed for males and females with the characteristics as defined above compared to the 'average' characteristics or mean values used in computing the marginal effects earlier.

8. CONCLUSIONS

This study examined the relationship between self-assessed health status and labour force participation using pooled unit-record data from ABS's five consecutive NHSs. It presented a descriptive analysis of labour force participation, health and selected other demographic and socioeconomic factors. A decomposition of age, period and cohort effects was carried to examine their separate effects on labour force participation. A logistic regression model was used to investigate the association between health status and participation in the labour force controlling for age, period, cohorts and a number of other demographic and socioeconomic variables.

The study showed that although a large majority of the Australian population enjoys good to excellent health status there were population sub-groups that were characterised by fair or poor self-assessed health. There was also an upward trend in some of the long-term health conditions, such as arthritis, asthma, diabetes and heart disease, while cancer stayed somewhat stable over the period under study.

The logistic regression analysis suggested that health status was an important factor associated with males' and females' participation in the labour force. People with fair or poor self-assessed health status had significantly lower likelihood of participation in the labour force compared to those with good or better health. The probability of participation in the labour force of an 'average' male (female) with fair or poor health was 0.099 (0.162) lower than that of an 'average' male (female) with good or better health, holding the other variables constant at their mean values. The marginal effect of health status on female's participation in the labour force was just second to that of the presence of dependent kid(s) that lowered an average female's probability of participation by 0.237 compared to the same female with no dependent kid(s).

Furthermore, major long-term health conditions were also found to have a negative relationship with both males' and females' participation in the labour force. The presence of heart disease (–0.0508), arthritis (–0.0442), diabetes (–0.0361), cancer (–0.0268), and asthma (–0.0123), with figures in the brackets being marginal effects, had statistically significant adverse influences on males' probability of participation in the labour force. For females, diabetes (–0.1261), arthritis (–0.0647), heart disease (–0.0653), cancer (–0.0431) and asthma (–0.0239) had statistically significant influence on their probability of participation in the labour force. The strong negative association between poor health status and participation in the labour force was found to be robust to the alternative health outcome indicators used in this study. From the marginal effects, adverse health status appeared to have a greater negative association with females' participation in the labour force compared with that of males.

In relation to the other variables, age appeared to play an important role in labour force participation with participation increasing initially with age for both the sexes before beginning to decline. By age 55–64 around two-thirds of males were still in the labour force compared to only a third of females.

Marital status, non-school qualifications, proficiency in spoken English, and Indigenous status were other variables that were found to significantly influence the likelihood of males' and females' participation in the labour force, while presence of dependent children and location influenced that of females' only. Presence of dependent children, proficiency in spoken English, education and indigenous origin had particularly strong influences on females' participation in the labour force.

Period appeared to have largely positive effect on females' and negative effect on males' participation in the labour force. After controlling for other variables, some cohort effect was observed for both males and females, as younger cohorts exhibited lower labour force participation compared to their older counterparts.

Given the significant negative association between health status and labour force participation, after controlling for diverse factors, identification of factors that influence self-assessed health status would be very useful in informed public debate and decision-making. The rich pooled dataset that has been created as part of this study can be used to address other key research questions relating to individual long-term health conditions, health risk factors, sick leave, doctor visits, medications etc. In future research efforts, the addition of data on selected variables from the Australian Health Survey to the pooled NHS dataset would expand the period and data coverage and add to the statistical strength of their analyses.

This study used individual level data as the basis of analysis with cohorts added as explanatory variables in the models to control for unobservable heterogeneity. Further work that uses cohorts as the basis of analysis could also be explored.

ACKNOWLEDGEMENTS

This paper was based on a research project called "National Health Survey Data Pooling and Analysis". The authors thank Greg Francis, Peter Rossiter, Peter Radisich, Damien Melksham, Anura Amarasinghe, Daniel Elazar and John Martin for their assistance while compiling and analysing the pooled data. We are grateful to Louise Gates, Michelle Ducat, Sue Phillips, Kirsty Leslie and Siu-Ming Tam for their useful comments on the analysis and the paper. We express our thanks to Ruel Abello for his guidance and useful comments and suggestions on the paper. Any errors in the paper remain the sole responsibility of the authors.

REFERENCES

Allison, P. (2001) *Why I Don't Trust the Hosmer-Lemeshow Test for Logistic Regression,* Statistical Horizons. <http://www.statisticalhorizons.com/hosmer-lemeshow>

Australian Bureau of Statistics (2001) *Measuring Wellbeing: Frameworks for Australian Social Statistics,* cat. no. 4160.0, ABS, Canberra. <http://www.abs.gov.au/ausstats/abs@.nsf/mf/4160.0>

—— (2004) *An Overview of Age Standardisation*, Information Paper, ABS, Canberra.

—— (2006) *National Health Survey: Summary of Results, 2004–05*, cat. no. 4364.0, ABS, Canberra. <http://www.abs.gov.au/ausstats/abs@.nsf/mf/4364.0>

—— (2008) *Year Book Australia, 2008*, cat. no. 1301.0, ABS, Canberra. <http://www.abs.gov.au/ausstats/abs@.nsf/mf/1301.0>

—— (2009) *National Health Survey: Summary of Results, 2007–08*, cat. no. 4364.0, ABS, Canberra. <http://www.abs.gov.au/ausstats/abs@.nsf/mf/4364.0>

—— (2010) *Year Book Australia, 2009–10*, cat. no. 1301.0, ABS, Canberra. <http://www.abs.gov.au/ausstats/abs@.nsf/mf/1301.0>

Australian Institute of Health and Welfare (2010) *Australia's Health 2010*, Australia's Health Sseries no. 12, Cat. no. AUS 122, AIHW, Canberra. <https://www.aihw.gov.au/WorkArea/DownloadAsset.aspx?id=6442452962>

Bartley, M. and Owen, C. (1996) "Relation between Socioeconomic Status, Employment and Health during Economic Change, 1973–93", *British Medical Journal*, 313(7055), pp. 445–449. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2351866/>

Bartley, M. (1994) "Unemployment and Ill Health: Understanding the Relationship", *Journal of Epidemiology and Community Health*, 48(4), pp. 333–337. <http://jech.bmj.com/content/48/4/333.abstract>

Bound, J.; Schoenbaum, M.; Stinebrickner, T.R. and Waidmann, T. (1998) "The Dynamic Effects of Health on the Labor Force Transitions of Older Workers", *Labour Economics*, 6(2), pp. 179–202. <http://www.sciencedirect.com/science/article/pii/S0927537199000159>

Benjamin, D.; Gunderson, M. and Riddell, C. (2002) *Introduction to Labour Market Economics*, McGraw-Hill Higher Education.

Brown, H.S. III; Pagan, J.A. and Bastida, E. (2005) "The Impact of Diabetes on Employment: Genetic IVs in a Bivariate Probit", *Health Economics*, 14(5); pp. 537–544.
<http://search.proquest.com/docview/232038371/abstract?source=fedsrch&accountid=26148>

Cai, L. and Cong, C. (2009) "Effects of Health and Chronic Diseases on Labour Force Participation of Older Working-Age Australians", *Australian Economic Papers*, 48(2), pp. 166–182.
<http://onlinelibrary.wiley.com/doi/10.1111/j.1467-8454.2009.00365.x/abstract>

Cai, L. and Kalb, G. (2004) *Health Status and Labour Force Participation: Evidence from the HILDA Data*, Working Paper Series No. 4/04, Melbourne Institute of Applied Economic and Social Research, Melbourne.

Cai, L. and Kalb, G. (2006) "Health Status and Labour Force Participation: Evidence from Australia", *Health Economics,* 15(3), pp. 241–261.
<http://onlinelibrary.wiley.com/doi/10.1002/hec.1053/abstract>

Chirikos, T.N. (1993) "The Relationship between Health and Labor Market Status", *Annual Review of Public Health*, 14, pp. 293–312.
<http://www.annualreviews.org/doi/abs/10.1146/annurev.pu.14.050193.001453>

Commonwealth Treasury (2010) *Intergenerational Report 2010: Australia to 2050: Future Challenges*,
<http://archive.treasury.gov.au/igr/igr2010/>

Contreras, D.; de Mello, L. and Puentes, E. (2011) "The Determinants of Labour Force Participation and Employment in Chile", *Applied Economics*, 43(21), pp. 2765–2776.
<http://www.tandfonline.com/doi/abs/10.1080/00036840903373303>

Crossley, T. and Kennedy, S. (2002) "The Reliability of Self-Assessed Health Status", *Journal of Health Economics*, 21(4), pp. 643–658.
<http://www.ncbi.nlm.nih.gov/pubmed/12146595>

Feudtner, C.; Hexem, K.R.; Shabbout, M.; Feinstein, J.A.; Sochalski, J. and Silber, J.H. (2009) "Prediction of Pediatric Death in the Year after Hospitalization: A Population-Level Retrospective Cohort Study", *Journal of Palliative Medicine*, 12(2), pp. 160–169.
<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2656437>

Fu, W.J. (2008) "A Smoothing Cohort Model in Age–Period–Cohort Analysis with Applications to Homicide Arrest Rates and Lung Cancer Mortality Rates", *Sociological Methods and Research*, 36(3), pp. 327–361.
<http://smr.sagepub.com/content/36/3/327.abstract>

Glenn, N.D. (2003) "Distinguishing Age, Period and Cohort Effects", in J.T. Mortimer and M.J. Shanahan (eds), *Handbook of the Life Course*, pp. 465–476, Kluwer Academic/Plenum, New York.

Grogan, L. and Koka, K. (2010) "Young Children and Women's Labour Force Participation in Russia, 1992–2004", *Economics of Transition*, 18(4), pp. 715–739. <http://onlinelibrary.wiley.com/doi/10.1111/j.1468-0351.2010.00390.x/abstract>

Gray, M.C. and Hunter, B.H. (2002) "A Cohort Analysis of the Determinants of Employment and Labour Force Participation: Indigenous and non-Indigenous Australians, 1981 to 1996", *Australian Economic Review*, 35(4), pp. 391–404. <http://search.proquest.com/docview/208190768/abstract?source=fedsrch&accountid=26148>

Heuer, C. (1997) "Modelling of Time Trends and Interactions in Vital Rates using Restricted Regression Splines", *Biometrics*, 53, pp. 161–177.

Holford, T.R. (1991) "Understanding the Effects of Age, Period and Cohort on Incidence and Mortality Rates", *Annual Review of Public Health*, 12, pp. 425–457. <http://www.annualreviews.org/doi/abs/10.1146/annurev.pu.12.050191.002233>

Hosmer, D.W.; Hosmer, T.; Le Cessie, S. and Lemeshow, S. (1997) "A Comparison of Goodness-Of-Fit Tests for the Logistic Regression Model", *Statistics in Medicine*, 16(9), pp. 965–980. <http://www.stat.duke.edu/~zo2/dropbox/goflogistic.pdf>

Hotchkiss, J.L. and Robertson, J.C. (2012) "Asymmetric Labour Force Participation Decisions", *Applied Economics*, 44(16), pp. 2065–2073. <http://www.tandfonline.com/doi/abs/10.1080/00036846.2011.558480>

Hotchkiss, J.L. and Robertson, J.C. (2006) *Asymmetric Labour Force Participation Decisions over the Business Cycle: Evidence from U.S. Microdata*, Working Paper, Federal Reserve Bank of Atlanta, No. 2006-8. <http://www.econstor.eu/bitstream/10419/70716/1/572282435.pdf>

Jones, M.K. and Latreille, P.L. (2009) "Disability, Health and the Labour Market: Evidence from the Welsh Health Survey", *Local Economy*, 24(3), pp. 192–210. <http://lec.sagepub.com/content/24/3/192.abstract>

Jose, A.; Ravindiran, R. and Abello, R. (2004) *Health Status – Labour Force Non-Participation Nexus: Evidenced from Pooled NHS Data*, Paper presented to the 33rd Annual Conference of Economists, University of Sydney, 27–30 September.

Kakinaka, M. and Miyamoto, H. (2012) "Unemployment and Labour Force Participation in Japan", *Applied Economics Letters*, 19(11), pp. 1039–1043. <http://www.tandfonline.com/doi/abs/10.1080/13504851.2011.613742>

Kalwij, A.S. and Vermeulen F.M.P. (2008) "Health and Labour Force Participation of Older People in Europe: What do Objective Health Indicators Add to the Analysis?", *Health Economics*, 17(5), pp. 619–638. <http://search.proquest.com/docview/232070013/abstract?source=fedsrch&accountid=26148>

Karmer, A. and Zimmerman, J.E. (2007) "Assessing the Calibration of Mortality Benchmarks in Critical Care: The Hosmer-Lemeshow Test Revisted", *Critical Care Medicine*, 35(9), pp. 2052–2056. <http://www.ncbi.nlm.nih.gov/pubmed/17568333>

Kumar, A. and Chessman, J. (2009) *Report on SEW Data Pooling Feasibility Study SEW Data Pooling Project – Improved Youth Education Participation and Attainment Rates Through Alternative Data Pooling Option*, unpublished report, Analytical Services Branch, Australian Bureau of Statistics, Canberra.

Kumar, A.; Rossiter, P. and Olczyk, A. (2009) "Children's Participation in Organised Sporting Activity", *Methodology Research Papers*, cat. no. 1351.0.55.028, Australian Bureau of Statistics, Canberra. <http://www.abs.gov.au/ausstats/abs@.nsf/mf/1351.0.55.028>

Lindsey, J.K. (1999) "On the Use of Corrections for Overdispersion," *Applied Statistics*, 48(4), pp. 553-561. <http://www.stat.sc.edu/~hansont/stat770/Lindsey1999overdispersion.pdf>

Mathers, C.D. and Schofield, D.J. (1998) "The Health Consequences of Unemployment: The Evidence", *Medical Journal of Australia*, 168(4), pp. 178–182. <https://www.mja.com.au/journal/1998/168/4/health-consequences-unemployment-evidence>

Mathers, C. (1994) "Health Differentials among Older Australians", *Health Monitoring Series No. 2*, Cat. no. AIHW 267, Australian Institute of Health and Welfare, Canberra. <http://www.aihw.gov.au/publication-detail/?id=6442466664>

McKenzie, D.J. (2005) *Disentangling Age, Cohort, and Time Effects in the Additive Model*, The World Bank, Washington, D.C. <http://siteresources.worldbank.org/DEC/Resources/obesmckenziefinal.pdf>

O'Brien, R.M. (2000) "Age Period Cohort Characteristic Models", *Social Science Research*, 29(1), pp. 123–139.

Nam, K. (2010) "The Effects of Having Children on Women's Labour Force Participation in Korea: An Analysis Using Instrumental Variables", *Review of Labour Economics and Industrial Relations*, 24(3), pp. 333–356. <http://onlinelibrary.wiley.com/doi/10.1111/j.1467-9914.2010.00487.x/abstract>

Oshio, T.; Oishi, A.S. and Shimizutani, S. (2011) "Social Security Reforms and Labour Force Participation of the Elderly in Japan", *Japanese Economic Review*, 62(2), pp. 248–271.
<http://onlinelibrary.wiley.com/doi/10.1111/j.1468-5876.2009.00505.x/abstract>

Prowse, V. (2009) "Why do People Work? Why do Some People Work While Others do not?", *Economic Review (UK)*, 27(1), p. 11.

Ravindiran, R.; Rawnsley, T. and Jose, A. (2002) *A Cohort Analysis of Unemployment and Participation Rates for Australia*, unpublished report, Analytical Services Branch, Australian Bureau of Statistics, Canberra.

Simon, J.G.; De Boer, J.B.; Joung, I.M.A.; Bosma, H. and Mackenbach, J.P. (2005) "How is your Health in General? A Qualitative Study on Self-Assessed Health", *European Journal of Public Health*, 15(2), pp. 200–208.
<http://eurpub.oxfordjournals.org/content/15/2/200.short>

Stronks, K.; van de Mheen, H.; van den Bos, J. and Mackenbach, J.P. (1997) "The Interrelationship between Income, Health and Employment Status", *International Journal of Epidemiology*, 26(3), pp. 592–600.
<http://ije.oxfordjournals.org/content/26/3/592.abstract>

World Health Organization (1946) *WHO Definition of Health*, Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, 19–22 June, 1946; signed on 22 July 1946 by the representatives of 61 States (Official Records of the World Health Organization, no. 2, p. 100) and entered into force on 7 April 1948.
<http://www.who.int/about/definition/en/print.html>

Yang, Y.; Schulhofer-Wohl, S.; Fu, W.J. and Land, K.C. (2008) "The Intrinsic Estimator for Age–Period–Cohort Analysis: What It Is and How to Use It", *American Journal of Sociology*, 113(6), pp. 1697–1736.

Wagstaff , A. and Van Doorslaer, E. (1994) "Measuring Inequalities in Health in the Presence of Multiple-Category Morbidity Indicators", *Health Economics*, 3(4), pp. 281–289.
<http://onlinelibrary.wiley.com/doi/10.1002/hec.4730030409/abstract>

Wilson, S.H. and Walker, G.M. (1993) "Unemployment and Health: A Review", *Public Health*, 107(3), pp. 153–162.

(All reported URLs accessed 6 February 2014)

APPENDIXES

A. CONSTRUCTION OF THE NHS POOLED DATASET

A pooled cross-sectional and time-series dataset (NHS Pooled Dataset) was created by pooling data from five repeated National Health Surveys (NHS) (NHS1989/90, NHS1995, NHS2001, NHS2004/05, NHS2007/08). It contains unit or person-level records sorted by survey year with unit records for NHS89 appearing first and NHS0708 last. The pooled dataset covers the population 0 years and above. A subset of this dataset was also created covering the population aged 18–64 years which was used for this research project. Additional variables were created on this dataset for descriptive analysis and modelling purposes. This section briefly describes the steps and processes involved in creating the pooled dataset, including discussion on data comparability issues and the key variables available on the dataset.

The main steps involved in creating the pooled dataset included the following: getting access to the relevant files for each of the survey years; identifying the required common variables from each survey; checking for comparability/consistency of these variables across the five surveys; harmonising, recoding, recreating and renaming variables across the five surveys; merging data from each of these five surveys using survey year and household ID to create the pooled dataset; and creating new variables on the pooled dataset as required.

While a large number of the required variables were obtained from the individual level files for each survey, some variables had to be obtained from the household level files e.g. household size, household income, SEIFA and for some surveys disease conditions had to be obtained from separate disease conditions files, e.g. for NHS2004/05 and NHS2007/08.

Assessing comparability and consistency across the surveys formed an important part of the data compilation stage. With the five surveys being almost twenty years apart, there was bound to be some conceptual, methodological and/or classification differences between the surveys. However, given the repeated nature of these surveys they were found to be comparable in many aspects, such as survey design, sample selection, scope, coverage, response rate, survey method, etc. as shown in Appendix B.[35] In instances where there were differences, various assessments (ABS, 2006, Jose *et al.*, 2004) have suggested that the results are broadly comparable across surveys. In compiling the common variables in the pooled dataset, we ensured that the variables were as consistent as possible across the surveys. For example, there was some

35 There was some difference in the reference period across the surveys but with each survey conducted between 10–12 months the chances of any differences across surveys resulting from seasonal differences should be minimal.

difference in the definition of heart disease between the first two surveys and the last three surveys which had a broader definition for heart disease compared to the earlier surveys. To ensure comparability of this variable across the five surveys the disease codes applicable to the earlier surveys were applied to the latter surveys to derive a consistent and comparable variable. Similarly, in compiling several other variables, an assessment was made across the surveys for differences in the definitions of these variables, differences in codes used for them, their measurements, whether they were continuous or categorical, if categorical, number of categories and cut-offs, differences in time frames used for the variables before compiling the variable. Where there were differences in categories, we collapsed them to a minimum number to make them consistent and comparable across the surveys, e.g. self-assessed health status[36], labour force status, highest non-school qualification, and income.

For some variables, there were differences in questionnaire wording or information collected and some assumptions had to be made to derive a comparable variable. For instance, in the 1989/90 and the 1995 surveys the question in relation to highest school qualification was the year left school, while in the later surveys the question specifically asked highest year of school completed. Based on discussion with the client, it was decided that if a person had left school at age 17 or later then he/she was deemed to have completed year 12 education.

Some variables were available on some surveys only. For example, the SEIFA variable was available from 1995 onwards, while exercise, consumption of fruits and vegetables were not available in the earlier surveys. Appendix C provides a list of key variables on the pooled dataset, its availability and consistency across the surveys.

Some other issues that were encountered in compiling the common variables cross the five surveys included the following:

- Difference in time frame like “in the last three or five years” in 2007/08 compared to “in the last two weeks” that was used in the other surveys. Information on exercise number of days in one week in 2007/08 compared to number of times in two weeks in 2004/05.
- Difference in question wording like, whether taken days away from “work” versus “work/study” in last two weeks.
- Differences in the measurement of income, like weekly personal cash income in 2001, 2004/05 and 2007/08 surveys compared to gross annual personal cash income in the 1989/90 and 1995 surveys.

36 For example for the self-reported health status variable it was available as a four-category variable (excellent, good, fair, poor) in NHS 1989 but was available as a five-category variable from NHS 1995 onwards. It was not possible to compile a four-category variable across all the five surveys as the good to excellent categories in the first survey were not aligning with these categories in the last four surveys. Hence we created a two-category variable (good to excellent, poor to fair) that was found to be comparable across all the surveys.

An issue of importance for the pooled dataset is what sample weights to use for deriving required statistics like counts/frequencies or proportions from the pooled sample. We can use individual year sample weights if we want to derive counts and proportions for single years but we need some pooled weights to derive counts and proportions from the pooled data. For the pooled dataset, we need to compute some modified sample weights that add to some meaningful population over the pooled period.[37] While there are several alternative methods for deriving the pooled weights, there does not seem to be any consensus as to an ideal method to be used. For our purposes, the method we have used involved adjusting each survey's original sample weight to reflect its relative sample size and population in the pooled dataset. More details on this method can be found in Kumar *et al.* (2009).

37 Note we cannot use the individual year weights to compute the total for the whole sample for variables of interest as this will give a number that will be larger than the population at any given year as it will simply add up the weights from the five surveys.

B. COMPARISON OF THE FIVE NATIONAL HEALTH SURVEYS

Survey information	*1989/90*	*1995*	*2001*	*2004/05*	*2007/08*
Sample size	54,241	53,828	26,863	25,906	20,788
Response rate	96%	97%	87%$	89%	91%
Survey design	Multistage area sampling	Multistage area sampling	Multistage area sampling	Multistage area sampling	Multistage area samplings
Scope	Covers population 0+ in private dwellings.	Covers population 0+ in private dwellings.	Covers population 0+ in private dwellings.	Covers population 0+ in private dwellings.	Covers population 0+ in private dwellings.
Coverage	Covers rural and urban areas across all states and territories but excludes very remote areas.	Covers rural and urban areas across all states and territories but excludes very remote areas.	Covers rural and urban areas across all states and territories but excludes very remote areas.	Covers rural and urban areas across all states and territories but excludes very remote areas.	Covers rural and urban areas across all states and territories but excludes very remote areas.
Sampling unit	Household	Household	Household	Household	Household
Sample selection	1 adult 18+ and 1 child (where applicable) selected for interview	1 adult 18+ and 1 child (where applicable) selected for interview	1 adult 18+ and 1 child (where applicable) selected for interview	1 adult 18+ and 1 child (where applicable) selected for interview	1 adult 18+ and 1 child (where applicable) selected for interview
Interview method	Face-to-face personal interview	Face-to-face personal interview	Face-to-face personal interview	Face-to-face personal interview	Face-to-face personal interview
Reference period	12 month period October 1989 to September 1990	12 month period January 1995 to January 1996	10 month period February to November 2001	10 month period August 2004 to July 2005	11 month period August 2007 to June 2008.
Mode of survey	Personal interview	Personal interview	Personal interview	Personal interview (CAI)	Personal interview (CAI)
Questionnaire wording	Most questions based on standard ABS questionnaire wording but some differences in questions possible.	Most questions based on standard ABS questionnaire wording but some differences in questions possible.	Most questions based on standard ABS questionnaire wording but some differences in questions possible.	Most questions based on standard ABS questionnaire wording but some differences in questions possible.	Most questions based on standard ABS questionnaire wording but some differences in questions possible.
Weighting method	Initial weights based on probability of selection benchmarked to age by sex by area of residence population totals to derive final sample weights.	Initial weights based on probability of selection benchmarked to age by sex by area of residence population totals to derive final sample weights.	Initial weights based on probability of selection benchmarked to age by sex by area of residence population totals to derive final sample weights.	Initial weights based on probability of selection benchmarked to age by sex by area of residence population totals to derive final sample weights.	Initial weights based on probability of selection benchmarked to age by sex by area of residence population totals to derive final sample weights.

C. LIST OF MAIN VARIABLES AVAILABLE ON THE POOLED NHS DATASET

The main pooled dataset covers the population 0 plus. A subset of this data covering the population 18–64 was used for this research project. Additional variables were created on this dataset for descriptive analysis and modelling. A list of key variables available on the main 0+ dataset is presented below. The pooled dataset contains 181,626 unit records or observations with the following breakdown by survey year: NHS1989/90 – 54,241, NHS1995 – 53,828, NHS2001 – 26,863, NHS2004/05 – 25,906, NHS2007/08 – 20,788.

Variable	*1989/90*	*1995*	*2001*	*2004/05*	*2007/08*	*Status*
Survey year	√	√	√	√	√	Specific
Person ID	√	√	√	√	√	Consistency established
Household ID	√	√	√	√	√	Consistency established
Sample weight – Person level	√	√	√	√	√	Consistent except naming
Replicate sample weight – Person level	X	X	√	√	√	2001 has 30 replicate weights
Pooled Weight	√	√	√	√	√	Created
Age (single years, 0+)	√	√	√	√	√	Consistent
Household size	x	x	√	√	√	Consistent
Number of adults	x	x	√	√	√	Consistency established
Number of children <15 age	√	√	√	√	√	Consistency established
Gender (Sex)	√	√	√	√	√	Consistent
Marital status (5 Categories)	√	√	√	√	√	Consistency established
Marital status dummy	√	√	√	√	√	Created
Indigenous status	√	√	√	√	√	Consistency established
Main language spoken at home	X	√	√	√	√	Consistency established
English proficiency	√	√	√	√	√	Consistency established
Country of birth	√	√	√	√	√	Consistency established
Educational qualification	√	√	√	√	√	Consistency established
State	√	√	√	√	√	Consistent
Capital city	√	√	√	√	√	Consistency established
Section of state	X	X	√	√	√	Consistency established
Remoteness	X	X	√	√	√	Consistency established
SEIFA CD (decile)	X	√	√	√	√	Consistency established
Employment status	√	√	√	√	√	Consistency established
Labour force status	√	√	√	√	√	Consistency established
Employment type	√	√	√	√	√	Consistency established
Hours per week	√	√	√	√	√	Consistency established
Employment sector	X	√	√	√	√	Consistency established

Occupation	√	√	√	√	√	Consistency established
Weekly Income #	√	√	√	√	√	Consistency established
Non-wage income	√	√	√	√	√	Consistency established
Equivalised household income	X	√	X	√	√	Consistency established
Health status (5 categories) 1. Excellent 2. Very good 3. Good 4. Fair 5. Poor	X	√	√	√	√	Consistency established
Health status (4 categories) 1. Excellent 2. Good 3. Fair 4. Poor	√	X	X	X	X	Consistency established
Health status (2 categories) 1. Excellent /Very Good/Good 2. Fair/Poor	√	√	√	√	√	Consistency established
Asthma (dummy 0/1)	√	√	√	√	√	Derived
Diabetes (dummy 0/1)	√	√	√	√	√	Derived
Arthritis (dummy 0/1)	√	√	√	√	√	Derived
High blood Pressure (dummy 0/1)	√	√	√	√	√	Derived
Heart disease (dummy 0/1)	√	√	√	√	√	Derived & made consistent
Cancer (dummy 0/1)	√	√	√	√	√	derived
Number of medications	√	√	√	√	√	Consistency established
No. of chronic conditions	X	X	√	√	√	Consistency established
Health concession card (dummy 0/1)	√	√	√	√	√	Consistency established
Number of visits to GP	?	?	?	√	√	Consistency to be established
Number of visits to OHP	?	?	?	√	√	Consistency to be established
Physical activity level	X	√	√	√	√	Consistency established
Level of alcohol consumption	X	√	√	√	√	Consistency established
Consumption of fruits	X	X	√	√	√	Consistency established
Consumption of vegetables	X	X	√	√	√	Consistency established
Self-assessed body mass index	√	√	√	√	√	Consistency established
Self-assessed body Weight (kg)	√	√	√	√	√	Consistency established
Self-assessed Height (cm)	√	√	√	√	√	Consistency established
Cohort group (3 year birth cohorts)	√	√	√	√	√	Created
Cohort group (6 year birth cohorts)	√	√	√	√	√	Created
Cohort group (10 year birth cohorts)	√	√	√	√	√	Created
Cohort group (4 special birth cohorts) (Born 1925–48, Born 1949–57, Born 1958–66, Born 1967–89)	√	√	√	√	√	Created

√ – the variable is available

X – the variable is missing

? – yet to be established

– In 1989/90 and 1995, weekly income is derived by dividing annual income by 52 weeks

D. LABOUR FORCE PARTICIPATION BY AGE, PERIOD AND COHORT

Labour Force Participation by Age, Period and Cohort, Persons aged 18–64 years

	Period				
	1989/90	*1995*	*2001*	*2004/05*	*2007/08*
Age group			(%)		
18–20 years	81.5 (C16)	77.1 (C18)	79.7 (C20)	80.9 (C21)	82.5 (C22)
21–23 years	89.9 (C15)	85.2 (C17)	86.5 (C19)	86.2 (C20)	84.1 (C21)
24–26 years	85.7 (C14)	86.3 (C16)	87.2 (C18)	83.2 (C19)	84.8 (C20)
27–29 years	82.1 (C13)	83.8 (C15)	83.0 (C17)	86.7 (C18)	86.7 (C19)
30–32 years	82.8 (C12)	81.6 (C14)	81.0 (C16)	85.2 (C17)	84.8 (C18)
33–35 years	82.6 (C11)	81.9 (C13)	82.0 (C15)	81.1 (C16)	85.2 (C17)
36–38 years	84.3 (C10)	82.8 (C12)	82.6 (C14)	82.1 (C15)	83.0 (C16)
39–41 years	85.4 (C9)	85.1 (C11)	84.3 (C13)	84.7 (C14)	87.5 (C15)
42–44 years	84.8 (C8)	86.1 (C10)	83.9 (C12)	85.6 (C13)	82.7 (C14)
45–47 years	82.5 (C7)	83.5 (C9)	84.5 (C11)	84.4 (C12)	88.4 (C13)
48–50 years	78.5 (C6)	82.0 (C8)	80.0 (C10)	83.9 (C11)	86.5 (C12)
51–53 years	73.7 (C5)	76.0 (C7)	76.5 (C9)	81.0 (C10)	84.3 (C11)
54–56 years	63.3 (C4)	66.8 (C6)	66.6 (C8)	72.4 (C9)	75.5 (C10)
57–59 years	51.2 (C3)	54.2 (C5)	58.6 (C7)	63.1 (C8)	63.9 (C9)
60–62 years	37.3 (C2)	36.5 (C4)	40.4 (C6)	46.5 (C7)	53.9 (C8)
63–64 years	24.8 (C1)	27.7 (C3)	29.8 (C5)	34.2 (C6)	43.7 (C7)

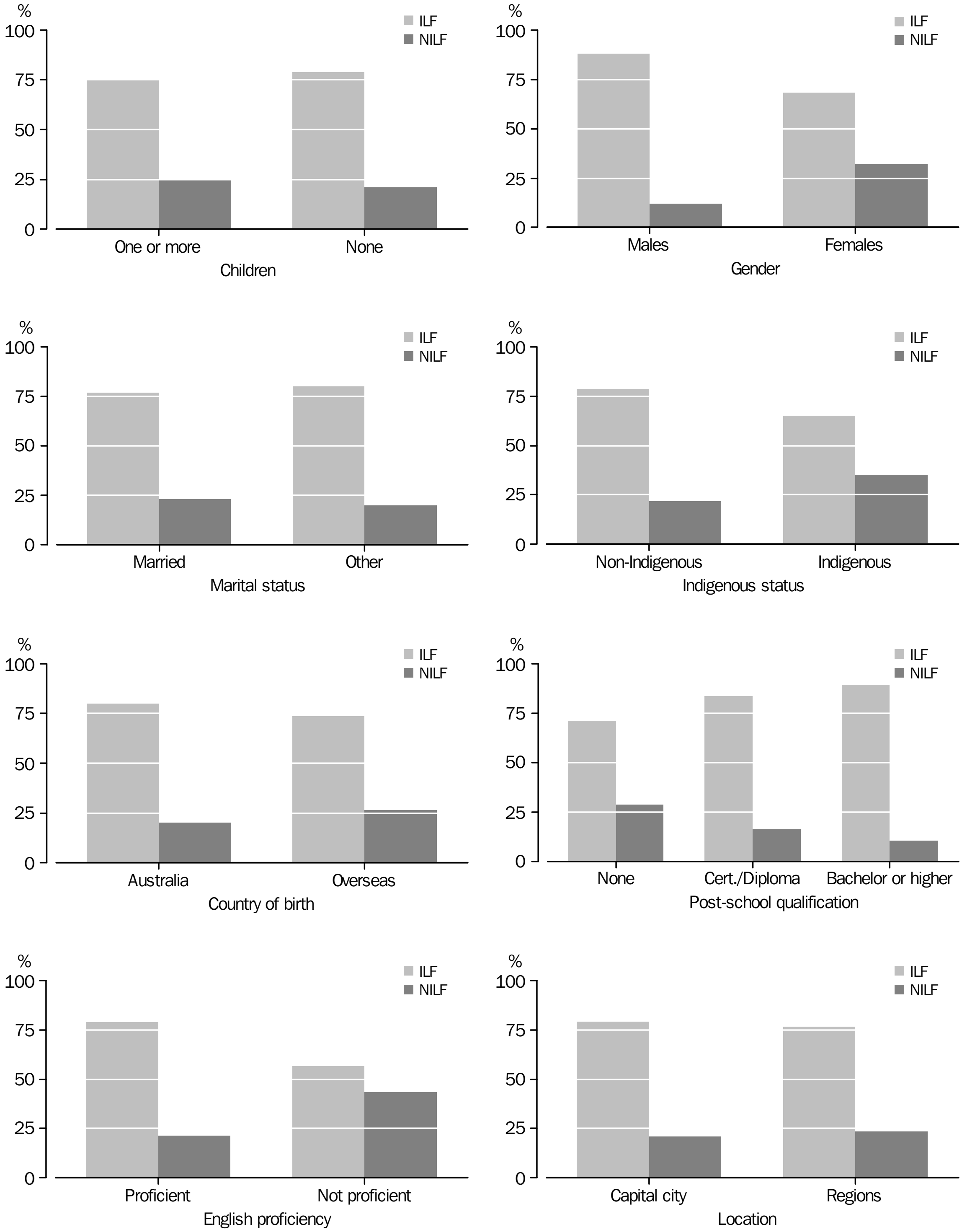
* Cohorts were defined in terms of three-year birth period. From the above 16x5 age-period groupings, we can identify 22 three-year birth cohorts as shown in the table. The cohorts are labelled C1–C22, with C1 being the oldest cohort (born during 1925–26) and C22 being the youngest (born during 1987–89). The shaded cells along the diagonals show how the three-year cohorts can be tracked over time. Since we do not have the same widths for age and period for the five NHSs, we do not get straight line diagonals but curves, as can be seen, for example, for C10 and C16. However, all the 22 cohorts satisfy the condition Cohort = Period – Age (or Age + Cohort = Period), that is, the cohort birth year is derived by subtracting the age from the period.

E. SUMMARY STATISTICS OF SELECTED VARIABLES IN THE POOLED DATASET

E.1 Summary statistics of selected variables in the pooled dataset (18–64 years)

	N	*Mean*	*Std Dev*	*Min*	*Max*
			FEMALES		
Presence of children	56,318	0.24	0.42	0	1
Age (years)	56,318	37.49	12.33	18	64
Gender	56,318	0.00	0.00	0	0
Married	56,318	0.58	0.49	0	1
ATSI origin	56,318	0.02	0.12	0	1
Australian born	56,318	0.75	0.43	0	1
Non-school qualification	47,815	1.63	0.73	1	3
Poor proficiency in spoken English	56,318	0.02	0.15	0	1
Lives in city	56,318	0.69	0.46	0	1
State	56,318	3.38	2.07	1	8
Period	56,318	2.56	1.34	1	5
Cohort	56,318	2.72	1.14	1	4
Health fair/poor	56,318	0.14	0.34	0	1
Arthritis	56,318	0.14	0.34	0	1
Asthma	56,318	0.11	0.31	0	1
Cancer	56,318	0.01	0.12	0	1
Diabetes	56,318	0.02	0.13	0	1
Heart	56,318	0.02	0.13	0	1
Labour force status	56,318	0.70	0.45	0	1
			MALES		
Presence of children	53,173	0.14	0.36	0	1
Age (years)	53,173	37.42	12.81	18	64
Gender	53,173	1.00	0.00	1	1
Married	53,173	0.56	0.50	0	1
ATSI origin	53,173	0.01	0.11	0	1
Australian born	53,173	0.74	0.44	0	1
Non-school qualification	44,985	1.71	0.74	1	3
Poor proficiency in spoken English	53,173	0.02	0.16	0	1
Lives in city	53,173	0.69	0.47	0	1
State	53,173	3.38	2.11	1	8
Period	53,173	2.52	1.39	1	5
Cohort	53,173	2.72	1.19	1	4
Health fair/poor	53,173	0.14	0.36	0	1
Arthritis	53,173	0.10	0.30	0	1
Asthma	53,173	0.08	0.27	0	1
Cancer	53,173	0.01	0.11	0	1
Diabetes	53,173	0.02	0.14	0	1
Heart	53,173	0.02	0.14	0	1
Labour force status	53,173	0.88	0.32	0	1
			ALL		
Presence of children	109,491	0.19	0.39	0	1
Age (years)	109,491	37.45	12.57	18	64
Gender	109,491	0.50	0.50	0	1
Married	109,491	0.57	0.50	0	1
ATSI origin	109,491	0.01	0.12	0	1
Australian born	109,491	0.74	0.44	0	1
Non-school qualification	92,800	1.67	0.74	1	3
Poor proficiency in spoken English	109,491	0.02	0.15	0	1
Lives in city	109,491	0.69	0.46	0	1
State	109,491	3.38	2.09	1	8
Period	109,491	2.54	1.37	1	5
Cohort	109,491	2.72	1.16	1	4
Health fair/poor	109,491	0.14	0.35	0	1
Arthritis	109,491	0.12	0.32	0	1
Asthma	109,491	0.09	0.29	0	1
Cancer	109,491	0.01	0.11	0	1
Diabetes	109,491	0.02	0.13	0	1
Heart	109,491	0.02	0.13	0	1
Labour force status	109,491	0.79	0.41	0	1

F. RELATIONSHIP BETWEEN LABOUR FORCE PARTICIPATION AND SELECTED VARIABLES



G. DISENTANGLING THE AGE, PERIOD AND COHORT EFFECTS

Age, period and cohort (APC) analysis is useful in studying outcomes or events occurring over time. In fact any phenomenon that has a time dimension has age, period and cohort effects (McKenzie, 2005). Age effects are the more common life cycle effects where age has a strong influence on the outcome of interest, e.g. labour force participation changes as people move through different stages of their life. Period effects are the effect of social and economic conditions prevailing in a given year or point in time. Cohort effects are the effects of the specific characteristics or experience of the cohorts on the outcome of interest, e.g. prevalence of smoking among different generations.

The APC accounting model is used when all of the three (age, period and cohort) are potentially of interest in the phenomenon under consideration. Decomposing these effects and separating them can help identify which of these effects, if any, are present and which effects are more dominant than the others. The APC approach has been widely used as a general methodology for estimating age, period, and cohort effects in demographic and social research (Yang *et al.*, 2008).

The APC decomposition model presented in Equation (2) Section 4 can be used to estimate the coefficients of each of the APC categories. However, the inclusion of all three APC variables in the model poses a problem in model estimation. The APC variables are not independent of each other, but any one variable is a linear combination of the other two. For instance, given A and P, we can determine the birth cohort (C), since $C = P - A$ or $P = C + A$ or $A = P - C$. This gives rise to the problem of perfect collinearity or the 'identification' problem in which case it is not possible to simultaneously estimate the true effects of APC, unless some additional constraints are imposed on the parameters of some of these variables. Thus, it is not always easy to disentangle the relative effects of each and the effect of one could be confounded by the other two. Decomposing the APC effects has proved to be a methodological challenge and several alternative methods have been proposed to resolve this problem. These methods include:

- assuming only two of the three APC variables affect the outcome[38]
- assuming that two age or two period or two cohort parameters are equal, which is also known as the coefficient constraints approach
- using proxy variables for one of the APC variables.[39]

---

38 For example it could be assumed that the cohort or the period effect is zero, on average, so only the age variable (because it is an important determinant of social behaviour) and one other variable is kept.

39 For example assuming cohort effect to be proportional to cohort size or using unemployment rate as a proxy for period effect.

All these methods, however, require strong theoretical assumptions and have issues/problems which may or may not be justified in particular circumstances. The main problems with these methods are that: there is generally an element of arbitrariness/value judgement involved; there is heavy reliance placed upon external information; and the parameter estimates are sensitive to the choice of the constraints imposed (Yang *et al.*, 2008). Thus, there does not seem to be any consensus as to the most appropriate method to use to resolve this identification problem.

The APC model in Equation (2), Section 4, can be written in the conventional matrix form using Y in place of M as:

$$Y = Xb \tag{1}$$

where Y is a vector of labour participation rate and X is the design matrix (a vector consisting of 1|0 dummy variables) and b denotes the vector of model parameters of the age, period and cohort.

$$b = \left(\mu, \alpha_1, \ldots, \alpha_{a-1}, \beta_1, \ldots, \beta_{p-1}, \gamma_1, \ldots, \gamma_{a+p-2}\right)^{\mathrm{T}}. \tag{2}$$

If the design matrix *X* is of full rank then we can solve for *b* as follows:

$$\hat{b} = \left(X^{\mathrm{T}} X\right)^{-1} X^{\mathrm{T}} Y. \tag{3}$$

But because of perfect linear relationship between age, period and cohort effects the design matrix is singular and not of full rank, i.e., it is one less than full rank and therefore we cannot find an inverse for $\left(X^{\mathrm{T}} X\right)$ to derive the model coefficients for the APC effects. This is the model identification problem of APC analysis. Therefore due to perfect collinearity between APC it is not possible to separately estimate the effects of cohort, age, and period without imposing at least one constraint on the coefficients in addition to the parameterization in (3) (Yang *et al.*, 2008).

Yang *et al.* (2008) provide an alternative strategy to the commonly used solution above. This method appears to offer a more satisfactory solution to the APC identification problem. This method is referred to as the Intrinsic Estimator (IE) method. This approach is based on the use of the Moore–Penrose generalised inverse to estimate the APC coefficients. The generalised inverse is a method of finding the inverse of a matrix when it is singular or of not full rank, i.e., where the number of rows is not equal to number of columns.

Under IE given a matrix *X* its generalised inverse X^{+} can be found that produces a unique solution for the b parameters as follows:

$$\hat{b} = \left(X^{+} Y\right). \tag{4}$$

The derivation of the IE is equivalent to conducting a principal components analysis (i.e. reducing dimensionality of the set of variables). Compared to other methods used, this method appears to remove the arbitrariness of the coefficient constraints approach, i.e., it restores objectivity to the analysis in that it lets the data decide the shape of the effects rather than imposing constraints based on judgements. This method also has certain optimality properties compared to the other estimable solutions. It may be worth emphasising that the intrinsic estimator does not solve the identification problem, but it provides a more objective way of imposing the constraints on the APC variables prior to estimating the coefficients.

It may be noted that separate APC effects can be estimated in cases where the collinearity problem does not arise, for example when there are different widths for age groups and period intervals. The cohorts formed could be based on wider intervals than the age groups. Such groupings can avoid the perfect collinearity problem and the APC effects in this case can be estimated since the matrix will be of full rank and the inverse of the design matrix exists. In our case, this was achieved by defining the cohorts more broadly that reduced the number of cohorts from 22 to 4. For the four-category cohort, we used the normal inverse to derive the APC coefficients unlike the 22-category cohort where there was collinearity and for which the generalised inverse had to be used to derive the APC coefficients. The coefficients derived under each cohort categorisation are valid.

The generalised inverse is used when normal inverse does not exist. The move from 22 to four cohorts does not mean that we can always find a convenient way to resolve the identification problem by simply manipulating the age and/or cohort groupings. Implicit in the move from 22 to 4 cohorts is that we are imposing constraints based on some personal judgement, in this case having broader cohorts may enable cohorts effects, if they exist, to be seen than narrower definition of cohorts. It may be noted that in forming the cohorts, we used some value judgement to resolve the problem of identification. The definition of cohorts should be based on some valid justification rather than trying to find a way to avoid the identification problem which is inherent in any event observed over time.

H. LOGISTIC REGRESSION MODEL ESTIMATES

H.1 Logistic regression model estimates for association between major long-term health conditions and labour force participation

	Males			*Females*		
Parameter	*Estimate*		*Marginal effects*	*Estimate*		*Marginal effects*
Intercept	–0.5039	*		–0.7785	***	
Has no arthritis	0.0000			0.0000		
Has arthritis	–0.7239	***	–0.0442	–0.3672	***	–0.0647
Has no asthma	0.0000			0.0000		
Has asthma	–0.2416	***	–0.0123	–0.1418	***	–0.0239
Has no cancer	0.0000			0.0000		
Has cancer	–0.4679	***	–0.0268	–0.2466	***	–0.0431
Has no diabetes	0.0000			0.0000		
Has diabetes	–0.5956	***	–0.0361	–0.6545	***	–0.1261
Has no heart disease	0.0000			0.0000		
Has heart disease	–0.7748	***	–0.0508	–0.3629	***	–0.0653
Has no children	0.0000			0.0000		
Has children	0.0583		0.0027	–1.2486	***	–0.2351
Age (years)	0.1916	***	–0.0021	0.1343	***	–0.0071
Age squared	–0.0031	***	0.997	–0.0023	***	
Not married	0.0000			0.0000		
Married	0.8376	***	0.044	–0.2369		–0.0383
Non-indigenous origin	0.0000			0.0000		
Indigenous origin	–0.6273	***	–0.0386	–0.7175	***	–0.1401
Good proficiency in spoken English	0.0000			0.0000		
Poor proficiency in spoken English	–0.9232	***	–0.0647	–1.0721	***	–0.2230
Overseas born	0.0000			0.0000		
Australian born	0.3798	***	0.0193	0.3273	***	0.0560
Lives outside capital city	0.0000			0.0000		
Lives in capital city	0.0367		0.0017	0.1710	***	0.0284
Has no non-school qualification	0.0000			0.0000		
Has certificate/diploma education	0.5816	***	0.0258	0.6332	***	0.0963
Has degree and above education	0.8777	***	0.0320	1.1438	***	0.1470
Period 1989/90	0.0000			0.0000		
Period 1995	–0.1946	***	–0.0096	0.0103		0.0017
Period 2001	–0.2518	***	–0.0127	0.5732	***	0.0834
Period 2004/05	–0.1721	**	–0.0084	0.6321	***	0.0913
Period 2007/08	–0.0325		–0.0015	0.8745	***	0.1180
Cohort born 1925–48	0.0000			0.0000		
Cohort born 1949–57	0.0008		0.0000	0.0742		0.0120
Cohort born 1958–66	–0.2596	***	–0.0129	–0.0751		–0.0124
Cohort born 1967–89	–0.4656	***	–0.0231	–0.3279	***	–0.0550
n	44,985			47,815		
Likelihood Ratio (p Value)	5,867		(<0.0001)	7,918		(<0.0001)
Deviance Value/DF (p Value)	0.90		(1.00)	1.18		(<0.0001)
Max-rescaled R^2	0.24			0.22		
% Concordant	80.2			74.3		

** and *** indicate the coefficient is statistically significant at 5% and 1%, respectively.

H.2 Logistic regression model estimates for association between presence of one or more major long-term health conditions and labour force participation

	Males			Females		
Parameter	*Estimate*		*Marginal effect*	*Estimate*		*Marginal effect*
Intercept	–0.4754	**		–0.7914	***	
Has no children	0.0000			0.0000		
Has children	0.0734		0.0033	–1.2428	***	–0.2329
Age (years)	0.1928	***	–0.0022	0.1364	***	–0.0074
Age squared	–0.00315	***		–0.00237	***	
Not married	0.0000			0.0000		
Married	0.8369	***	0.0412	–0.2324	***	–0.0374
Non-ATSI origin	0.0000			0.0000		
ATSI origin	–0.6515	***	–0.0404	–0.7454	***	–0.1458
Good proficiency in spoken English	0.0000			0.0000		
Poor proficiency in spoken English	–0.9224	***	–0.0644	–1.0694	***	–0.2216
Overseas born	0.0000			0.0000		
Australian born	0.3776	***	0.0191	0.3260	***	0.0555
Lives outside capital city	0.0000			0.0000		
Lives in capital city	0.0440		0.0021	0.1711	***	0.0283
Has no non-school qualification	0.0000			0.0000		
Has certificate/diploma education	0.5776	***	0.0255	0.6373	***	0.0963
Has degree and above education	0.8965	***	0.0324	1.1506	***	0.1468
Has no diseases§	0.0000			0.0000		
Has one or more diseases §	–0.6372	***	–0.0354	–0.2944	***	0.0499
Period 1989/90	0.0000			0.0000		
Period 1995	–0.2046	***	–0.0101	0.0039		0.0006
Period 2001	–0.2679	***	–0.0135	0.5622	***	0.0815
Period 2004/05	–0.1987	***	–0.0098	0.6128	***	0.0883
Period 2007/08	–0.0791		–0.0038	0.8517	***	0.1148
Cohort born 1925–48	0.0000			0.0000		
Cohort born 1949–57	–0.0054		–0.0002	0.0692		0.0111
Cohort born 1958–66	–0.2605	**	–0.0129	–0.0761		–0.0125
Cohort born 1967–89	0.4415	***	–0.0218	–0.3147	***	–0.0524
n	44,985			47,815		
Likelihood Ratio (p Value)	5,704		(<0.0001)	7,791		(<0.0001)
Deviance Value/DF (p Value)	0.94		(1.00)	1.23		(<0.0001)
Max-rescaled R^2	0.23			0.22		
% Concordant	79.9			74		

** and *** indicate the coefficient is statistically significant at 5% and 1%, respectively.

§ This refers to the five major diseases considered in table H.1.

FOR MORE INFORMATION . . .

INTERNET	**www.abs.gov.au** The ABS website is the best place for data from our publications and information about the ABS.
LIBRARY	A range of ABS publications are available from public and tertiary libraries Australia wide. Contact your nearest library to determine whether it has the ABS statistics you require, or visit our website for a list of libraries.

INFORMATION AND REFERRAL SERVICE

Our consultants can help you access the full range of information published by the ABS that is available free of charge from our website, or purchase a hard copy publication. Information tailored to your needs can also be requested as a 'user pays' service. Specialists are on hand to help you with analytical or methodological advice.

PHONE	1300 135 070
EMAIL	client.services@abs.gov.au
FAX	1300 135 211
POST	Client Services, ABS, GPO Box 796, Sydney NSW 2001

FREE ACCESS TO STATISTICS

All statistics on the ABS website can be downloaded free of charge.

WEB ADDRESS	www.abs.gov.au