Australian Government
Productivity Commission

Influences on Indigenous Labour Market Outcomes

Productivity Commission Staff Working Paper

December 2011

Annie Savvas
Clare Boulton
Ellie Jepsen

The views expressed in this paper are those of the staff involved and do not necessarily reflect the views of the Productivity Commission.

ISBN 978-1-74037-375-3

This publication is available from the Productivity Commission website at www.pc.gov.au. If you require part or all of this publication in a different format, please contact Media and Publications.

Publications Inquiries:
Media and Publications
Productivity Commission
Locked Bag 2 Collins Street East
Melbourne VIC 8003

Tel: (03) 9653 2244
Fax: (03) 9653 2303
Email: maps@pc.gov.au

General Inquiries:
Tel: (03) 9653 2100 or (02) 6240 3200

An appropriate citation for this paper is:

Savvas, A., Boulton, C. and Jepsen, E. 2011, *Influences on Indigenous Labour Market Outcomes*, Productivity Commission Staff Working Paper, Melbourne.

JEL code: J490

The Productivity Commission

The Productivity Commission is the Australian Government's independent research and advisory body on a range of economic, social and environmental issues affecting the welfare of Australians. Its role, expressed most simply, is to help governments make better policies, in the long term interest of the Australian community.

The Commission's independence is underpinned by an Act of Parliament. Its processes and outputs are open to public scrutiny and are driven by concern for the wellbeing of the community as a whole.

Further information on the Productivity Commission can be obtained from the Commission's website (www.pc.gov.au) or by contacting Media and Publications on (03) 9653 2244 or email: maps@pc.gov.au

Contents

Acknowledgements

The authors wish to thank the following people for their help and advice in the production of this paper. The paper was refereed by Prof Anne Daly at the University of Canberra, and Dr Nicholas Biddle at the Centre for Aboriginal Economic Policy Research at ANU. At the Productivity Commission, Dr Jenny Gordon and Lawrence McDonald acted as referees. The paper has also benefited from comments made by Dr Patrick Jomini, Lisa Gropp, Dr Patrick Laplagne, Peter Daniel, Anthony Shomos, Leo Soames, Maurice Glover and Matthew Forbes.

The paper use a confidentialised unit record file (CURF), accessed from the Australian Bureau of Statistics (ABS) via its Remote Access Data Laboratory (RADL). The findings and views reported in this paper, however, are those of the Productivity Commission staff involved and should not be attributed to either the ABS or the Productivity Commission.

Abbreviations

Abbreviations

ABS	Australian Bureau of Statistics
CDEP	Community Development Employment Projects
COAG	Council of Australian Governments
IIA	Independence of Irrelevant Alternatives
LMO	Labour Market Outcome
MNL	Multinomial Logit
MNP	Multinomial Probit
NATSISS	National Aboriginal and Torres Strait Islander Social Survey
NIRA	National Indigenous Reform Agreement
OID	Report on *Overcoming Indigenous Disadvantage: Key Indicators*
PC	Productivity Commission
SCRGSP	Steering Committee for the Review of Government Service Provision
SEIFA	Socio-Economic Index for Areas
SEM	Simultaneous Equations Model

Glossary

Cultural activities	Cultural activities include: fishing; hunting; gathering wild plants/berries; making Aboriginal and Torres Strait Islander arts or crafts; performing Aboriginal or Torres Strait Islander music, dance, theatre; and writing or telling Aboriginal or Torres Strait Islander stories.
Cultural events	Cultural events, ceremonies or organisations include: ceremonies; NAIDOC week activities; sports carnivals (other than NAIDOC); festivals or carnivals involving arts, craft, music of dance (other than NAIDOC); Aboriginal/Torres Strait Islander organisations; and funerals/sorry business.
Endogeneity bias	The bias affecting the coefficients of an estimated equation in which one (or more) of the explanatory variables is correlated with the error term.
Human capital	The set of attributes, including skills and knowledge, that influence a person's ability to perform work and contribute to economic value.
Kessler Psychological Distress Scale	A non-specific scale of psychological distress designed to reflect levels of negative emotional states.
Labour force participation	A participant in the labour force is an Indigenous person aged between 15 years and over, and who is either in employed (either in non-CDEP or CDEP employment) or unemployed. The alternative is not in the labour market.

Labour market outcomes	Refers to the four discrete labour market outcomes for Indigenous people used in multinomial regression analysis: mainstream (non-CDEP) employment, unemployment, CDEP participation and not in the labour force.
Marginal effect	For a binary variable: the effect on the dependent variable (the predicted probability of an LMO) of the binary variable changing from 0 to 1, holding all other explanatory variables constant. For a continuous variable: the effect on the dependent variable of a one-unit change in the continuous variable, holding all other explanatory variables constant.
Panel data	Repeated observations over time on the characteristics of the same individual.
Self-assessed health status	A summary measure of a person's overall health status, as determined and reported by that person.
Simultaneity	A situation arising when two variables being modelled influence each other.
Simultaneous equations model	A econometric model consisting of two (or more) equations, to be estimated jointly.
Social capital	Includes the networks, norms, values and understandings that facilitate cooperation within or among groups.
Unobserved heterogeneity (omitted variable bias)	Describes the case when unobserved characteristics of a person jointly influence two (or more) of the variables being modelled, including the dependent variable.

Key points

- This paper uses data from the 2008 National Aboriginal and Torres Strait Islander Social Survey (NATSISS) to model the links between various personal characteristics and activities, and four Indigenous labour market outcomes. As causality may run in either direction, results are interpreted as associations.
- Results typically differ between men and women, in part because of different responsibilities in relation to child rearing, broader family responsibilities and other specialisation in unpaid work. Of these differences, the most notable were those for educational attainment. Results suggest that for women, attainment of year 10 or above is associated with a higher probability of employment and labour force participation, whereas this is less evident for men.
- Results confirm findings from previous research that other human capital factors, such as good health, are positively associated with Indigenous employment and labour force participation.
- Arrest and imprisonment were found to be negatively associated with Indigenous employment and labour force participation. A history of arrest was found to have a larger negative association with employment for women compared to men.
- Models that omit personal characteristics that might determine an individual's labour supply (such as ability, motivation and preferences) are likely to produce biased estimates of the effect of the human capital factors that are included in the model (such as education, health and disability).
  - One benefit of including variables in the model that represent cultural and social engagement is that they may act as proxies for unobserved personal characteristics, reducing the impact of this source of bias.
- The cultural and social engagement factors were found to have statistically significant associations with Indigenous labour market outcomes for women, but not for men.
  - For women there were positive associations between employment and labour force participation, and engagement in social cultural events and the provision of support outside the household.
- A greater understanding of the links between social capital and labour market outcomes could be explored using alternative econometric models. However, there are limits to the indicators that can be derived from the data.

1 Introduction

This staff working paper examines factors that potentially influence Indigenous labour market outcomes (LMOs). It uses regression analysis, and builds on the simple model that was presented in *Overcoming Indigenous Disadvantage: Key Indicators* (OID Report) *2011* (SCRGSP 2011), by including additional variables. The analysis uses the 2008 National Aboriginal and Torres Strait Islander Social Survey (NATSISS), a rich source of information on the characteristics of Indigenous people, including data on LMOs and many factors that might contribute to them (appendix table A.1 contains a description of the variables included in the analysis). The analysis does not make comparisons with non-Indigenous Australians.

Empirical analysis can be used to test and quantify relationships that have been developed in theory. For example, policy makers might be interested in those factors that have the greatest association with Indigenous people's decisions to participate in the labour market and their success in obtaining a job.

The purpose of this analysis is to quantify those associations. The aim is to add variables that represent social and cultural factors to the basic model in the 2011 OID Report to obtain insight into the effects of unobserved personal characteristics, and whether the way Indigenous people engage with their community and culture affects their LMOs.

After controlling for a number of personal and demographic characteristics, the econometric modelling could help answer the following questions:

- Is a person's health and education strongly associated with LMOs?
- Does contact with the criminal justice system have an association with a person's LMO?
- Do social and cultural factors associate with LMOs?
- Do results differ for men and women?

The statistical methods used cannot establish the *direction* of causality between these factors and LMOs, but evidence of statistically significant relationships can provide a basis for future research to identify causality.

The outline of this chapter is as follows: the policy context is discussed in section 1.1; the LMOs examined are described in section 1.2; recent trends in Indigenous LMOs are presented in section 1.3; the reasons for undertaking empirical analysis are explained in section 1.4; and an outline of the paper is provided in section 1.5.

1.1 Policy context — an integrated, multidimensional approach to targeting Indigenous disadvantage

The framework that COAG has developed to address Indigenous disadvantage, as outlined in the National Indigenous Reform Agreement (NIRA) (COAG 2011), is an integrated strategy with an emphasis on a whole-of-government approach to achieving objectives. A key focus of the work is to address the multiple dimensions and causes of disadvantage experienced by Indigenous Australians. It includes six high level targets for closing the gap of Indigenous disadvantage and the building blocks required to achieve them. The aim is to improve performance against the targets by addressing disadvantage on a range of policy fronts.

An important target is to 'halve the gap in employment outcomes between Indigenous and non-Indigenous Australians within a decade' (COAG 2011, p. 3). This has been interpreted as halving the gap in the employment to population ratio by 2018 (CRC 2010). This target reflects the value placed on participation in paid work in the community. Employment contributes to improved living standards and overall wellbeing. Being employed leads to higher income for families and communities, which in turn has a positive influence on the health and education of children (that is, there is an intergenerational benefit). Employment also enhances self-esteem, increases opportunities for self-development, influences interaction at the family and community levels, and promotes social inclusion.

The building block approach agreed by COAG, as it relates to employment outcomes, is shown in table 1.1 This approach aims to make changes to factors that affect high level target outcomes linked to economic participation.

Table 1.1 **Examples of National Indigenous Reform Agreement targets, building blocks and outputs**

1. *Halve the gap in employment outcomes between Indigenous and non-Indigenous Australians within a decade*		
Building Blocks	**COAG Agreements**	**Outputs**
Early Childhood	Indigenous Early Childhood Development National Partnership	Establishment of a minimum of 35 Children and Family Centres in urban, regional and remote areas with high Indigenous populations
		Provision of early learning, child care and parent and family support services to Indigenous families at or through each of the Children and Family Centres
Schooling	Low SES School Communities National Partnership	Provision of innovative and tailored learning opportunities and external partnerships with parents, other schools, businesses and communities
Health	National Healthcare Agreement	Chronic disease management
	Preventive Health National Partnership	Chronic disease management, including good health, fitness and nutrition
	Closing the Gap in Indigenous Health Outcomes National Partnership	Chronic disease management and prevention, including good health, fitness and nutrition
Economic Participation	Indigenous Economic Participation National Partnership	Focus on industry sectors with jobs growth potential (e.g. health, education, construction and government services)
		Increase access to employment and training services (extend intensive assistance program to Indigenous job seekers, wage assistance programs, and continue and extend the STEP program)
	National Agreement for Skills and Workforce Development	Increase access to employment and training services – extend intensive assistance program to Indigenous job seekers
		Build aspirations and foundation skills of unemployed and those outside the labour force
	Improving Teacher Quality National Partnership	Focus on industry sectors with jobs growth potential (e.g. education)
		Build professional pathways for Indigenous people and Indigenous education workers who wish to progress to teaching

Table 1.2 **Examples of National Indigenous Reform Agreement targets, building blocks and outputs**

<table>
<tr><td colspan="3">1. Halve the gap in employment outcomes between Indigenous and non-Indigenous Australians within a decade</td></tr>
<tr><td>Building Blocks</td><td>COAG Agreements</td><td>Outputs</td></tr>
<tr><td>Healthy Homes</td><td>Remote Indigenous Housing National Partnership</td><td>Local investment in construction – government procurement includes Indigenous participation</td></tr>
<tr><td rowspan="4">Safe Communities</td><td>National Healthcare Agreement</td><td>Mental health promotion programs (including coping skills)</td></tr>
<tr><td>Preventive Health National Partnership</td><td>Addressing alcohol / substance abuse and harm</td></tr>
<tr><td rowspan="2">Closing the Gap in Indigenous Health Outcomes National Partnership</td><td>Mental health promotion</td></tr>
<tr><td>Diversionary programs / skills learning within juvenile justice programs</td></tr>
</table>

Source: COAG (2011).

1.2 What labour market outcomes are examined?

In the labour market literature, LMOs that have been the subject of research include income, hours worked and labour force status, with the latter encompassing labour force participation, employment and unemployment. This study does not examine income or hours worked (which are areas for possible future research).

Labour force participation describes the economically active population or the formal supply of labour. It is defined as the number of people contributing to, or willing to contribute to, the supply of labour, and is often expressed as a percentage of the working age population (usually aged 15 to 64). It comprises two mutually exclusive groups:

- the employed (people who have worked for at least one hour in the reference week)
- the unemployed (people who are without work, but are actively looking for work and available to start work within four weeks).

The remainder of the working age population is described as being not in the labour force.

Labour force status in this project is examined as four labour market outcomes — ‘mainstream (non-CDEP) employment’, ‘unemployment’, ‘CDEP participation’ and ‘not in the labour force’. Unless otherwise specified, references to ‘employment’ are to ‘non-CDEP employment’. The analysis of labour force status of Indigenous people is complicated by the classification of participants in the Community Development Employment Project (CDEP) Scheme as employed in the 2008 NATSISS (box 1.1). People who have received wages for participating in CDEP for at least one hour in the reference week are included in the group of employed people in ABS surveys. Subsequent changes to the operation of the CDEP scheme will require CDEP participants to be classified as unemployed which would distort any historical comparison of Indigenous labour force trends using the 2008 NATSISS data and later data collections.

Box 1.1 **Community Development Employment Projects (CDEP)**

The original aim of the CDEP program — introduced in 1977 — was to create local employment opportunities in remote Indigenous communities where the labour market might not otherwise offer employment. The program was later extended to all areas. However, a recent restructuring of the CDEP program has seen its focus shift back to supporting employment opportunities in remote Indigenous communities.

For statistical purposes, in the 2008 NATSISS, the ABS classified known participants in CDEP as employed rather than as unemployed. Consequently the employment rate for Indigenous people appears higher than it would be if participants in the CDEP program were classified as unemployed. It is important to consider CDEP when analysing historical labour force and unemployment data because, at the time data were collected:

- CDEP participant payments comprised a mix of both wages and income support payments such as NewStart Allowance
- CDEP had elements of both unemployment and employment, especially in remote and very remote areas. Some CDEP activities were similar to those undertaken by participants in Work for the Dole, while other activities were essential roles in municipal services, health care, community services, education and other sectors that would be considered employment in mainstream communities and organisations. However, through the National Partnership Agreement on Indigenous Economic Participation agreed in early 2009, COAG committed to converting around 2000 CDEP positions to ongoing jobs in government service provision (COAG 2011, p. 5).

Following the collection of the 2008 NATSISS data significant changes to CDEP were announced. Since then, CDEP has ceased operating in non-remote locations where the economy was already reasonably established, with services to Indigenous job seekers in those areas now provided through Job Services Australia and the Indigenous Employment Program. Commencing on 1 July 2009, new CDEP participants received corresponding income support payments rather than wages, with existing CDEP participants continuing to access CDEP wages until 30 June 2011 before transferring to the new payment arrangements.

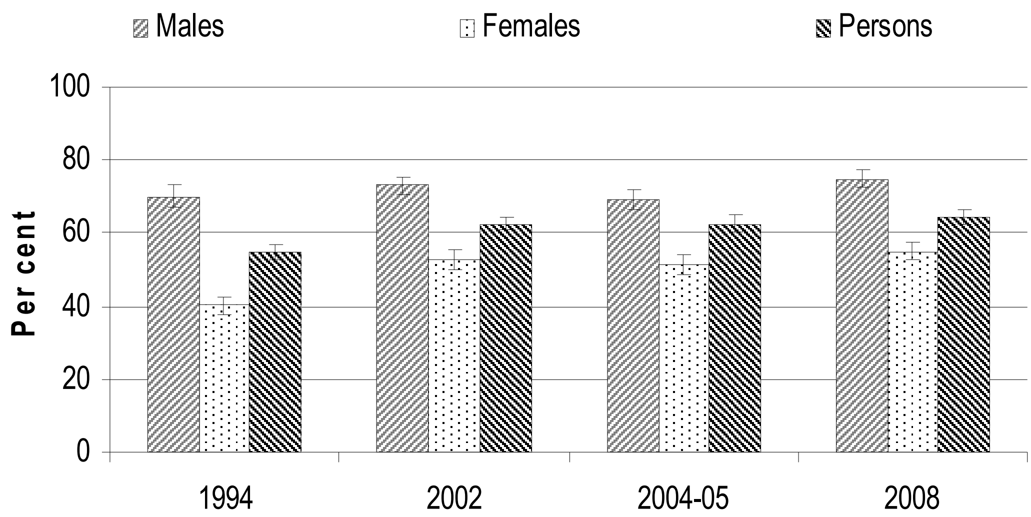
Source: SCRGSP (2011).

1.3 Recent history of Indigenous labour market outcomes

The OID Report (SCRGSP 2011) contains an analysis of trends in Indigenous and non-Indigenous labour market outcomes. These include labour force participation (LFP) rates, employment to population ratios, unemployment rates, and CDEP participation rates.

Between 1994 and 2008, for people aged 15 to 64 years, labour force participation increased from 40.2 to 55.0 per cent for Indigenous women and from 70.0 to 74.9 per cent for Indigenous men (figure 1.1). From 2004-05 to 2008, the gap between Indigenous and non-Indigenous labour force participation decreased from 17.6 to 14.4 percentage points (SCRGSP 2011).

Figure 1.1 **Indigenous labour force participation, people aged 15–64 years, 1994 to 2008[a,b]**



[a] Error bars represent 95 per cent confidence intervals around each estimate. [b] Labour force participation is the number of employed plus those who were unemployed and available for work expressed as a percentage of people aged 15–64 years.

Source: ABS (unpublished) NATSIS 1994; ABS (unpublished) NATSISS 2002; ABS (unpublished) NATSIHS 2004-05; ABS NATSISS 2008; ABS NHS 2007-08; SCRGSP (2011).

Between 2004-05 and 2008, for those aged 15 to 64, the employment to population ratio increased for Indigenous people (from 50.7 to 53.8 per cent), and for non-Indigenous people (from 74.2 to 76.0 per cent). Overall, there was no significant change in the gap in the employment to population ratio between Indigenous and non-Indigenous people over this period (from 23.5 to 22.2 percentage points between 2004-05 and 2008) (SCRGSP 2011).

The unemployment rate for Indigenous people between 1994 and 2008 decreased significantly from 28.1 to 17.1 per cent for women and 32.8 to 16.3 per cent for men. Long-term unemployment also fell between 2004-05 and 2008 both as a proportion of the labour force and as a proportion of all unemployed (SCRGSP 2011).

- Long-term unemployment as a proportion of the Indigenous labour force decreased from 5.2 to 4.3 per cent.
- Although the number of Indigenous long-term unemployed only decreased slightly from 8707 in 2004-05 to 8678 in 2008, as a proportion of all Indigenous unemployed it fell from 33.3 to 26.0 per cent, reflecting a larger number of Indigenous people entering the labour force during this period.

CDEP participation has fallen since 2002, reflecting government policy of phasing out CDEP in non-remote areas. As a proportion of the Indigenous population, CDEP participation decreased from 17.7 per cent of males in 2002 to 7.7 per cent in 2008. For females, the fall was from 9.7 to 4.5 per cent over the same period (SCRGSP 2011).

1.4 Why do empirical analysis?

The trends in LMOs examined above assist in determining the extent to which objectives are being achieved in relation to Indigenous disadvantage but provide no explanation for those outcomes.

The OID Report examines patterns of multiple disadvantage and observes that:

> Where analysis shows that a particular population who experience one type of disadvantage also experience another kind of disadvantage, the two aspects of disadvantage are assumed to be linked or associated in some way; for example, low levels of educational attainment appear to be linked with high levels of unemployment (SCRGSP 2011, p. 13.2).

Analyses in the OID Report illustrate that Indigenous people who are unemployed or not in the labour force are more likely to be disadvantaged in other aspects compared to non-Indigenous people. They examine links between unemployment and labour force participation, and education, income, housing, health and health risk factors, crime, difficulty speaking English and removal from family.

They are a type of bivariate analysis — because they only look at two variables at a time — which are best described as ‘associations’ that do not indicate cause and effect relationships. Other studies have also used this bivariate approach to examine patterns of disadvantage (Hunter and Borland 1997; Hunter and Gray 1999). However, associations can vary depending on the demographic group being examined. For example, LMOs vary significantly between men and women, as women tend to take on more of the unpaid work related to child rearing and caring for other family members. The associations will also vary depending on other factors such as a person's age, education and where they live.

This paper develops a framework for undertaking analysis that examines relationships between LMOs and factors of potential interest to policy makers, while controlling for other factors. The framework is based on a microeconomic model of labour demand and supply, uses human and social capital theory to select modelled factors, and uses multivariate regression analysis to provide numerical estimates of the relationships, while controlling for other factors.

However, as noted in Laplagne et al. (2007), modelling the theoretical link between LMOs and these factors has a number of problems:

- Variables available in datasets are often imperfect proxies for human and social capital.
- Some elements of human and social capital are not observable, such as motivation, ability and preferences.
- The causal relationships are complicated, and the direction of causality is not always one way (for example, health affects LMOs but LMOs might also affect health simultaneously).
- There may be errors in the measurement of variables.

In some areas, research has provided evidence of the links between particular factors, for example:

- education and income levels are estimated to account for between one-third and one-half of the gap between Indigenous and non-Indigenous people's self-assessed health status (AIHW 2004; Booth and Carroll 2005)
- socioeconomic differences account for between one-third and two-thirds of the gap in early childhood outcomes (Leigh and Gong 2008)
- eleven modifiable risk factors account for almost half of the gap in disease burden (including tobacco, obesity, physical inactivity, high blood cholesterol and high blood pressure (Vos et al. 2007)).

Research on the underlying *causes* of Indigenous disadvantage in other policy areas is not as well developed.

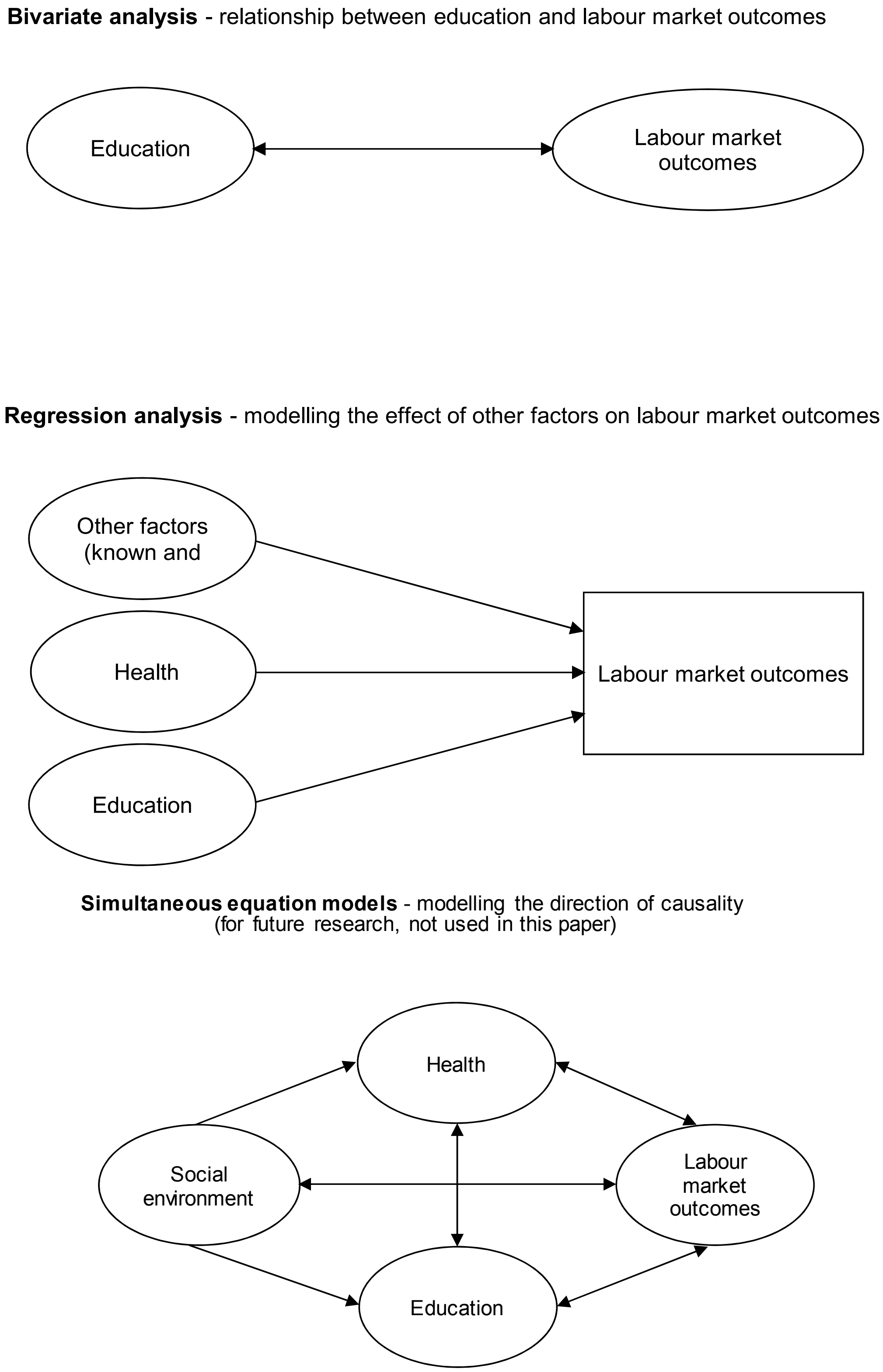
Laplagne et al. (2007), in looking at the effects of health and education on labour force participation, used a simultaneous equations model (to account for endogeneity between human capital factors and participation) and panel data (to account for unobserved factors). These models are complex, and panel data on the characteristics of Indigenous people are not available so these approaches are not undertaken here.

Instead, a broad range of variables is considered, some of which are intended to be proxies for unobservable characteristics. The results — the marginal effect of an explanatory variable on a particular dependent variable — should be interpreted as 'associations' between explanatory variables and dependent variables, rather than estimates of cause and effect. Nonetheless, the regression analysis results provide more robust numerical estimates of the associations than bivariate analysis, because it estimates the effect of variables separately, allowing for other control variables (such as age and marital status), as well as a range of other determining factors) to be held constant. Diagrams explaining the different types of analyses described above are shown in figure 1.2.

1.5 Outline of the study

The outline of this paper is as follows: the analytical approach is discussed in chapter 2; the econometric model, interpretation of results and some qualifications are discussed in chapter 3; the variables used in the econometric model that are suggested by the theoretical framework are described in chapter 4; the results are discussed in chapter 5; and possible future directions for research are provided in chapter 6. An appendix contains detailed results, which are also be found on the Commission's website (www.pc.gov.au) in excel format.

Figure 1.2 **Modelling the relationship between explanatory factors and labour market outcomes**



[a] *Source*: Adapted from Shomos (2010).

2 Analytical framework and literature

This chapter presents a framework for analysing determinants of Indigenous labour market outcomes (LMOs). The theory behind the determinants of LMOs is discussed in section 2.1 and a review of the relevant literature is presented in section 2.2.

2.1 What influences labour market outcomes?

A person's LMO, as described in chapter 1, reflects:

- their willingness and ability to supply labour to the paid labour market
- the demand for labour
- the interaction of labour supply and demand.

Individuals make their decision on whether to participate in the paid labour market based on whether the wage offered exceeds their reservation wage, that is, the lowest wage that a person will accept to undertake a particular job rather than remain unemployed or outside the labour force. A person's reservation wage depends on their perceptions of the costs and benefits of work. The costs include job search costs, transport to and from work, and the opportunity costs of paid employment, including unpaid work and leisure. The benefits include wages (relative to welfare benefits), as well as the status and enhanced wellbeing that an individual might obtain from working.

The reservation wage also depends on a person's expectations of future wages, and their family and living arrangements. A person might not participate in the labour force because they have some other means of financial support, for example, income earned by other household members or investment income. Or they might not be able to work because they have a disability or are in poor health. Decisions about labour force participation may involve consideration of how many hours to supply (and might involve a choice between full or part time paid work).

People also engage in productive activities outside the paid labour market, including:

- work in the household (such as caring for children and other family members, and housework) and community or volunteer work (such as organising and/or participating in social and cultural activities).
- for some Indigenous people, 'traditional activities' such as the production of ceremonial art or the pursuit of a traditional hunter or gatherer lifestyle.

The output of unpaid work can be exchanged in informal markets that allow people to meet their needs without monetary compensation. However, these activities and outputs are not always recorded as employment, even where there is payment 'in kind', but especially where the exchange is implicit. Such arrangements are common in many communities, not just Indigenous communities.

Some people also become discouraged jobseekers, and withdraw from the labour market. Discouraged jobseekers are those who would like to work but are not actively looking for work. For discouraged jobseekers, the costs of searching for a job are high compared to the probability of finding a job. They might observe there are no jobs, or no jobs for people with their skills, in close proximity to where they live and be reluctant to relocate to areas where there are more job opportunities. Discouraged workers might have expectations that they will be discriminated against on the basis of age, gender, race, or their contact with the criminal justice system.

Labour demand is determined by the cost and benefits of producing a firm's output. The benefits are the profits from the goods or services that workers produce. The costs to employers include the cost of finding suitable workers (search costs) and retaining them (wages), relative to other inputs to production. All else given, if wages rise, employers will substitute away from labour inputs.

At the aggregate or national level, labour demand will be influenced by macroeconomic effects which are not uniform across the economy. Employers are less likely to locate in areas where the costs of labour are relatively high or potential workers have low skill levels, there are relatively high levels of crime, low levels of economic activity (demand for goods and services) or high operational costs (for example those related to transport and rent). These factors will contribute to lower labour demand in some areas relative to others. Therefore, some areas might not have many jobs of the type expected to be seen in conventional labour markets (Biddle and Webster 2007).

Labour is not an homogenous commodity, and different workers have different skills and productivity. Finding suitable workers is costly, and the desirable attributes of workers are not always observable. In the absence of perfect information about potential staff, employers use networks, and signals such as education and interaction with the justice system, to identify suitable workers and lower their search costs. Employers (legally) discriminate on the basis of skills and knowledge, but sometimes (illegally) discriminate on the basis of gender, race or age.

There are various theories about influences on the supply and demand of labour. According to human capital theory, a person's productivity is determined by their personal attributes (box 2.1). People are endowed with some attributes (such as innate ability), and acquire and maintain others (such as skills and health). The acquisition and maintenance of attributes is investment in human capital.

Box 2.1 **Human capital theory**

Human capital can be described as the set of attributes that makes it possible for individuals to work and contribute to production (Forbes et al. 2010). It includes knowledge, skills, health, work experience, motivation and work ethic. The neoclassical human capital theory of wage determination in Mincer (1974) explains differences in wages as being determined by the quantity of skills (years of schooling and work experience) possessed by an individual in a competitive labour market.

Human capital may be observed (such as skills) or unobserved (such as motivation and work ethic). Employers may look for signals of a person's human capital, such as their highest education level and work experience. Employers will employ another unit of labour if the marginal revenue produced is greater than the marginal cost. Employers are willing to pay more to relatively higher skilled people because the marginal revenue they contribute is higher than that produced by a relatively lower skilled person.

Human capital improvements (through attainment of more education, skills and work experience as well as improvements in health status) can be expected to increase labour productivity and supply. At the aggregate level, higher workforce participation and labour productivity increases gross domestic product, consumption and community wellbeing.

While innate ability is an endowment, the acquisition and maintenance of many personal attributes can be influenced by a person's social environment. Social capital theory can thus provide some additional insights into LMOs (box 2.2).

Social networks might lower job search costs, making it more likely that a person will participate in the labour market. Stone et al. (2003) proposed that networks influence LMOs by lowering job search costs, providing support to sustain being employed and influencing preferences for work. Social networks might also lower the cost of recruiting suitable employees for businesses, thus increasing labour demand as it reduces the cost of labour relative to capital.

Box 2.2 **Social capital theory**

Social capital as an area of policy research has increasingly attracted interest from academics and governments. While there is no universal agreement on exactly what social capital is, some progress has been made in clarifying the issues and developing some broad definitions. The OECD (2001, p.4) takes social capital to include the 'networks, norms, values and understandings that facilitate cooperation within or among groups'.

Social capital can be thought of as enabling people to secure benefits by virtue of membership in social networks or other social structures. A distinction can be made between the 'structure' (the characteristics of networks) and the 'content' of social capital (the quality of relationships, including trust, reciprocity, tolerance and diversity) (Brough and Bond 2009). Trust, as an aspect of social capital, may be a key component of personal and institutional relationships that determine motivations or opportunities for work.

A distinction is made between bonding and bridging social capital (PC 2003):

- Bonding social capital refers to relations among relatively homogenous groups (such as ethnic groups) that strengthen ties within the particular group. It can either increase labour market activity (because of the support structures it provides) or decrease it (because of the commitments it entails).
- Bridging social capital refers to relations that strengthen ties across heterogeneous groups. It tends to increase labour market activity.

Social norms and networks can have both positive and negative influences on a person's motivation to obtain employment and their preferences for paid versus unpaid work:

- Social networks can lower job search costs by making it easier for a person to locate jobs through informal mechanisms or personal contacts or by providing signals to employers through the informal recommendations of friends.
- Involvement in social networks might be seen by employers as an attractive attribute of a worker that enhances their productivity.
- Social norms or networks affect people's preferences for unpaid work through expectations about family and community commitments such as unpaid carer responsibilities (particularly caring for children, the elderly and people with a disability), and through preference or responsibilities to participate in cultural or traditional activities.
- Social norms can also be downward levelling, where the norm of inactivity contributes to poor labour market outcomes in a community over generations.

Source: Brough and Bond (2009); PC (2003); OECD (2001); Putman (2000).

Some types of social networks may have negative effects on LMOs. For example, Stone et al. (2003, p. 6) note that:

> As well, where individuals are embedded within networks of family, friends, community and institutional ties that support the normative aspects of work, these are likely to reinforce the value of work for that individual, thereby acting to increase a person's likelihood of being employed. Some authors have emphasised the possible 'negative' consequences of some types of social capital (Cox 1997; Portes 1998). For example, some networks may be governed by norms of behaviour that are inconsistent with maintaining employment. That is, where relationships between individuals and institutions are generally negative, or if informal networks are characterised by a non-work ethic. This latter point is consistent with welfare discourse from the United States that emphasises ghettoisation and intergenerational welfare dependence as undermining fulfilment of individuals' responsibility to work (see, for example, Murray 1994; for an Australian example, McCoull and Pech 2000).

2.2 Literature review

There have been numerous empirical analyses of the determinants of LMOs, both in Australia and overseas that use the discrete choice modelling framework (see Cai and Kalb (2006) for a review). There have been several studies of Indigenous LMOs, using the NATSISS (1994 and 2002) and the Census (box 2.3).

There are a number of problems associated with the estimated effects of explanatory variables using these models. In particular, they are affected by omitted variable bias because many attributes such as motivation, preferences and innate ability that affect LMOs are inherently unobservable. Also it is difficult to capture the importance of community and culture to Indigenous people in numerical measures.

Many studies include as explanatory variables the observable attributes of suitable employees used by employers to choose workers, such as education, proficiency with the English language and work experience, as proxies for ability and skills. Other models are more innovative and include proxies for unobservable characteristics. For example, Borland and Hunter (2000) included whether the person voted at the most recent election as an indicator of civic engagement, and Hunter and Gray (2001) included volunteer work as an indicator of social capital. The authors of these studies conclude that the significance of these types of variables suggests that they should be included in models where possible.

Box 2.3 **Empirical studies on Indigenous labour market outcomes**

There is extensive research on the factors associated with labour market outcomes (LMOs). Empirical studies vary in terms of the data source, modelling approach, LMOs of interest and variables used. A common approach is to use binary or multinomial discrete choice models to estimate the marginal effects of explanatory variables on categories of LMOs, usually labour force participation, unemployment and employment.

Daly (1995) used 1991 Census data and a multinomial logit regression model to examine determinants of labour force status of both Indigenous and non-Indigenous men and women, using factors such as age, education, dependent children and location of residence. Being Indigenous was found to substantially decrease the probability of being in full or part-time employment and increased the probability of being unemployed or not in the labour force.

Borland and Hunter (2000) used 1994 NATSISS data and a two stage equations approach to examine the effect of skills attainment, family structure, location and socioeconomic variables on employment status. In particular, the study focussed on the interaction between history of arrest and employment status.

Biddle and Webster (2007) used the 2002 NATSISS and a set of binary probit models to examine the effect of factors associated with employment and unemployment. In addition, a multinomial probit model was estimated with the dependent variable consisting of four LMOs, employment, unemployment, CDEP participation and not in the labour force.

Hunter and Daly (2008) used the 2002 NATSISS and a binary probit model to examine the effects of arrest and fertility on the probability of labour force participation (LFP) of Indigenous women. A sequential two stage regression model was used to address simultaneity bias that might arise by causation in both directions between both arrest and LFP and fertility and LFP.

Overcoming Indigenous Disadvantage: Key Indicators 2009 (SCRGSP 2009) used 2006 Census data and binary logit models to examine the effect of education and non-education factors, (such as personal and family characteristics and location), on labour force participation. These effects were estimated separately for Indigenous and non-Indigenous men and women.

Stephens (2010) used 2002 NATSISS data and a multinomial logit model to estimate the marginal effects of a number of variables on four LMOs. These variables related to location, demographic characteristics, education, health, culture, criminal history and housing issues, such as overcrowding. The results suggested there was strong association between employment and socio-cultural factors such as living in a mixed race household, speaking an Indigenous language and cultural participation.

Biddle and Yap (2010) used Census data to examine differences in associations between demographic and other variables and LMOs (including income).

In the 2008 NATSISS data, there are some reasonable proxies for unobserved personal characteristics, community engagement and culture that might reduce omitted variable bias in the model and lead to better numerical estimates of the associations between various factors and LMOs. These are discussed in more detail in chapter 4.

3 Econometric model

Regression analysis is a statistical tool that is used to measure the links between an explanatory variable and a dependent variable while holding all other variables constant. This chapter describes some of the technical aspects of the regression analysis using the 2008 National Aboriginal and Torres Strait Islander Social Survey (NATSISS) dataset.

The outline of this chapter is as follows: some of the benefits of multivariate analysis are outlined in section 3.1; the econometric model, and how the results can be interpreted, are explained in sections 3.2 and 3.3, respectively; econometric issues that should be considered in interpreting the results are discussed in section 3.4.

3.1 Benefits of regression analysis

Multivariate regression analysis is applied in this paper — the advantages of this technique, compared to bivariate analysis are presented in box 3.1. Notwithstanding the qualifications (discussed in section 3.4), multivariate regression analysis is likely to produce more accurate numerical estimates of relationships between explanatory variables and a dependent variable than those obtained from bivariate analysis, such as cross-tabulations of data.

Box 3.1 **Bivariate analysis vs. multivariate regression analysis**

Regression analysis allows the effects of multiple factors to be identified separately. Bivariate analysis only measures associations between two variables and therefore does not account for the concurrent effects of other relevant factors on the outcome of interest. For example, a cross-tab shows that the labour force participation rate of people with non-school qualification is higher than that for people without a non-school qualification:

	Without a non-school qualification	*With a non-school qualification*	*All people*
Labour force participation rate (per cent)	55.8	80.5	65.0

Source: Productivity Commission estimates based on the NATSISS (2008)

One interpretation of these data is that increasing the number of people with a non-school qualification would increase the labour force participation rate. However, this interpretation does not account for the influence of non-school qualification type, nor does it take into account other determining factors, such as individuals' gender, location or innate ability. Investigating the influence of multiple factors on LMOs using bivariate analysis is difficult and often not very informative — a crosstab that examined multiple and separate subgroups would quickly become unwieldy, and results would be difficult to present and interpret.

Alternatively, bivariate analysis can be performed with a simple regression model using labour force participation as the dependent variable and non-school qualification as the explanatory variable. Using the same data as the crosstab, the model predicts that changing educational attainment from having no non-school qualification to having a non-school qualification increases the probability of labour force participation by approximately 24.7 percentage points. This is the same as the difference between the labour force participation rates of the two education based subgroups, and similarly, should be interpreted as an association rather than a causal change.

The addition of other (control) factors to such a regression model (such as gender and location) is relatively simple and the results relatively straightforward to interpret. This is known as multivariate regression analysis.

Multivariate regression analysis allows the associations between factors of interest with LMOs to be identified separately, and thus is likely to improve the accuracy of estimates of the size of the associations compared to bivariate analysis. This could be, for example, the association between the probability of an individual being employed and their educational attainment, while controlling for their age and other factors that might influence their employment prospects. It also allows for different aspects of particular factors, such as levels of education, to be investigated.

3.2 Model selection

The choice of regression model depends on the data available. In this case, the data are from the NATSISS survey, which provides information on many characteristics, including labour market outcomes (LMOs). The LMOs considered in this analysis — mainstream employment, Community Development Employment Projects (CDEP) participation, unemployment and not in the labour force — are discrete (non-continuous), unordered (for example ‘not in the labour force’ is not ranked higher or lower than being ‘unemployed’) and mutually exclusive. There are two possible modelling options:[1]

- estimate each outcome independently, using a set of four binary models
- estimate the four outcomes in the same multinomial model.

The advantages of each approach are outlined in box 3.2.

Box 3.2 **Comparing multinomial and binary models**

The major advantage of the multinomial model is that it allows for consideration of all labour market outcomes (LMOs) jointly. In a multinomial model, the sum of the predicted probabilities of the (mutually exclusive) LMOs equals one for the base person. This is not likely to hold when four separate binary models are estimated independently.

Another advantage is that the results for each equation are directly comparable because they are based on the same sample of individuals (and therefore the values at which the other variables are held are the same) and the parameters of the model are estimated simultaneously.

However, estimating several regressions simultaneously using a multinomial model requires consideration of whether or not the error terms in the equations move together (or co-vary).

In practical terms, multinomial models take more time to estimate, and the presentation and interpretation of the results are more difficult because of the greater complexity of the model. Determining whether or not a variable is significant is also complicated by the appearance of the same variable in multiple equations — the variable may be significant in one equation, but not in others.

[1] There are other modelling options, such as a nested model, but these are more difficult to estimate and unlikely to produce significantly different results.

A simplified representation of the multinomial model is:

$$LMO = c + \beta_1 Health + \beta_2 Education + \beta_3 Personal + \beta_4 Crime + \beta_5 Other + \varepsilon$$

Where:

- LMO is a vector of the four labour market outcomes (mainstream employment, CDEP participation, unemployment and not in the labour force)
- Health is a vector of variables representing self-assessed health status
- Education is a vector of variables representing educational attainment
- Personal is a vector of control factors, demographic and other characteristics (including variables such as age, marital status, labour force experience and locational factors)
- Crime is a vector of variables representing history of imprisonment or arrest
- Other is a vector of indicators relating to an individual's social and cultural environment
- c is a constant and ε is the error (residual) term in the model.

The choice of explanatory variables in the model is discussed in chapter 4.

Probit or logit?

This analysis uses a multinomial model in a similar way to Biddle and Webster (2007), Hunter and Gray (2001), and Stephens (2010). In order to estimate a multinomial model, an assumption must be made about the distribution of the error term. If it can be assumed that the error terms of the four equations are unrelated (zero-covariance), the preferred model would be a multinomial logit (MNL) model.[2] If this assumption is not appropriate, a multinomial probit (MNP) model, which assumes the error terms to be multivariate normal, could be used. Allowing the error terms to co-vary means that MNP models are often computationally more difficult to estimate using maximum likelihood estimation than MNL models.

The MNL model requires the assumption of the Independence of Irrelevant Alternatives (IIA). There are logic and statistical tests to determine the validity of

2 See Greene (2003) for more information on the assumed distribution of the error terms in MNL models.

this assumption. While one statistical test was conducted,[3] and indicated that the IIA assumption might be appropriate, Fry and Harris (1998) suggest that the test is biased towards this conclusion. Therefore, greater weight was placed on the logic test, which did not support the use of the MNL model (box 3.3), and so a MNP model was used.[4]

Box 3.3 **The IIA assumption**

The Independence of Irrelevant Alternatives (IIA) assumption is that the ratio of probabilities between outcomes does not change with the introduction of another choice. In other words, the *relative* probability of one alternative does not depend on the existence of other alternatives.

To investigate whether the IIA assumption is likely to hold in the context of this analysis, the following hypothetical situation is considered. Initially, the possible labour market outcomes for a particular group of people are mainstream employment, unemployment, or non-participation in the labour force. Then the option of participation in the CDEP scheme is introduced to this group. For IIA to hold it would be necessary for the ratio of probabilities for each outcome in the initial choice set to remain the same once CDEP becomes available. This would require that proportionally equal numbers of the group would move from each of the three labour market outcomes into CDEP employment.

However, there is reason to believe that this might not happen. According to Stephens (2010), the characteristics of Indigenous people who are unemployed and the characteristics of those participating in the CDEP scheme are similar. This would indicate that a disproportionately large number of people would transition from unemployment to CDEP when CDEP is added to the choice set. Some people who are not in the labour force may also transition, but it is unlikely that many in mainstream employment would do so. If this is true, it would invalidate the IIA assumption in the context of this analysis, and potentially weaken any conclusions drawn from the estimation of a MNL model.

3.3 Interpretation of results

Multinomial models are non-linear in the parameters and therefore coefficients are not readily interpretable. It is more common to report the marginal effects of

3 The Hausman test was conducted (with each outcome in turn omitted from the choice set). Other tests were not possible with the ABS's Remote Access Data Laboratory.

4 A comparison of the results from the MNP and MNL models indicated the difference in the magnitude and significance of the marginal effects was relatively small for most variables.

variables, which are derived from the estimated coefficients and predicted probabilities.

Base predicted probability — the probability of a 'base person' (described in box 3.4) being in a particular LMO. The base person in this study is someone who is 37 years old, is married, lives in a non-remote area, is in good health, has a year 10 or 11 education and no non-school qualifications, has not been arrested in the last five years and has never been in jail.

Marginal effects — the change in the value of a dependent variable (in this case, the base predicted probability of a LMO) that is associated with a marginal change in an explanatory variable, holding all other explanatory variables constant. The results are presented as percentage point changes relative to the base predicted probability.

Marginal effects are generally calculated with continuous variables held at mean values and binary variables held at the mode (most common) values, consistent with the characteristics of the base person.[5]

In interpreting the marginal effects, the following approaches were used:

- For binary variables, other than the education variables, the marginal effects represent the percentage point change in probability of a LMO associated with a change in a characteristic variable from '0' to '1', holding the value of all other explanatory variables constant.[6]
- For the education variables, the marginal effects represent the percentage point change in the probability of a LMO associated with an increase in education compared to having year 10 or 11 but no non-school qualification, holding the value of all other explanatory variables constant.

5 The base man and woman are married, even though a majority of women in the sample are not married.

6 The convention in econometric analysis is to define binary variables in terms of the least common characteristic. For example, the variable indicating whether someone has a degree is defined as 'has a degree'. This description is chosen because most people do not have a degree (DEGREE = 0). The marginal effect associated with having a degree is the change in the predicted probability of a LMO resulting in a change from not having a degree to having a degree (DEGREE = 1). Some variables are defined in terms of not having the characteristic (does not provide support, is not married) because the most common characteristic for people in the sample is that they provide support, and are married. The definitions are intended to provide a consistent basis for interpretation rather than having any connotation of desirability about being in one category over another.

- For the continuous variables, the marginal effects represent the percentage point change in the probability of a LMO associated with a small increase in the variable from its mean value, holding the value of all other explanatory variables constant (known as the instantaneous rate of change in the predicted probability
- associated with the explanatory variable). These values are reported in the appendix (tables A.3 and A.4).
- For age and work experience, a squared term is included to account for a non-linear relationship with LMOs.
  - For work experience, the total marginal effect was calculated as a one year increase from the 25th percentile, mean and 75th percentile values to facilitate the interpretation of these results.

Statistical significance

Statistical significance tests are used to gauge the reliability of estimates. In the results tables, the stars next to the estimated marginal effects represent the level of statistical significance based on individual Wald tests. (In the attachment tables, the standard errors are also reported.) One, two and three stars represent significance at the 10, 5 and 1 per cent levels, respectively.[7] In the charts, error bars indicate 95 per cent confidence intervals for the estimates.

If an explanatory variable is determined not to be statistically significant (for example, for some of the results with no stars) it does not necessarily mean that there is no relationship between the variables. Rather, there is not sufficient evidence, based on the survey sample, to indicate that a relationship exists. For example, if there are relatively few people in a sample with a particular characteristic, or in a particular outcome category, it may be difficult to detect a statistically significant association between variables. This is particularly an issue for the CDEP and unemployment outcomes in this model, because the number of people in these categories in the data set is relatively small.

3.4 Econometric issues

The regression results provide estimates of the sign and magnitude of relationships between the LMOs and the explanatory variables. However, several qualifications need to be understood in interpreting and using the results of these kinds of models.

7 If a marginal effect is significant at the 5 per cent level (two stars), then there is 95 per cent certainty that there is a (non-zero) relationship between the factor and the LMO.

Box 3.4 **Characteristics of the base person**

In this analysis, the characteristics of the base person were chosen so that they would be most representative of the NATSISS sample. The base person:

- is around 37 years old
- is married[8]
- has 1.2 dependent children (for women), and 0.8 (for men)
- lives in a single family household
- lives in a non-remote area
- does not live in his/her homelands
- lives in a household located in an area with a low SEIFA index of relative disadvantage (in decile 3)
- has no difficulty communicating in English
- is in good general health
- has low levels of psychological distress
- does not have a severe or profound disability
- has a year 10 or year 11 education and no non-school qualification
- has around 15 years of work experience for men and 11 years for women
- has not been arrested in the past five years,
- has never been in jail
- did not participate in three or more types of cultural activities in the last 12 months (such as fishing, hunting, gathering or making traditional crafts)
- did not attend three or more types of cultural events in the last 12 months (such as ceremonies, festivals, NAIDOC week activities)
- provided support outside the household in the last 12 months
- did not provide unpaid childcare outside the household in the last 12 months.

8 Most women (52 per cent) in the estimation sample were unmarried and most men (54 per cent) were married. For ease of comparison of the results, the reference man and woman are both married, which is the most common category overall.

These qualifications are particularly important when seeking to use the results to quantify the increase in the proportion of the Indigenous population in a particular LMO, say employment or labour force participation, that might be expected from a change in an explanatory variable (for example, from meeting COAG targets for year 12 attainment). A large and significant marginal effect for an explanatory variable does not necessarily mean that the explanatory variable causes the LMO, or that a change in a particular factor will necessarily result in a change in the probability of a LMO of the magnitude implied by the marginal effect.

Some fundamental assumptions of regression analysis relate to:

- the coverage of the survey (sample selection bias)
- whether factors not included in the model influence LMOs (omitted variable bias)
- the direction of causality between the dependent and explanatory variables (endogeneity due to simultaneity)
- whether self-assessed health status is measured accurately (endogeneity due to rationalisation)
- whether the explanatory variables are related to each other (multicollinearity).

Sample selection bias

Sample selection bias occurs when the characteristics of the survey respondents are different from those of the in-scope population. When each person in the relevant population has an equal chance of being selected for the survey, the survey sample is described as being an 'equal probability of selection' sample design.[9] In practice, this is difficult, and most surveys have various forms of sample selection bias arising from some groups having different probabilities of being selected in the sample, or from undercoverage where particular groups have no probability of being selected. In particular, people in non-private dwellings (including prisons) were excluded from the 2008 NATSISS. The NATSISS has a relatively high level of undercoverage[10] and potentially different selection probabilities for population

---

9 See table A.7 for the sample size.

10 The undercoverage rate reported by the ABS is 53 per cent for the 2008 NATSISS. This is higher than the undercoverage rates reported by the ABS for the 2004-05 National Aboriginal and Torres Strait Islander Health Survey (42 per cent) and the Monthly Population Survey (12 per cent). More information on NATSISS sampling and non-sampling errors can be found at ABS (2010b). Some of the other significant factors contributing to undercoverage relate to

groups, for example, in remote and non-remote areas. This is likely to result in the estimated results being biased due to sample selection.

Sample selection bias is often corrected by applying weights to the data.[11] Weights have the effect of repeating observations that are representative of an under-sampled group of people (ABS 2010b). It is accepted practice to weight sample descriptive statistics (such as means), but it is less clear whether weighted data should be used in regressions (Gelman 2007). Instead, the factors that influence the sample selection bias are often included as explanatory variables in the regression. In this analysis, the variable that might address sample selection bias is remoteness.

In this study, weighted data were not used because the added level of complexity makes it difficult to interpret results. This means that the estimated results are best described as representing the associations for individuals in the sample, and results cannot necessarily be extrapolated to the whole Indigenous population. Nonetheless, results for a particular type of individual are likely to be valid for other individuals displaying the same characteristics.

A comparison of the results using unweighted data showed that application of (person) weights:

- increased the standard errors of many variables and therefore reduced their statistical significance
- changed the size of the marginal effects by small amounts.

Endogeneity due to unobserved heterogeneity (omitted variable bias)

A model's results may be biased when the dependent variable and an explanatory variable are linked by a third variable that is not included in the model. This results in unobserved heterogeneity or omitted variable bias. An example is education and LMOs. A person's educational attainment and LMOs may both be influenced by personal attributes, such as motivation, aptitude and preferences, some of which cannot readily be captured by surveys. Omitting these attributes from the model could result in the marginal effects for education (and possibly other explanatory variables) on LMOs being biased, since they might capture, in part, the effects of the omitted attributes.

non-identification of Indigenous people, non-response to survey questions and errors identifying some private dwellings.

11 This includes applying a geographical adjustment to initial weights to account for different undercoverage rates and the characteristics of Indigenous people living in these areas and calibrating the weights with population benchmarks (ABS 2010b).

Results of other studies (Cai 2010; Laplagne et al. 2007) — using models of labour force participation and health, and Australian survey data — support the hypothesis of unobserved heterogeneity, especially for females, and concluded that the marginal effects derived from models that do not adjust for this kind of bias are likely to be upper bound estimates.

Endogeneity bias due to simultaneity

Results may also be biased when the direction of causality runs both ways between the dependent variable and an explanatory variable. An example is health and LMOs. People who are in good health are more likely to be able to work, but it is also true that working could affect a person's health (in some cases positively; in other cases negatively). This means that a person's health affects their LMO, but their LMO also affects their health. This form of endogeneity bias is known as simultaneity.

The impact of this type of bias on model estimates is unknown. The marginal effects could under or over-estimate the true effect of a change in health on LMOs, depending on whether the relationship between LMOs and health is negative or positive.

An alternative type of model, such as a simultaneous equations model (SEM), might be considered where dependent and explanatory variables are likely to be interdependent. SEMs can be used to test for interdependency between a dependent and explanatory variable. If necessary, these models can also correct for the endogeneity bias that would otherwise arise from a simpler, single equation model.

Some studies have found evidence of endogeneity in models of labour force status and health. Cai (2010) found that good health had a positive and significant effect on labour force participation for men and women. It was reported that for men, there was a negative and significant reverse effect (of labour force participation on health), implying that treating health as an exogenous variable could lead to underestimates of the effect of health on labour force participation.

Laplagne et al. (2007), using a SEMs model, also found evidence of simultaneity affecting the relationship between self-assessed health status and labour force participation for men aged 15–49 years.

Endogeneity bias due to rationalisation

Laplagne et al. (2007) note that it is difficult to determine whether the relationship between labour force participation and self-assessed health results from a simultaneous relationship between the two, or from rationalisation endogeneity. Rationalisation endogeneity exists where self-assessed health status does not reflect true health status. For example, people might misreport their health status to justify non-participation in the labour force. Unlike simultaneity bias arising from interdependent variables or unobserved heterogeneity, rationalisation endogeneity cannot be corrected (Laplagne et al. 2007).

Laplagne et al. (2007, p. 44) found:

> For older females, the coefficient on labour force participation is positive and significant, which suggests rationalisation endogeneity might be present … Although rationalisation might be a cause of the positive coefficient for older females, it could also reflect positive 'true' endogeneity (simultaneity).

This result is consistent with Cai (2010), reporting a positive and marginally significant relationship between labour force participation and self-reported health status for women. Cai (2010, p. 84) suggests that this may be evidence of rationalisation endogeneity, although it is also hypothesised that:

> … the positive estimate for the labour force status variable for females may not be due to justification, rather it may be due to self-selection into labour force status and the selection of jobs by women. That is, those women who choose to be in the labour force are in good health and they are in jobs that are less likely to harm their health.

Multicollinearity

Multicollinearity occurs when explanatory variables are highly correlated with each other. This might occur when people suffer multiple disadvantage (for example, poor health, poor education and a criminal record) and the factors associated with disadvantage are all included in the model. Multicollinearity will not reduce the predictive power of the model as a whole, but the magnitude of individual coefficients may be unreliable, standard errors of their associated coefficients may be inflated, and as a result some explanatory variables might appear insignificant when they are not (Greene 2003).[12]

---

[12] There is a trade-off involved in balancing the potential impacts of multicollinearity by excluding some highly correlated explanatory variables, against the possibility of introducing omitted variable bias.

4 Data and variables

Econometric models of the determinants of labour market outcomes (LMOs) are generally based on a demand and supply framework, and the various theories (described in chapter 2) that guide the choice of variables. Empirical labour market models commonly include basic features or characteristics of individuals, such as their age, marital status, education and work experience. More detailed datasets, such the National Aboriginal and Torres Strait Islander Social Survey (NATSISS), include indicators related to health, and cultural and socioeconomic environment that also influence LMOs.

Theory may be used to predict whether a particular factor has a positive or negative influence on a LMO, or, in technical terms, the expected sign of the marginal effect of an explanatory variable. However, in some cases, different theories will suggest different expected signs. Some of the variables in this model are chosen based on more than one theory, and the expected sign of the marginal effect is, *a priori*, ambiguous. It is then an empirical matter as to whether the positive or negative effect dominates.

The outline of this chapter is follows: the 2008 NATSISS dataset and the sample used is described in section 4.1; the variables that are used in this analysis are described in section 4.2; and some of the data limitations are described in section 4.3.

4.1 ABS 2008 NATSISS

This analysis was conducted using the 2008 NATSISS dataset. This is the third survey of this type to be conducted by the ABS and includes Indigenous people who were permanent residents of private dwellings. The NATSISS excludes people who are not permanent residents of private dwellings (including visitors), and people living in institutions (such as hospitals, nursing homes and, prisons) and special dwellings (such as hotels and boarding houses).

The NATSISS dataset has information relating to characteristics of the household, as well as individual characteristics of at least one adult and one child who are residents of that household. For remote areas, information relates to two adults and two children in each household.

Estimation sample

The model excludes full-time students because they are limited in the extent of their possible involvement in the labour force. The estimation sample was further restricted to people aged between 15 and 64 (the usual working age population).

Some labour supply factors are highly deterministic, for example, the labour supply decisions of men and women tend to be different due to factors such as child rearing and other family responsibilities. This supports estimating separate regressions for men and women.

After accounting for these exclusions and for missing observations, the estimation samples were around 2700 men and 3500 women.

4.2 Explanatory variables

This section outlines the variables used in the analysis in this paper and explains the theoretical reasons for their inclusion. It also explains some measurement issues related to the use and derivation of particular variables, including approaches taken in economics literature. The variables presented in this section include gender, age and family structure of the household, and others that fall under the following categories: location, health, skills and knowledge, contact with the criminal justice system, and social and cultural environment. Descriptions of the variables used are in table A.1. Descriptive statistics are in table A.2.

Gender

Patterns of LMOs tend to differ between men and women. Women are more likely to balance paid work and child rearing and other family responsibilities. They often leave the labour market to care for young children, while men rarely leave the labour market to care for children for long periods (Orzechowska-Fischer 2004). While many women return to the labour market within a year of having a child, others return when their children reach school age, or they never return (Boushey 2005; Goldin 2006). As discussed in chapter 1, in 2008, Indigenous labour force participation (as a proportion of the Indigenous working age population aged 15 to 64) was 75 per cent for men, and 55 per cent for women (SCRGSP 2011).

For these reasons, this analysis uses separate models for men and women. This approach is common in the literature (Hunter and Gray 1999; Stephens 2010) although some researchers have used gender as an explanatory variable in their models (Biddle and Webster 2007).

Age

Engagement with the labour market is influenced by a person's age because there are life-cycle effects on their preferences for labour force participation and the value of a person's human capital. Young people might not participate in the workforce and be employed until after completion of education or skills training. As people age, labour force participation and employment may decrease as their health and physical ability to work declines. Eventually, people retire and leave the labour market. Some effects which are linked to age, such as studying, health and physical ability, are separately accounted for by excluding students and by explicitly including other factors in the model. The inclusion of age variables therefore reflects the influence of other stages in a person's lifecycle on LMOs, such as skills formation, child rearing and early retirement.

A mathematical function that captures the effect of age on labour market outcomes is a quadratic function, comprising age and age squared (Bartus 2005). The total effect of age on a LMO is a combination of the age and age-squared effects evaluated at the mean age (and the mean age squared) of the working population.

Other studies have used a different specification for age in econometric models, including dummy variables to reflect age cohorts (Biddle and Webster 2007; Cai 2010; Forbes et al. 2010; Laplagne et al. 2007; Stephens 2010). Other researchers have included an age-cubed term for women (Shomos 2010).

Family structure and household characteristics

Family structure is a significant determinant of a person's household responsibilities which in turn affects their availability and incentives to participate in the labour market. Family relationships can be characterised as bonding social capital and can either increase or decrease labour market activity (see box 2.2). For example, preferences for, or commitments to undertake unpaid work make it less likely that a person will participate in the labour force and or be in full time employment. At the same time, such arrangements might support other family members to participate in paid work. The total utility of the household may be maximised when those household members who have the ability to earn the highest wages specialise in paid work while other household members specialise in unpaid work.

The direction of these effects is ambiguous because the existence of a particular family structure does not indicate whether an individual's relationships and responsibilities are likely to either impede or encourage labour force participation. The inclusion of variables related to family structure, in part, controls for these relationships. It becomes an empirical matter whether the associations between these variables and LMOs is positive or negative, and whether the associations differ markedly for men and women.

Marital status

The responsibilities associated with marriage or being in a de-facto relationship might change incentives to participate in the market for paid work if couples specialise between paid and unpaid work to maximise household utility. Other couples may choose to share responsibilities. Being married or in a de facto relationship might also be less of an influence on LMOs *per se*, but rather a proxy for a number of characteristics that are desirable to both partners and employers (Lattimore 2007).

Being married is generally found to be positively associated with employment and labour force participation for men, but has been found to have both positive and negative associations for women in different studies (Borland and Hunter 2000; Hunter and Gray 1999).

Dependent children

The presence of young children in a household or family indicates that there may be constraints on the household members participating in the labour force because of child rearing responsibilities, especially when those children are below school age. Having one or more dependent children has generally been found to be associated with a lower probability of employment for women (Hunter and Gray 1999). For men, the association between having one or more dependent children and the probability of employment is less clear — for example, Hunter and Gray (1999) found a positive association, while Stephens (2010) found a negative association. This study estimates the number of dependent children by using information about the number of children in the household and family structure.

Multi-family household

The presence of more than one family living in a person's household can have a number of influences on their preferences for different types of work. It might mean that the demands of unpaid work (carer responsibilities and domestic work) are greater than those in a single family household, which can constrain the ability of some household members to participate in the labour market. On the other hand, such household structures may provide support for those who are best placed to specialise in paid work. In such cases, the influences of living in a multifamily household on labour force participation might be similar to marriage relationships.

However, the incidence of a person living in a multifamily household might also be an outcome of being either unemployed or not in labour force, as people who do not earn an income might live with other families due to a lack of financial resources. Therefore the association between this variable and LMOs may be subject to a degree of endogeneity. It may also indicate that relatively fewer resources are available to each household member that may improve their employment prospects, such as access to education or transport. Therefore this variable could be considered a proxy for disadvantage.

Unpaid childcare

Some people provide unpaid childcare outside their household. This might constrain their ability to participate in the labour market for paid work. The offsetting benefit of the support, that might increase the child's parent or carer's labour market activity, will accrue to someone else.

Location

Where a person lives might affect their LMOs, which reflects both labour supply and labour demand factors.

Remoteness

Remoteness, as defined by the ABS Australian Standard Geographical Classification (ASGC) (ABS 2010a), is based on the physical distance by road to the nearest urban centre. Remote areas tend to have lower levels of economic activity and access to services. Consequently labour demand may be lower in remote areas relative to less remote areas, or there may not be much diversity in the type of jobs available or the range of skills required to perform those jobs. Education opportunities and health services may also be less accessible, resulting in lower levels of human capital for people living in remote areas.

Some studies have used multiple variables to indicate degrees of remoteness (Biddle and Webster 2007; SCRGSP 2009). Due to data constraints, this study uses a binary remoteness indicator based on the ASGC (remote and non-remote).[1]

Indigenous people who live in remote areas are less likely to be in mainstream employment and more likely to participate in CDEP (Biddle and Webster 2007; Borland and Hunter 2000; SCRGSP 2011). Labour force participation and unemployment among Indigenous people tends to be lower in remote areas (due to CDEP participation being classified as employment by the ABS).

Lives in homelands

Living in homelands is included in the model for a number of reasons. For example, it might indicate:

- a preference for participating in traditional activities outside the mainstream labour market
- that a person lives in a remote area with low levels of labour market activity[2]
- that a person is reluctant to move locations to obtain work.

These factors are expected to contribute to poor LMOs. This is supported by Stephens (2010) who found that for males, 'living in homelands' had a negative marginal effect on being employed. Nonetheless, the marginal effect of living in homelands on employment was positive for women living in non-remote areas, but negative for women living in remote areas (Stephens 2010).

Socioeconomic status

The socioeconomic status of the area in which people live could influence their LMOs because they have poor access to material and social resources, reducing their capacity to participate in society. This is the definition used by the ABS to define relative socioeconomic disadvantage (ABS 2006).

[1] The dataset that was available at the time this study was undertaken only had a binary indicator of remoteness. Another dataset has since become available with five classifications of remoteness (major cities, inner regional, outer regional, remote and very remote).

[2] Some people who live in non-remote areas may report that they live in their homelands if the place they live in is part of their traditional country, even though much of it is likely to be owned or occupied by non-Indigenous people. Indigenous people who live in remote areas were more likely to report that they lived in their homelands than those who lived in non-remote areas (SCRGSP 2011).

The NATSISS includes data on the decile of the socioeconomic index of relative disadvantage of the household (1 is most disadvantaged, 10 is least disadvantaged). This index is derived from Census variables related to disadvantage, such as low income, low educational attainment, unemployment, dwellings without motor vehicles and households that have one parent, pay low rent or rent from government and community organisations. Importantly, the index incorporates the proportion of the local community who are Indigenous and therefore might introduce additional endogeneity.

Health and disability

Health and disability are elements of human capital that are often included in models of LMOs. This is because good physical and mental health are attributes valued by employers, and disability can affect people's ability to participate in the workforce and be employed. Health and disability might also affect a person's labour supply decision (Cai 2010). This may be because a person decides to supply less labour because of poor health or disability, or more labour to be able to pay for health treatments.

The measurement of health is difficult in surveys. A common approach is to use self-assessed health status, although this may result in rationalisation endogeneity (box 4.1). There are also a number of difficulties measuring disability, especially for Indigenous Australians. More information on the extent of disability amongst Indigenous Australians can be found in PC (2011).

Good physical and mental health tends to be positively related to labour force participation with healthy workers generally having higher productivity and greater incentives to work. Having a disability is associated with a lower probability of labour force participation. Laplagne et al. (2007) cite analysis that suggests that poor mental health, in particular, is associated with a lower probability of labour force participation than good mental health. Indicators of health status that have been used in some models of Indigenous LMOs include self-assessed general health, disability and mental health status, as well as health risk factors such as smoking and alcohol consumption. Stephens (2010) found there was a reduced probability of employment and labour force participation associated with smoking, having a disability and poor self-assessed health status.

Causality between health and labour market outcomes can occur in both directions, resulting in potential endogeneity bias. Labour force participation can affect a person's health positively or negatively. It may increase a person's general health and wellbeing, or lead to deterioration in health due to stress, fatigue or physical demands, potentially leading to withdrawal from the labour market. The potential simultaneity bias that arises from complex interactions between health and labour force participation has been accounted for using simultaneous equations models (Cai 2010; Laplagne et al. 2007).

Box 4.1 **Measuring the effects of health status for labour market research**

One issue that arises in studies of the effects of health on participation, productivity and wages is the measurement of health status. Some researchers use data based on formal diagnosis of particular medical conditions. For example, the 2003 HILDA survey asked respondents:

> Have you ever been told by a doctor or nurse that you have any of the long-term health conditions listed below? [The list of conditions included arthritis, asthma, cancer, chronic bronchitis, emphysema, diabetes, heart disease and high blood pressure] (AC Nielsen 2003, p. 10)

Other studies rely on individuals' self-reported general health. Self-reported general health can be derived from direct responses to survey questions regarding a person's health status. For example, the NATSISS asks respondents whether 'in general' they would say their health is 'excellent, very good, good, fair, or poor'.

In the context of labour market research, this kind of health measure can be prone to 'rationalisation endogeneity', which occurs when a person uses their self-assessed health as a rationalisation for their labour market status. Cai and Kalb (2005) found mixed evidence of rationalisation behaviour in previous studies, and also found that self-assessed health status is highly correlated with diagnosed conditions.

Alternatively, measures of general health can be derived from responses to questions about how well people are able to perform certain tasks (such as climbing stairs and carrying groceries) and how they feel (for example, 'how much bodily pain have you felt during the past four weeks').

Source: Forbes et al. (2010); ABS (2010b).

The use of self-assessed health status as a health indicator can result in rationalisation endogeneity where a person misreports their true level of health possibly to justify non-participation in the workforce (Laplagne et al. 2007).

This study uses self-assessed health status, psychological distress as measured by the Kessler scale and whether the respondent has a severe or profound disability as indicators of health and disability.[3]

[3] See table A.1 for definitions.

Skills and knowledge

Human capital theory, discussed in chapter 2, suggests that higher levels of skills and knowledge improve employment prospects. These attributes may not always be observable, and employers look for observable characteristics that indicate a person's level of skills, such as education, English language skills and work experience, to assess the suitability of an employee.

Education

There are two main reasons why education influences LMOs. First, education develops academic, vocational and social skills that are valuable to employers. Second, education acts as a signal to employers that a person has skills that are for the most part unobservable, including ability and motivation. This information reduces job-hiring costs for employers.

Previous studies find that people with higher education levels (especially degrees and other non-school qualifications) are more likely to be in the labour force, less likely to be unemployed and for those who are employed, less likely to participate in the CDEP program (Biddle and Webster 2007; Borland and Hunter 2000; Stephens 2010). The magnitude of the association between education and LMOs is different for men and women (the marginal effects for women tend to be larger and more consistently statistically significant).

There are many ways education can be specified in econometric models, including years of education, levels of education, or pathways through the education system. This study classifies people into one of seven possible categories of pathways though the education system, depending on their highest level of schooling and non-school qualification. This approach was chosen because Indigenous people tend to engage differently with the education system compared to non-Indigenous people. For example, Indigenous people have:

- higher attainment of certificate levels 1 and 2 courses
- higher attainment of certificate 3 and 4 courses by people aged over 35
- higher participation in Technical and Further Education Institutions or vocational education and training (SCRGSP 2011).

This suggests that Indigenous people tend to obtain lower levels of non-school qualifications, and do so at a later age, possibly without having completed year 12 compared to non-Indigenous people. Biddle and Yap (2010) found that Indigenous Australians aged 25 years and over had a significantly higher probability of attending education than non-Indigenous Australians with otherwise identical (observable) characteristics.

English language skills

Employers value English language skills. Studies of LMOs tend to include whether the respondent has a non-English speaking background. This will not always capture English language skills, as many people from non-English speaking backgrounds are able to communicate well in English. This study includes a variable on whether the respondent indicated they had difficulty communicating in English.

Work experience

Skills and knowledge are also enhanced by work experience through on-the-job training and workplace-specific courses. Work experience is a reasonably observable and verifiable attribute valued by employers. Unlike many other studies that use potential work experience (estimated as a person's age less the number of years spent studying), this analysis includes the (self-reported) number of years in the workforce.

Like age, work experience is often modelled using a quadratic function, comprising work experience and work experience-squared terms, and the total effect of work experience on a labour market outcome is a combination of both, evaluated at the mean value of years of work experience (Bartus 2005). The results reported in the appendix tables are for an instantaneous change in work experience and work experience-squared. The results reported in chapter 5 (table 5.1) are for the total effect of an increase in work experience of one year from the mean value (and the value of years of work experience at the 25th and 75th percentiles).

The inclusion of both age and work experience terms in the model raises issues of collinearity, which, as explained in chapter 3, can inflate standard errors, and make explanatory variables appear insignificant when they are not. However, the alternative of excluding work experience or age raises issues about omitting relevant variables from the model. The inclusion of both age and work experience variables can be justified on the basis that a person's age may not be a reasonable proxy for their years of work experience, partly because Indigenous people spend longer periods of their lifecycle not in the labour force than non-Indigenous people. This implies the correlation between the two terms is likely to be lower than for non-Indigenous people (SCRGSP 2011).

A comparison of the results presented in this paper with results using the same model with the work experience terms omitted found that the marginal effects of the education and crime variables were larger in the model that excluded work experience. These differences were relatively larger for women. This suggests that excluding work experience from the model risks, among other things, overstating the magnitudes of the associations between some factors and LMOs.

Contact with the criminal justice system

Contact with the criminal justice system might affect both labour supply and demand. On the demand side a criminal history might affect a person's chances of obtaining employment if a history of arrest or jail is used as screening device by potential employers, or where businesses may be deterred from locating in regions with relatively high crime rates, resulting in reduced job opportunities. On the supply side, a criminal history might influence a person's motivations or perceptions of the benefits of work.

Borland and Hunter (2000) examined the effect of arrest on labour force status by controlling for unobserved characteristics of people arrested in the preceding five years using 1994 NATSISS data. Arrest was found to be associated with a lower probability of employment for both Indigenous women and men (13 per cent and 18 per cent respectively). However these results may be biased if an individual's arrest record is influenced by their employment status, or if there are unobservable factors affecting both their employment status and arrest record. For example, the analysis found a strong positive relationship between alcohol consumption and arrest, so a history of arrest may be an indicator of other characteristics of an individual that may have a negative effect on labour supply. Hunter and Daly (2008) used a two-stage model to address potential simultaneity bias that might arise from the interaction between arrest, fertility and labour force participation.

The analysis in this paper uses two variables related to contact with the criminal justice system — arrest in the last five years and whether the respondent has ever been in jail. These variables are correlated, because in most cases the people who have been in jail have also been arrested. However, some people may not have been arrested in the last five years, but have been in jail in their lifetime. The marginal effects associated with a history of imprisonment cannot, in this study, be interpreted as additional to the marginal effects of having being arrested.

Social and cultural environment

As well as drawing on human capital theory, this paper also considers factors related to social and cultural environment, which social capital theory suggests might influence the LMOs (box 2.2). Attributes, such as the motivation to obtain employment, and preferences for paid versus unpaid work, can be influenced by a person's social and cultural environment. Some studies that include variables relating to social and cultural factors suggest that they reduce omitted variable bias due to their correlation with characteristics which are largely unobservable (Hunter and Gray 2001; Stephens 2010). The theoretical links between living in homelands and providing unpaid childcare and LMOS are discussed above. Some indicators of community engagement, that might be proxies for motivation, are discussed below.

Preferences

Preferences for a particular lifestyle might influence LMOs. For example, some individuals might:

- be part of a household or community where they specialise in, or have a preference for, unpaid work
- prefer, and are able, to live in their traditional lands, which might mean they live in an area with low labour demand or lack of jobs that match their skills, and be unwilling to move to obtain employment
- prefer to live a more traditional lifestyle and can be sustained through traditional activities.

Some of the variables described above can act as indicators of such influences. For example, being married and the presence of dependent children might increase the probability of someone allocating significant time to unpaid work within a household. The 2008 NATSISS contains other information that might reflect a person's preference for a traditional lifestyle. These include:

- whether they recognise and currently live in their homelands (discussed above as a locational variable)
- whether they participate in traditional cultural activities (three or more types in the last 12 months).

It might be expected that a person who allocates a large proportion of their time to unpaid work, or activities associated with a more traditional lifestyle, will place a relatively low value on labour market engagement, and choose not to participate in the labour market for paid work. Hunter and Borland (1997) found that hunting, fishing and gathering activities had no effect on the probability of employment, suggesting there was little substitution between these traditional activities and market work. However, Hunter and Gray's (1999) results suggested there might be some substitution, especially for women. However, these activities might confer other benefits to the individual or the community.

Motivation

An important personal attribute that influences LMOs is a person's motivation. Motivation also affects educational and health outcomes — which in turn affects LMOs — and is influenced by the social environment. It is difficult to observe, and a common approach is to use panel data to control for this and other unobserved characteristics. In the absence of panel data on Indigenous people, proxies for unobserved characteristics can be used. These proxies are characteristics that might be correlated with unobservable attributes.

Employers look for signs of motivation, such as education and work experience. Motivation might also be indicated by a person's engagement with their community. For example, a person who engages with their community through volunteer work or by attending social cultural events, is likely to be motivated in other ways, including to participate in the labour market for paid work. The benefit of including these proxies is that marginal effects for other explanatory variables are likely to be less biased because at least some of the influence of the unobserved attributes on other variables in the model is accounted for.

There are costs and benefits to this sort of community engagement. It could create networks that lower job search costs but it might also become a barrier to labour market engagement where such activities are time consuming. For example, Stephens (2010) found that for males, having attended a cultural event was negatively associated with the probability of employment. Importantly, he noted that the association between LMOs and variables such as cultural participation might be inflated because of a high degree of unmeasured correlation with remoteness. Hunter and Gray (1999) found that participation in volunteer work increased the probability of non-CDEP employment for both Indigenous men and women.

4.3 Some data limitations

While the NATSISS is a rich source of data on the characteristics of Indigenous people, it has limitations, which constrain the explanatory power of the analysis and introduce some possible data measurement errors.

Labour demand is an important determinant of LMOs, but the NATSISS does not have many variables that can be used as reasonable proxies. This study includes variables related to geographic location which can be considered proxies for labour demand. Remoteness is one such indicator, but this is a poor proxy because it can be argued that some people live in non-remote areas where there is weak demand for labour (or for particular skills or occupations).

The SEIFA index is a measure of the relative disadvantage of the collection district the household is located in. This index is derived using a number of variables, including the number of Indigenous people living in the area which may in themselves be the result of employment outcomes and therefore may introduce a degree of endogeneity bias in the model estimates.

Another limitation of the data is the lack of a variable representing the number of dependent children a respondent has. The number of dependent children used in the analysis was derived using variables including household type, family composition and relationships in households, and number of people aged 0 to 14 in the household. Because it is difficult to determine from the data who are parents in multifamily households, and because the NATSISS variable had an upper bound of five children, the estimate of the number of dependent children is likely to be an underestimate.

In a perfect world, survey data would contain information on innate ability, motivation, and preferences which theory suggests are important determinants of labour market outcomes; but in practice these attributes are difficult to measure. The NATSISS contains some information on characteristics that might be correlated with these unobservable attributes. This means that models can include only proxies for these characteristics.

5 Results and discussion

This chapter presents the results of the regression analysis of a dependent variable with four labour market outcomes (LMOs) — mainstream employment, Community Development Employment Projects (CDEP) participation, unemployment and not in the labour force.[1]

The discussion of results is in terms of the associations between an explanatory variable and a category of the dependent variable (that is, a LMO). An explanatory variable's estimated marginal effect is the change in the base predicted probability of a LMO that is associated with a change in the explanatory variable, holding all other explanatory variables constant. Chapter 3 contains a discussion on the interpretation of marginal effects (section 3.3). Most of the factors examined have the expected sign, and many are significant at the 1 and 5 per cent levels of confidence, although few variables were significant for the LMOs 'CDEP participation' and 'unemployment'.

These numerical estimates should be viewed as associations — they do not imply causal relationships. This chapter focuses on the estimated associations between 'mainstream employment' and 'not in the labour force', as these LMOs represent the majority of the survey sample. For ease of expression, the results for 'not in the labour force' are usually discussed in terms of 'labour force participation'.

Results for health, education, personal and demographic characteristics, and social and cultural variables are discussed in sections 5.1 to 5.4. A summary of the key findings is at the end of this chapter (box 5.1). The full set of results can be found in appendix A, including all estimated marginal effects and regression coefficients, and their standard errors, a description of the variables, descriptive statistics and some model diagnostics.

1 Models were estimated with Stata and processed using the ABS's Remote Access Data Laboratory (RADL).

5.1 Associations with health and disability

The association between two elements of health — poor general and mental health — and severe or profound disability, and the probability of an Indigenous person being employed and not in the labour force are reported in figure 5.1.

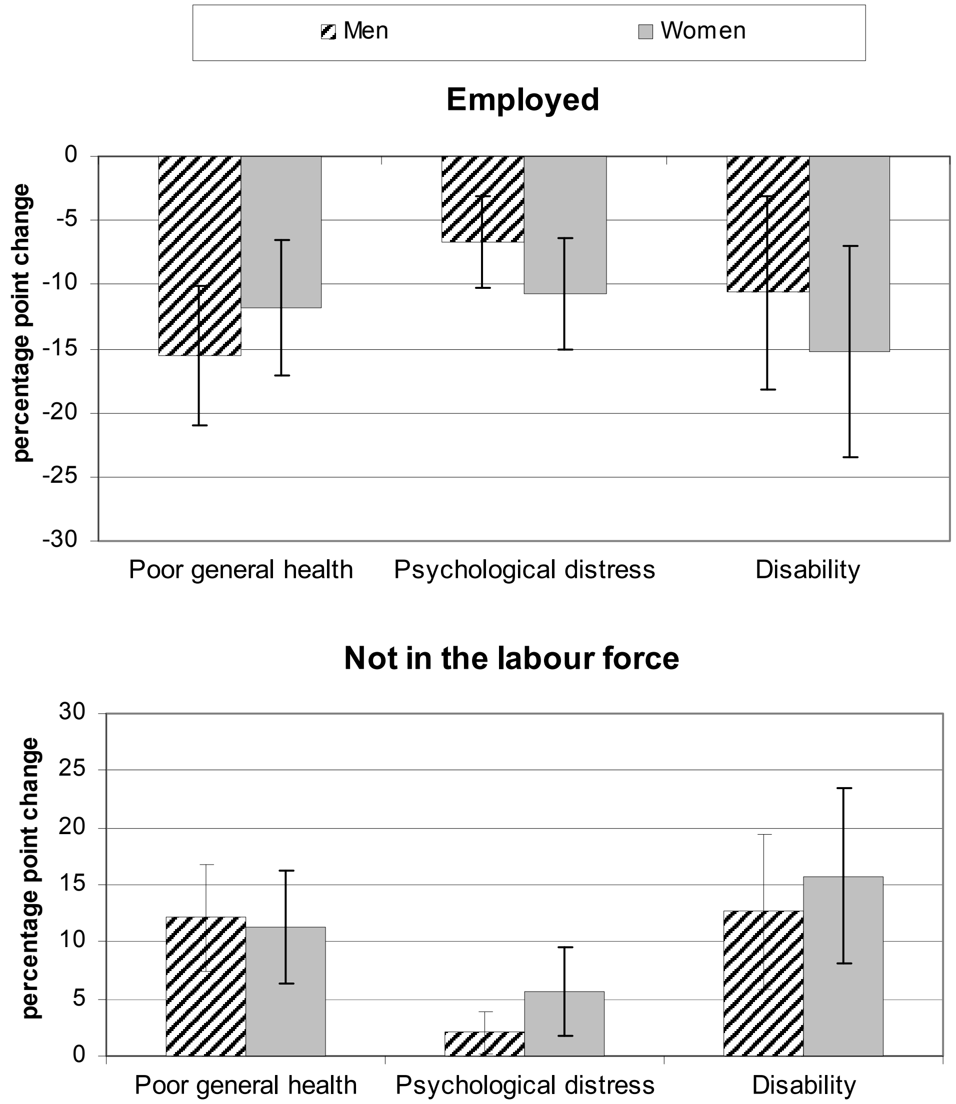
The results suggest that disability and both elements of health have statistically significant associations with LMOs. The most notable negative associations with the probability of Indigenous labour force participation and employment are between poor (self-reported) general health for Indigenous men and having a disability for Indigenous women. Holding all other modelled factors constant:

- Indigenous men and women in poor general health were 12 and 11 percentage points respectively less likely to participate in the labour force, and 16 and 12 percentage points respectively less likely to be employed, compared with those in good general health.
- Indigenous men and women with high levels of psychological distress were 7 and 11 percentage points respectively less likely to be employed, and 2 and 5 percentage points respectively less likely to be in the labour force.
- Indigenous men and women with a severe or profound disability were 11 and 15 percentage points respectively less likely to be employed than those without.

Indigenous men with a disability were 3 percentage points less likely to be unemployed (tables A.3 and A.4). This counter-intuitive result might be explained by the increased probability of Indigenous men with a disability being out of the labour force (13 percentage points) and that people with a disability who select in the labour force are more likely to be employed than unemployed.

The results are broadly consistent with results presented by Stephens (2010) and Biddle and Webster (2007). These studies found that similar health indicators have negative associations with employment and labour force participation. In contrast to this study, Stephens (2010), using the 2002 National Aboriginal and Torres Strait Islander Social Survey (NATSISS) data and a multinomial logit model, found that having a disability had a slightly larger negative effect on the probability of employment for Indigenous men compared to Indigenous women — having a disability decreased the probability of employment by 13 percentage points for men and 10 percentage points for women.

Figure 5.1 **Marginal effects of health and disability indicators, Indigenous people, 2008** [a,b,c]



a General health and disability are self assessed. Mental health is measured in terms of psychological distress using responses to questions from the Kessler Psychological Distress Scale. The estimated marginal effect indicates the change in the predicted probability of a labour market outcome for changing from: being in good, very good or excellent health, to being in fair or poor health (self assessed); having low/moderate to having high/very high levels of psychological distress (as indicated by the Kessler Psychological Distress Scale); and not having a severe or profound disability to having a severe or profound disability. **b** The probability of the base person being employed is 89 per cent for men and 67 per cent for women. The probability of the base person not participating in the labour force is 4 per cent for men and 26 per cent for women. Definitions of all variables are in table A.1. **c** The bars attached to each estimate indicate the 95 per cent confidence interval of the estimate.

Data source: Productivity Commission estimates based on NATSISS (2008).

Biddle and Webster (2007) estimated marginal effects for Indigenous people using a single multinomial probit model (without separating the sample into men and women), and also found a strong negative relationship between having a disability, and the probability of employment (both mainstream employment and CDEP participation) and labour force participation.

5.2 Associations with educational attainment

Most studies find that higher levels of educational attainment are associated with improved LMOs, holding other factors constant. This study examines the marginal effects of educational attainment relative to having completed year 10 or 11 and no non-school qualification. For this analysis, a 'pathways' approach was used to combine years of schooling with non-school qualifications, recognising that people may take different pathways through the education system. This approach is particularly relevant to the way that Indigenous people tend to engage with the education system compared to non-Indigenous people — on average, Indigenous people have lower levels of non-school qualifications and obtain them at a later age, often without having completed year 12.

The associations between having a non-school qualification was estimated separately for people who completed year 12, who completed year 10 or 11 only, and who had not completed year 10. The association between having a degree was estimated separately from the other education categories. The estimated associations between education, and the probability of being employed and not in the labour force are reported in figure 5.2.

For Indigenous women, the associations between education, and employment and labour force participation were statistically significant and relatively large. Any education of year 10 or above was associated with an increased probability of being in the labour force and being employed, and having a non-school qualification without year 10 was also associated with increased probability of labour force participation. Obtaining a degree had the largest positive association with the probability of being in the labour force (16 percentage points) and with being employed (19 percentage points) compared to Indigenous women with year 10 or 11 schooling but no other qualification. The increase in the probability of being employed associated with having a year 12 plus a non-school qualification, compared to those with year 10 or 11 education only, was also large (17 percentage points).

For Indigenous men, the results were more varied and relationships between some education variables and improved LMOs were weaker. Compared to Indigenous men with year 10 or 11 and no non-school qualification:

- Indigenous men with year 12 plus a non-school qualification were 3 percentage points more likely to be in the labour force, and 6 percentage points more likely to be employed
- Indigenous men who did not complete year 10 and had no non-school qualification were 4 percentage points less likely to be employed and in the labour force.

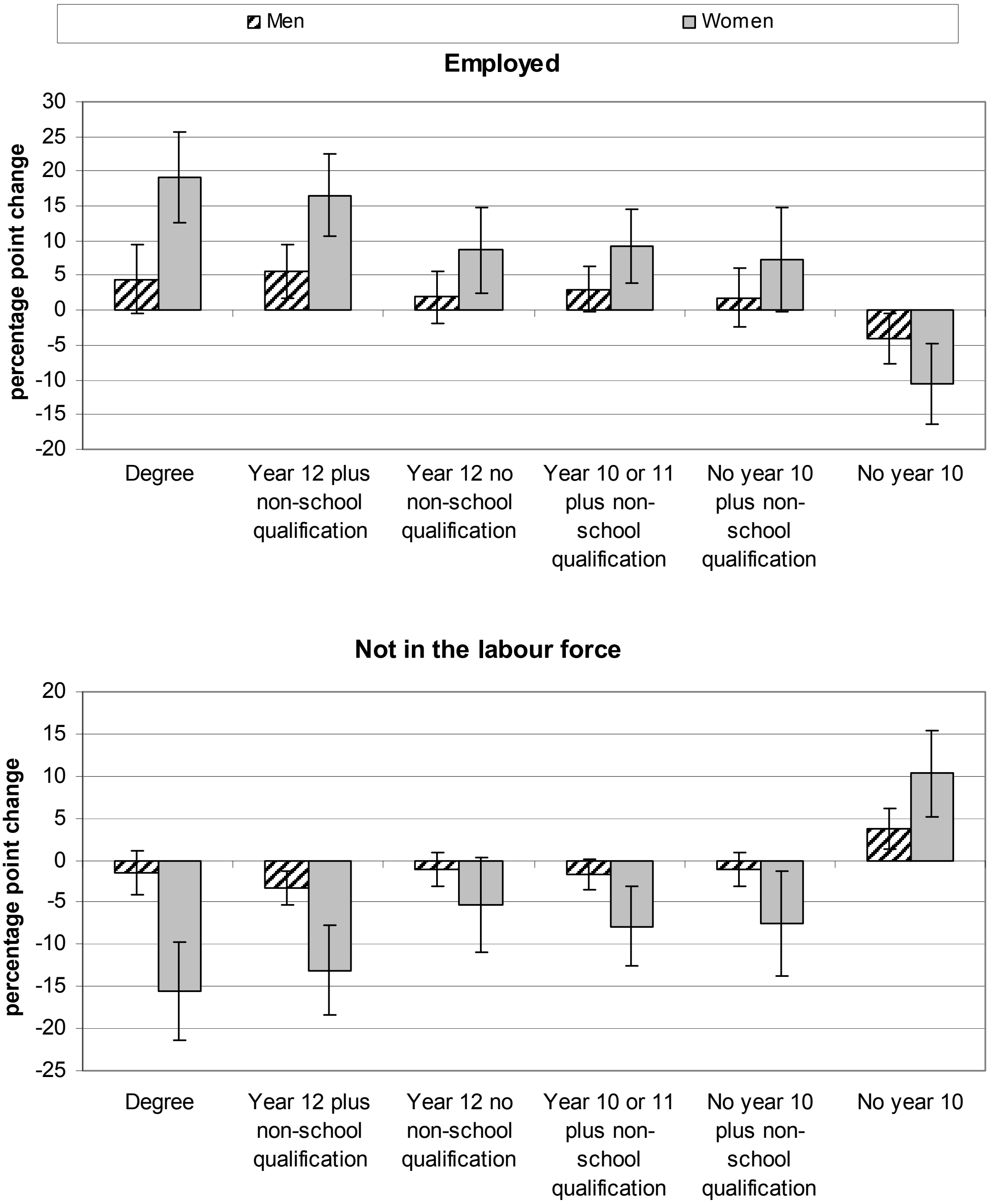
The results for other education outcomes were not statistically significantly different (at the 5 per cent level) from completing year 10 or 11. On the face of it, these results suggest that there is weak evidence that higher levels of educational attainment affect LMOs for men, although the results may be reflective of small subsample sizes (particularly for men with degrees), rather than the true relationship between educational attainment and LMOs. It may also indicate that other factors are important in determining the LMOs of Indigenous men.

There was generally no statistically significant relationship between education and both CDEP and unemployment outcomes for men and women (most likely due to high standard errors resulting from the small sample size for these outcomes) (table A.4 and A.5).

Results in this analysis are broadly consistent with other similar studies. However, results vary depending on the education levels examined and the choice of reference category.

Stephens (2010) used a reference group which had completed year 10 but with no further qualifications and found that higher education levels were associated with a higher probability of employment and labour force participation. In particular, having a degree or certificate was associated with a 15 percentage point increase in the probability of employment for Indigenous men and a 42 percentage point increase for women. Stephens (2010) also found negative associations between education levels of year 9 and below and mainstream employment. This might reflect other factors that affect LMOs apart from educational attainment, such as family dysfunction or social exclusion (Stephens 2010).

Figure 5.2 **Marginal effects of education, Indigenous people, 2008** [a,b,c]



a Education variables reflect the combinations of years of secondary schooling and non-school qualifications. The estimated marginal effect indicates the change in the predicted probability of a labour market outcome for a change from year 10 or 11 and no non-school qualification. **b** The probability of the base person being employed is 89 per cent for men and 67 per cent for women. The probability of the base person not participating in the labour force is 4 per cent for men and 26 per cent for women. Definitions of all variables are in table A.1. **c** The bars attached to each estimate indicate the 95 per cent confidence interval of the estimate.

Data source: Productivity Commission estimates based on NATSISS (2008).

5.3 Associations with personal and demographic characteristics

The associations between a range of personal and demographic characteristics, and labour force participation and employment are presented in table 5.1. These characteristics include years of work experience, location (remoteness, living in homelands and the Socio-Economic Index for Areas (SEIFA)), having difficulty with the English language, and history of arrest and imprisonment.

Table 5.1 **Marginal effects of selected personal and demographic characteristics, Indigenous people 2008** [a]

Percentage point change in predicted probability

		Men		*Women*	
Explanatory variable	*Explanatory variable unit*	*Employed*	*Not in the labour force*	*Employed*	*Not in the labour force*
Work experience[b]	Years	2 ***	-1 ***	4 ***	-3 ***
Remoteness	Binary	-10 ***	-2 **	-0.1	-4 **
Lives in homelands	Binary	-3 *	1	-7 ***	5 **
SEIFA[c]	No.	2 ***	-0.4 **	3 ***	-2 ***
Difficulty with English language	Binary	-4	6 *	-7	7
History of arrest	Binary	-8 ***	2 **	-17 ***	9 ***
History of imprisonment	Binary	-8 ***	3 **	-1	3

*** = significant at 1 per cent level (a 1 in 100 possibility that the result is due to chance); ** = significant at 5 per cent level (a 5 in 100 possibility that the result is due to chance); * = significant at 10 per cent level (a 10 in 100 possibility that the result is due to chance). No stars indicate that the variable is not statistically significant. **a** The probability of the base person being employed is 89 per cent for men and 67 per cent for women. The probability of the base person not participating in the labour force is 4 per cent for men and 26 per cent for women. Definitions of all variables are in attachment table 1. The marginal effects of binary personal and demographic characteristics are measured relative to a person not having the characteristic. For example, the estimated marginal effect for remoteness indicates the change in the predicted probability of a labour market outcome for changing from living in a non-remote area to living in a remote area. **b** The estimated marginal effect indicates the change in the predicted probability of a labour market outcome for a one year increase in work experience over the average number of years of experience (15 years for men and 11 years for women). **c** The estimated marginal effect for SEIFA indicates the change in the predicted probability of a labour market outcome associated with a change from living in an area with a SEIFA score in decile 3 to 4.

Data source: Productivity Commission estimates based on NATSISS (2008).

Work experience

An additional year of work experience was associated with an increase in the likelihood of employment of 2 and 4 percentage points for Indigenous men and women respectively, compared to someone with average years of experience (around 15 years for men and 11 years for women). For Indigenous women, this was reflected in a 3 percentage point increase in the probability of participating in the labour force. For men, the associated increase in the likelihood of labour force participation was around 1 percentage point.

The association between work experience and LMOs is not linear, that is, it is different depending on the number of years of work experience the respondent has. The increase in the predicted probability of mainstream employment from an additional year of work experience is higher for those with fewer than average years of experience and lower for those with more than average years of experience.[2]

Location

The remoteness of an Indigenous person's place of residence can affect their LMOs for a number of reasons (chapter 4). Remoteness is associated with an increased probability of participating in the labour force of 2 percentage points for Indigenous men and 4 percentage points for Indigenous women. For Indigenous men living in a remote area, the probability of CDEP participation increases by 13 percentage points, relative to Indigenous men living in a non-remote area (tables A.3 and A.4). This reflects the strong association between remoteness and the availability of CDEP.

Separate from any effects of remoteness on Indigenous LMOs, living in traditional homelands is associated with a reduced probability of employment for Indigenous men and women (3 and 7 percentage points, respectively). This may signal cultural attachment or a preference for living a traditional lifestyle, or more limited employment opportunities in traditional homelands.

The relative level of socioeconomic disadvantage of the region in which a person's residence is located might be an indicator of labour demand. It may also indicate access to material and social resources that might affect labour force participation, such as income, education and transport. Living in a relatively less disadvantaged

2 To illustrate this, the marginal effects were calculated at 25th and 75th percentile values of experience. Twenty five per cent of women have three or less years of work experience and 75 per cent have 18 years or less. Twenty five per cent of men have six years of work experience or less and 75 per cent have 25 years or less. For women with fewer years of work experience, the probability of mainstream employment increased by 6 percentage points, and for men increased by 7 percentage points. For men with more than average years of work experience, an additional year increases probability of mainstream employment by 0.3 percentage points, and for women increases by 1 percentage point.

area (as indicated by a high SEIFA score) was found to increase the probability of employment by 2 percentage points for Indigenous men and 3 percentage points for Indigenous women, compared to those living in a relatively more disadvantaged area.

Contact with the criminal justice system

A history of arrest in the last five years had a highly significant negative association with employment. The association between arrest and the probability of employment were substantially larger for Indigenous women than men (17 percentage points compared with 8 percentage points).

The negative association between a history of arrest and labour force participation were also larger for women compared with men (9 percentage points compared with 2 percentage points). This is consistent with findings from Stephens (2010) who also found larger negative associations for women.

The relatively smaller association between arrest and labour force participation relative to employment (for both men and women) might mean that arrest did not fully discourage labour force participation, but affected the ability to obtain a job. The association between arrest and unemployment were significant for both Indigenous men and women (5 and 6 percentage points respectively (tables A.3 and A.4)).

After controlling for other factors, including recent history of arrest, an Indigenous man who had been imprisoned in his lifetime had a reduced probability of employment of 8 percentage points and a reduced probability of participation in the labour force of 3 percentage points compared to an Indigenous man who has never been imprisoned. The results indicate no significant association between imprisonment and LMOs for women, although this may reflect the small number of women in the sample who had been in jail.

In interpreting these results, it is expected that the variables for arrest and imprisonment are collinear as generally people who have been in jail have also been arrested. However, it should be noted that the NATSISS data includes observations for some people who have not been arrested in the last five years, but have been in jail in their lifetime, and others who have been in jail without being arrested for other reasons.

The effects of a history of criminal activity on a person's LMO are also likely to be subject to endogeneity bias. Previous studies show that people who are unemployed are more likely to commit crimes (Freeman 1999).

5.4 Associations with social and cultural factors

As discussed in chapter 4, some attributes, such as the motivation and ability to obtain employment, and preferences for paid versus unpaid work, can be influenced by a person's social and cultural environment. The marginal effects of selected social and cultural factors associated with labour force participation and employment are presented in table 5.2.

Table 5.2 **Marginal effects of selected social and cultural factors, Indigenous people, 2008** [a, b]

Percentage point change in predicted probability

		Men		*Women*	
Explanatory variable	*Explanatory variable unit*	*Employed*	*Not in the labour force*	*Employed*	*Not in the labour force*
Does not provide support outside household	Binary	-3 *	2 *	-10 ***	10 ***
Lives in a multifamily household	Binary	-6 ***	4 **	-6 **	6 **
Provides unpaid child care	Binary	-2	-0.3	-7 ***	4 *
Participation in traditional cultural activities	Binary	0.2	0.5	-7 **	3
Participation in social cultural events	Binary	1	-2 **	10 ***	-8 ***

*** = significant at 1 per cent level (a 1 in 100 possibility that the result is due to chance); ** = significant at 5 per cent level (a 5 in 100 possibility that the result is due to chance); * = significant at 10 per cent level (a 10 in 100 possibility that the result is due to chance). No stars indicate that the variable is not statistically significant. **a** The probability of the base person being employed is 89 per cent for men and 67 per cent for women. The probability of the base person not participating in the labour force is 4 per cent for men and 26 per cent for women. Definitions of all variables are in table A.1. **b** The marginal effects of binary social and cultural characteristics are measured relative to a person not having the characteristic. For example, the marginal effect for 'lives in a multifamily household' indicates the change in the predicted probability of a labour market outcome for changing from living in a single family household to living in a multifamily household.

Data source: Productivity Commission estimates based on NATSISS (2008).

For Indigenous women, the largest associations relate to providing support (to relatives or others) outside the household and to participating in social cultural events. Of the factors considered, only some have statistically significant associations for Indigenous men.

Indigenous men and women who provide support outside the household were more likely to be employed (3 and 10 percentage points respectively) or in the labour force (2 and 10 percentage points respectively). It might be that providing support outside the household means that a person has greater exposure to social networks or groups that positively influence motivation to participate in paid work or lower job search costs (chapter 4). Providing support could also constrain household members' participation in the labour force. The results suggest that the dominant association is the positive effect of enhanced networks and motivation.

Living in a multifamily household (as defined by two or more families) has a negative association with employment (6 percentage points for both men and women) and labour force participation (4 and 6 percentage points for men and women, respectively).

Providing unpaid childcare outside the household was found to have a statistically significant association with the probability of employment and labour force participation for Indigenous women but not for Indigenous men. Indigenous women who provided unpaid childcare were 7 percentage points less likely to be employed and 4 percentage points less likely to be in the labour force.

Participation in three or more types of traditional cultural activities (such as hunting, gathering or making traditional crafts) was associated with a reduced probability of employment of 7 percentage points for Indigenous women. This suggests there might be some substitution between paid and unpaid work for Indigenous women.

For Indigenous women, participation in three or more types of social cultural events, ceremonies or organisations in the last 12 months has a large positive association with the probability of employment (10 percentage points) and labour force participation (8 percentage points). For Indigenous men, participation in these social cultural events has only a small positive association with labour force participation. These results suggest that participation in such events could signal engagement with positive social networks and norms.

As discussed in chapter 4, there are a few studies that examine the effects of social and cultural factors on Indigenous LMOs. Consistent with this study, Stephens (2010) found that living in homelands had a negative association with the probability of employment for Indigenous men in remote and non-remote areas, and a small negative association for Indigenous women in remote areas. In contrast to this study, Stephens found that participation in a 'cultural event' was negatively associated with the probability of employment for Indigenous men and women. Hunter and Borland (1997) found that hunting, fishing and gathering activities had no effect on the probability of employment, suggesting there was little substitution between these traditional activities and market work. These inconsistencies might be explained by differences in the type of cultural participation variables used, as well as differences in modelling approaches and data sources.

Box 5.1 **Summary of key findings**

Educational attainment

- This study demonstrates that there are statistically significant associations between education and labour market outcomes (LMOs) for Indigenous Australians. This is consistent with human capital theory that improving education has the potential to improve LMOs.
- Results indicate that higher levels of education have different outcomes for Indigenous women and men.
  - The association between employment and labour force participation (LFP), and all levels of education at or above year 10 was statistically significant for women.
  - For men, having year 12 and a non-school qualification (other than a degree) has a statistically significant association with employment and LFP (compared to having only year 10 or 11). However, the associations between other levels of education and employment and LFP were not statistically significant.

Health and disability

- This study found statistically significant associations between employment and LFP, and physical and mental health. It also found statistically significant (negative) associations between employment and LFP, and severe or profound disability.
- The results in this study are consistent with the hypothesis that improving Indigenous health, and reducing disability, might improve LMOs.

Contact with the criminal justice system

- There are statistically significant (negative) associations between contact with the criminal justice system and Indigenous LMOs. Men who have been arrested or imprisoned, and women who have been arrested, are less likely to be in the labour force and employed than those who have not.
- These results are consistent with the hypotheses from other researchers that employers might use a person's criminal record as a signal of their desirability as an employee and that contact with the criminal justice system might discourage people from participating in paid work.

Cultural and social engagement

- There were statistically significant associations, for women, between engagement with the community and culture, and employment and LFP.
  - Women who participate in social cultural events (such as festivals and carnivals involving arts, music and dance) and provide support (other than unpaid childcare) outside their household are more likely to be employed and in the labour force than those who do not.
  - Women who undertake traditional cultural activities (such as hunting, and traditional arts and crafts) are less likely to be employed, although these activities might accrue other benefits to the participant or the community.
- For men, the equivalent associations were weak.

6 Conclusion

This paper has examined the relationships between Indigenous labour market outcomes (LMOs) and selected variables using 2008 National Aboriginal and Torres Strait Islander Social Survey (NATSISS) data. It is a companion to section 13.2 of *Overcoming Indigenous Disadvantage: Key Indicators 2011* (OID Report) (SCRGSP 2011), providing some policy context for the analysis, an explanation of the theory underpinning variable selection, and an interpretation of the results (a summary of the key findings can be found in box 5.1). It extends the analysis used in the OID Report to examine the effect of additional variables to control for social and cultural influences.

The main benefit of using data obtained from the NATSISS, compared to the Census, is that the NATSISS allows examination of Indigenous health status and criminal history, as well as social and cultural factors, which are not available in the Census. Controlling for these factors improves the estimated marginal effects because it reduces the omitted variable bias in the model. It also confirms that a broad range of factors is likely to be associated with Indigenous LMOs. The NATSISS also allows examination of participation in the Community Development Employment Project (CDEP) Scheme, which has been an important component of Indigenous employment in some areas, although its importance is expected to reduce in coming years.

The findings of this study have some implications for further research in terms of other areas for study, modelling approaches and data collection.

Alternative modelling approaches

This paper provides an interpretation of the results of a multivariate model which has greater explanatory power than bivariate analysis in examining LMOs. The paper examines the advantages and disadvantages of using a multinomial probit (rather than logit) model and unweighted (rather than weighted) data. It also sets out some qualifications to the results, and provides an extensive set of results for potential use in further analysis.

As discussed, studies of this type suffer from numerous potential sources of endogeneity, meaning that the results are best interpreted as associations. More research could be undertaken to develop statistical models to allow for the causality that is implied by theory to be tested.

The analysis provides a basis for future research using alternative model specifications and can be extended to other issues of interest. The technique could be applied to investigate relationships between individuals' characteristics and factors that may be associated with outcomes of interest other than LMOs, for example, year 12 attainment, which is another COAG performance benchmark.

Some studies have explored other LMOs, such as discouraged worker effects. Others have used different modelling approaches, such as interaction terms, binary models, two-stage estimation and simultaneous equation models.

Another area of potential further study is in analysing the variation in determinants of Indigenous labour market outcomes between remote and non-remote areas, given the heterogeneity of Indigenous populations across regions. This could involve the estimation of separate models based on disaggregated samples of Indigenous people in remote and non-remote areas (as was done by Stephens (2010)).

Despite the many caveats, the comparison of results from alternative model specifications has shown that the conclusions are robust to modelling choices. That said, there might be benefit developing an agreed methodology for future studies, especially if results are to be comparable over time.

Overcoming data limitations

While the NATSISS provides a wealth of information about Indigenous people that can be used to research a broad range of subject areas related to the wellbeing of Indigenous Australians, changes to data collection could improve its application to labour market research.

A major limitation of the NATSISS data is the lack of indicators of labour demand across different geographic areas. Remoteness is a variable that partly controls for labour demand. A more direct indicator of labour demand would be related to the level of economic activity and labour market conditions. Potential indicators of local labour market conditions could include local employment to population ratios or unemployment rates disaggregated by regions such as Census Collection Districts. The inclusion of more direct indicators is likely to improve the explanatory power of the model in estimating the effects of the economic variables, especially if the latter are defined over large heterogeneous regions.

This study examined a subset of LMOs related to labour force status. An important extension would be to examine the relationships between various factors and Indigenous wages (which can be interpreted as an indicator of productivity). However, the NATSISS data do not allow for the identification of income earned from wages and salaries separate from other income sources, including government pensions and allowances, disaggregated by remoteness area. These data limitations constrain the ability to undertake research in this area.

The variables used to represent social and cultural factors would also benefit from information about the extent and quality of social networks, with a view to identifying whether they might reduce job search costs and improve LMOs. While the dataset includes information on participation in cultural events and activities and whether the respondent provides support outside the household, there is no information indicating whether these activities are likely to compete with or complement participation in the paid labour market.

The dataset also includes information on the age, education level and Indigenous status of the respondent's friends. This provides an indication of the diversity of their social network, but not whether it positively or negatively influences motivations to participate in the labour force.

Further research into labour market outcomes would also benefit from information on the education level and employment status of other household members and close family, as these might influence motivations to participate in the labour force and preferences for paid work compared to unpaid work.

One of the conclusions from modelling exercises of this type is the importance of controlling for unobservable factors such as innate ability and motivation, which reduces omitted variable bias. One way of doing this is to conduct longitudinal surveys of Indigenous people or link data across multiple datasets. Governments are examining opportunities for developing linked datasets.[1]

The modelling in this study does not allow for comparisons between Indigenous and non-Indigenous LMOs. It is expected that many of the results of this study would be similar for non-Indigenous people, but empirical analysis of the differences in outcomes between Indigenous and non-Indigenous people is an important priority

1 During consultations for the 2011 OID Report, government agencies advised the Steering Committee that they are examining opportunities for linking data across multiple administrative data sets. Knowing the relationships between factors such as health, education, income and housing could help governments to develop more effective policies and programs. However, the practical application of data linkage may take several years, because of the technical challenges in linking data and the need to address concerns about privacy that might arise (SCRGSP 2011).

for government. Comparisons of Indigenous and non-Indigenous outcomes are made in the 2011 OID using data from the National Health Survey (NHS).[2] However previous NHS collections do not provide information on arrest and imprisonment, or the social and cultural elements that have been examined by this study. Governments may wish to examine the feasibility of collecting this information to allow the gap between Indigenous and non-Indigenous labour market outcomes to be more rigorously examined.

Comparisons with non-Indigenous people could potentially also be undertaken using the Census, which was conducted on 9 August 2011, in a similar manner to the work done using 2006 Census data for the 2009 OID Report (SCRGSP 2009).

2 The NHS has education and labour market information, as well as health conditions and self-assessed health status.

A Appendix

Table A.1 Definition of variables, Indigenous people aged 15 to 64 years, 2008[a]

Variable name	*NATSISS variable*	*Variable type*[b]	*Definition*
Dependent variable			
Labour market outcome			
Mainstream (non-CDEP) employment	EMPSTAC, WONCDEP	..	Had a job or business for a minimum of one hour in the week prior to the survey interview, excluding CDEP participants
CDEP participation	EMPSTAC, WONCDEP	..	CDEP participant
Unemployed	EMPSTAC	..	Unemployed (not employed, but actively looking for work and able to start work in the four weeks prior to the survey interview)
Not in the labour force	EMPSTAC	..	Not in the labour force (neither employed or unemployed)
Explanatory variables			
Population and demographic characteristics			
Age	AGEC	no.	Age in years
Age squared	AGEC	no.	Age in years squared
Married	SMS	binary	Is married. This is defined as 'social marital status'. A marriage exists when two people live together as husband and wife, or partners, regardless of whether the marriage is formalised through registration
Dependent children[c]	FAMCOMPH, STHHTYC, RELHHLC, DEPS14C	no.	Number of dependent children aged 0 -14 (derived from NATSISS variables including household type, family composition of households, relationship in household and number of persons aged 0 to 14 in the household)
Lives in multifamily household	STHHTYC	binary	Lives in a household with two or more families
Remote	ARIAC	binary	Lives in a remote area as defined by the Australian Standard Geographical Classification
Lives in homelands	CULIQ3	binary	Recognises and lives in homelands

(Continued next page)

Table A.1 (Continued)

Variable name	NATSISS variable	Variable type[b]	Definition
SEIFA	CDDIS	no.	Decile value based on the Socio-economic Index of Relative Disadvantage for the Census collection district of the household (a high number indicates less disadvantage)
Difficulty with English	WDIFFENG	binary	Has difficulty communicating with English speakers
Self-assessed health			
Poor general health	SAHQ1	binary	Is in fair or poor health (self assessed)
Psychological distress	K5GRP	binary	Has high or very high levels of psychological distress (has a Kessler score between 12 and 25; a high score indicates high levels of psychological distress)
Disability	DSTAT	binary	Has a severe or profound core-activity limitation - a specified condition for which the person requires help or supervision in one or more core activities (for example, self-care, mobility or communication)
Education			
Degree or higher	HIGHSCHC, HINSCLC	binary	Has a degree or higher (with or without year 12)
Year 12 plus non-school qualification	HIGHSCHC, HINSCLC	binary	Has completed year 12 and has a diploma or certificate qualification
Year 12, no non-school qualification	HIGHSCHC, HINSCLC	binary	Has completed year 12 and has no non-school qualification
Year 10 or 11 plus non-school qualification	HIGHSCHC, HINSCLC	binary	Has completed year 10 or 11 and has a diploma or certificate qualification
Year 10 or 11, no non-school qualification	HIGHSCHC, HINSCLC	binary	Has completed year 10 or 11 and has no non-school qualification
No year 10 plus non-school qualification	HIGHSCHC, HINSCLC	binary	Did not complete year 10 and has a diploma or certificate qualification
No year 10, no non-school qualification	HIGHSCHC, HINSCLC	binary	Did not complete year 10 and has no non-school qualification
Work experience			
Years in paid employment[d]	DUREMP	no.	Years in paid employment
Years in paid employment squared	DUREMP	no.	Years in paid employment squared

(Continued next page)

Table A.1 (Continued)

Variable name	*NATSISS variable*	*Variable type*[b]	*Definition*
Crime			
Arrested	WARST5Y	binary	Was arrested by police in last five years
Imprisoned	WJAILLT	binary	Has been imprisoned in lifetime. Includes time spent in gaol without being arrested (for example, respondent was placed in gaol for their own safety) and in gaol awaiting a court hearing. Excludes time spent in custody of a 'night patrol', in a police lock up or visiting other people in gaol
Social and cultural factors			
Participation in traditional cultural activities	NCULACT	binary	Participated in three or more types of cultural activities in the last 12 months (including fishing, hunting, gathering, telling stories or making traditional crafts)
Participation in social cultural events	NCULEVNT	binary	Attended three or more types of cultural events in the last 12 months (including ceremonies, Indigenous organisations, sports carnivals, festivals, funerals or sorry business and NAIDOC week activities)
Does not provide support outside household	WSPTREL, WSPTANY	binary	Did not provide support to anyone (including relatives) outside the household
Provides unpaid childcare outside household	FCSPQ7	binary	Provides unpaid childcare to anyone outside the household in the last 4 weeks

[a] For more information on the definitions in this section, see the NATSISS Users' Guide. [b] For the binary variables, the definition describes a person who has been coded to the value '1'. The excluded category for the education variables is year 10 or 11 and no non-school qualifications. The base person is described in box 3.4. [c] The number of dependent children has been derived using variables including household type, family composition of households, relationship in household and number of persons aged 0 to 14 in the household. Because it is difficult to determine who are parents in multifamily households from the data, and because the NATSISS variable had an upper bound of five children, the estimate of the number of dependent children is likely to be an underestimate. [d] Years of experience in paid employment is total actual time in paid employment. This variable has an upper bound of 25 years.

.. Not applicable.

Source: ABS National Aboriginal and Torres Strait Islander Social Survey (NATSISS) 2008; ABS (2010c).

Table A.2 **Selected descriptive statistics of variables used in the models of labour market outcomes, Indigenous people aged 15 to 64 years, 2008**[a,b,c]

Variable name	*Variable type*	*Men*		*Women*	
		Mean	*S.D.*	*Mean*	*S.D.*
Explanatory variables					
		Unweighted			
Population and demographic characteristics					
Age	no.	37.08	12.90	36.79	12.68
Age squared	no.	1 541.50	1 011.23	1 513.93	998.87
Married	binary	0.46	0.50	0.52	0.50
Dependent children	no.	0.80	1.27	1.17	1.35
Lives in multifamily household	binary	0.12	0.32	0.14	0.35
Remote	binary	0.35	0.48	0.34	0.47
Lives in homelands	binary	0.28	0.45	0.25	0.43
SEIFA	no.	3.10	2.45	2.88	2.34
Difficulty with English	binary	0.05	0.21	0.03	0.18
Self assessed health					
Poor general health	binary	0.23	0.42	0.24	0.43
Psychological distress	binary	0.28	0.45	0.37	0.48
Disability	binary	0.07	0.25	0.08	0.28
Education					
Degree or higher	binary	0.04	0.21	0.07	0.25
Year 12 plus non-school qualification	binary	0.07	0.26	0.08	0.28
Year 12, no non-school qualification	binary	0.10	0.29	0.09	0.29
Year 10 or 11 plus non-school qualification	binary	0.17	0.37	0.15	0.35
Year 10 or 11, no non-school qualification	binary	0.28	0.45	0.31	0.46
No year 10 plus non-school qualification	binary	0.07	0.26	0.05	0.23
No year 10, no non-school qualification	binary	0.27	0.44	0.25	0.43
Work experience					
Years in paid employment	no.	14.63	8.91	10.70	8.78
Years in paid employment squared	no.	293.43	252.06	191.53	228.94
Crime					
Arrested	binary	0.25	0.43	0.11	0.31
Imprisoned	binary	0.20	0.40	0.04	0.19

(Continued next page)

Table A.2 (Continued)

Variable name	Variable type	Men Mean	Men S.D.	Women Mean	Women S.D.
Social and cultural factors					
Participation in cultural activities	binary	0.21	0.41	0.19	0.39
Participation in cultural events	binary	0.25	0.44	0.29	0.46
Does not provide support outside household	binary	0.28	0.45	0.23	0.42
Provides unpaid childcare outside household	binary	0.12	0.32	0.24	0.43
		Weighted			
Population and demographic characteristics					
Age	no.	35.17		35.86	
Age squared	no.	1 400.91		1 449.68	
Married	binary	0.47		0.51	
Dependent children	no.	0.78		1.11	
Lives in multifamily household	binary	0.17		0.20	
Remote	binary	0.26		0.25	
Lives in homelands	binary	0.27		0.25	
SEIFA	no.	3.40		3.07	
Difficulty with English	binary	0.04		0.03	
Self-assessed health					
Poor general health	binary	0.22		0.23	
Psychological distress	binary	0.28		0.36	
Disability	binary	0.07		0.07	
Education					
Degree or higher	binary	0.04		0.06	
Year 12 plus non-school qualification	binary	0.07		0.09	
Year 12, no non-school qualification	binary	0.11		0.10	
Year 10 or 11 plus non-school qualification	binary	0.16		0.14	
Year 10 or 11, no non-school qualification		0.29		0.30	
No year 10 plus non-school qualification	binary	0.06		0.06	
No year 10, no non-school qualification	binary	0.26		0.25	
Work experience					
Years in paid employment	no.	13.58		10.36	
Years in paid employment squared	no.	266.98		184.77	

(Continued next page)

Table A.2 (Continued)

Variable name	*Variable type*	*Men*		*Women*	
		Mean	*S.D.*	*Mean*	*S.D.*
Crime					
Arrested	binary	0.25		0.10	
Imprisoned	binary	0.18		0.03	
Social and cultural factors					
Participation in traditional cultural activities	binary	0.20		0.17	
Participation in social cultural events	binary	0.24		0.27	
Does not provide support outside household	binary	0.30		0.25	
Provides unpaid childcare outside household	binary	0.11		0.23	

a See table A.1 for definitions of variables. **b** The mean (or average) is the sum of all values for each variable in the sample, divided by the number of observations. For binary variables, the mean indicates the proportion of the sample population in the category that is defined as '1'. Weighted means were calculated by running weighed regressions at the mean using person weights. **c** The standard deviation is a measure of the variability of variable observations from the mean value. A low standard deviation indicates that observations tend to be close to the mean while a large standard deviation indicates that observations are 'spread out' over a large range of values.

Source: Productivity Commission estimates based on ABS NATSISS 2008.

Table A.3 **Marginal effects of explanatory variables in the model of labour market outcomes, Indigenous men aged 15 to 64 years, 2008**[a,b,c,d,e]

Variable name	Variable type	Employed		CDEP		Unemployed		NILF		Value held at
		M.E.	*S.E.*	*M.E.*	*S.E.*	*M.E.*	*S.E.*	*M.E.*	*S.E.*	
		ppt	*ppt*	*ppt*	*ppt*	*ppt*	*ppt*	*ppt*	*ppt*	
Explanatory variables										
Population and demographic characteristics										
Age	no.	-2.0 ***	0.5	0.0	0.1	1.1 ***	0.4	1.0 ***	0.3	37.1
Age squared	no.	0.0 **	0.0	0.0	0.0	0.0 **	0.0	0.0	0.0	1 375.1
Married	binary	-10.6 ***	2.1	0.1	0.2	2.7 **	1.3	7.8 ***	1.6	–
Dependent children	no.	-2.0 ***	0.5	-0.2	0.1	1.3 ***	0.4	0.9 ***	0.3	0.8
Lives in multifamily household	binary	-6.4 **	2.5	0.3	0.3	1.7	1.6	4.3 **	1.7	–
Remote	binary	-9.7 ***	3.0	12.8 ***	3.2	-1.4	1.0	-1.7 **	0.7	–
Lives in homelands	binary	-2.7 *	1.5	0.8 *	0.4	0.8	1.1	1.1	0.9	–
SEIFA	no.	1.7 ***	0.4	-0.5 **	0.2	-0.8 ***	0.3	-0.4 **	0.2	3.1
Difficulty with English	binary	-3.9	3.6	0.5	0.5	-2.0	1.7	5.5 *	2.9	–
Self assessed health										
Poor general health	binary	-15.5 ***	2.8	-0.1	0.2	3.5 **	1.6	12.1 ***	2.4	–
Psychological distress	binary	-6.7 ***	1.9	0.0	0.2	4.7 ***	1.5	2.1 **	0.9	–
Disability	binary	-10.6 ***	3.9	0.6	0.7	-2.6 *	1.5	12.7 ***	3.5	–
Education										
Degree or higher	binary	4.4 *	2.5	-0.6 **	0.3	-2.3	2.1	-1.5	1.4	–
Year 12 plus non-school qualification	binary	5.5 ***	2.0	-0.5 *	0.3	-1.7	1.6	-3.4 ***	1.0	–
Year 12, no non-school qualification	binary	1.9	1.9	-0.2	0.2	-0.7	1.4	-1.1	1.0	–
Year 10 or 11 plus non-school qualification	binary	3.0 *	1.6	-0.3	0.2	-1.0	1.3	-1.7 *	0.9	–

(Continued next page)

Table A.3 (Continued)

Variable name	*Variable type*	*Employed*		*CDEP*		*Unemployed*		*NILF*		*Value held at*[b]
		M.E.	*S.E.*	*M.E.*	*S.E.*	*M.E.*	*S.E.*	*M.E.*	*S.E.*	
		ppt	*ppt*	*ppt*	*ppt*	*ppt*	*ppt*	*ppt*	*ppt*	
No year 10 plus non-school qualification	binary	1.8	2.2	-0.4	0.2	-0.3	1.8	-1.1	1.1	–
No year 10, no non-school qualification	binary	-4.1 **	1.8	0.0	0.2	0.4	1.2	3.7 ***	1.2	–
Work experience										
Years in paid employment	no.	3.3 ***	0.5	0.1	0.1	-1.7 ***	0.4	-1.7 ***	0.3	14.6
Years in paid employment squared	no.	0.0 ***	0.0	0.0	0.0	0.0 **	0.0	0.0 ***	0.0	214.0
Crime										
Arrested	binary	-7.5 ***	2.0	0.7 *	0.4	4.7 ***	1.6	2.1 **	1.1	–
Imprisoned	binary	-8.1 ***	2.3	0.0	0.2	5.1 ***	1.9	3.1 **	1.2	–
Social and cultural factors										
Participation in traditional cultural activities	binary	0.2	1.6	0.0	0.2	-0.7	1.1	0.5	0.9	–
Participation in social cultural events	binary	1.1	1.5	0.5	0.4	0.0	1.1	-1.6 **	0.7	–
Does not provide support outside household	binary	-2.8 *	1.5	0.1	0.2	1.0	1.1	1.6 *	0.9	–
Provides unpaid childcare outside household	binary	-1.6	2.0	0.0	0.2	1.9	1.7	-0.3	1.0	–

M.E. = marginal effect. S.E. = standard error. ppt = percentage point change. *** = significant at 1 per cent level (a 1 in 100 possibility that the result is due to chance); ** = significant at 5 per cent level (a 5 in 100 possibility that the result is due to chance); * = significant at 10 per cent level (a 10 in 100 possibility that the result is due to chance). No stars indicate that the variable is not statistically significant (chapter 3). no. = number. CDEP = Community Development Employment Projects. NILF = Not in the Labour Force. – Nil or rounded to zero.

a See table A.1 for definitions of variables. **b** The results for this study were derived using the Stata statistical package and the ABS' Remote Access Data Laboratory. A multinomial model is used because there are four possible LMOs. There are other discrete choice modelling options, such as binary and nested models. There are two types of discrete choice models — logit and probit models. The difference between the two is largely theoretical. A probit model is used because there is reason to believe that the LMOs are not independently determined, based on there being reasonable doubt that the data would pass a logic test for the Independence of Irrelevant Alternatives assumption, which is necessary for a logit specification (see Fry and Harris (1998)). **c** Marginal effects are derived from the coefficients of the regression and the predicted probabilities. The coefficients are reported in tables A.5 and A.6, and the predicted probabilities in table A.7. Formulae for the marginal effects can be found in the Stata help manuals. **d** Marginal effects are calculated with continuous variables held at mean values and binary variables held at the mode (most common) values (with the exception of marital status in the model for women which is held at the value '1' to make it consistent with the model for men). For continuous variables, the marginal effects represent the percentage point change in the probability of a labour market outcome associated with a one unit increase in the variable from its mean value, holding the value of all other explanatory variables constant. For binary variables other than the education variables, the marginal effects represent the percentage point change in probability of a labour market outcome associated with a change from ‘0’ to ‘1’, holding the value of all other explanatory variables constant. For the education variables, the marginal effects represent the percentage point change in the probability of a labour market outcome associated with an increase in education compared to having year 10 or 11 but no non-school qualification, holding the value of all other explanatory variables constant. [e] Economic theory indicates the factors that might influence or determine LMOs. However, regression analysis only reflect associations between the explanatory and dependent variables, which are the net effect after accounting for endogeneity and other bias (chapter 4).

Source: Productivity Commission estimates based on NATSISS 2008.

Table A.4 **Marginal effects of explanatory variables in the model of labour market outcomes, Indigenous women aged 15 to 64 years, 2008[a,b,c,d,e]**

Variable name	*Variable type*	*Employed*		*CDEP*		*Unemployed*		*NILF*		*Value held at*
		M.E.	*S.E.*	*M.E.*	*S.E.*	*M.E.*	*S.E.*	*M.E.*	*S.E.*	
		ppt	*ppt*	*ppt*	*ppt*	*ppt*	*ppt*	*ppt*	*ppt*	
Explanatory variables										
Population and demographic characteristics										
Age	no.	0.6	0.6	0.0	0.0	0.5 *	0.3	-1.0 **	0.5	36.8
Age squared	no.	0.0 ***	0.0	0.0	0.0	0.0 *	0.0	0.0 ***	0.0	1 353.3
Married	binary	-1.3	2.0	-0.1	0.2	2.3 **	1.1	-0.8	1.7	–
Dependent children	no.	-7.0 ***	0.9	-0.1	0.1	-1.0 **	0.5	8.1 ***	0.7	1.2
Lives in multifamily household	binary	-6.2 **	3.1	0.4	0.4	0.2	1.3	5.7 **	2.7	–
Remote	binary	-0.1	2.5	6.2 ***	1.9	-2.2 **	0.9	-3.9 **	2.0	–
Lives in homelands	binary	-7.7 ***	2.5	0.9 *	0.5	2.0	1.3	4.9 **	2.2	–
SEIFA	no.	2.8 ***	0.5	-0.6 **	0.2	-0.6 **	0.2	-1.7 ***	0.4	2.9
Difficulty with English	binary	-6.7	6.9	0.5	0.6	-0.8	2.6	7.0	6.1	–
Self assessed health										
Poor general health	binary	-11.8 ***	2.7	0.0	0.2	0.5	1.2	11.3 ***	2.5	–
Psychological distress	binary	-10.8 ***	2.2	0.0	0.2	5.1 ***	1.4	5.7 ***	2.0	–
Disability	binary	-15.2 ***	4.2	-0.3	0.3	-0.2	1.7	15.8 ***	3.9	–
Education										
Degree or higher	binary	19.1 ***	3.3	-0.5 *	0.3	-3.0 *	1.7	-15.6 ***	3.0	–
Year 12 plus non-school qualification	binary	16.6 ***	3.0	-0.1	0.3	-3.3 **	1.4	-13.1 ***	2.7	–
Year 12, no non-school qualification	binary	8.6 ***	3.1	-0.1	0.3	-3.2 **	1.3	-5.3 *	2.8	–

(Continued next page)

Table A.4 (Continued)

Variable name	Variable type	Employed		CDEP		Unemployed		NILF		Value held at[b]
		M.E.	S.E.	M.E.	S.E.	M.E.	S.E.	M.E.	S.E.	
		ppt	ppt	ppt	ppt	ppt	ppt	ppt	ppt	
Year 10 or 11 plus non-school qualification	binary	9.2 ***	2.7	-0.1	0.3	-1.3	1.2	-7.9 ***	2.4	–
No year 10 plus non-school qualification	binary	7.3 *	3.8	-0.3	0.3	0.6	2.0	-7.6 **	3.2	–
No year 10, no non-school qualification	binary	-10.7 ***	2.9	0.2	0.3	0.2	1.2	10.3 ***	2.6	–
Work experience										
Years in paid employment	no.	6.5 ***	0.5	0.2 *	0.1	-0.9 ***	0.2	-5.8 ***	0.5	10.7
Years in paid employment squared	no.	-0.1 ***	0.0	0.0 *	0.0	0.0 **	0.0	0.1 ***	0.0	114.4
Crime										
Arrested	binary	-16.6 ***	3.8	1.4 *	0.8	6.2 ***	2.2	9.0 ***	3.3	–
Imprisoned	binary	-0.9	5.9	-0.2	0.4	-1.8	1.9	2.8	5.2	–
Social and cultural factors										
Participation in traditional cultural activities	binary	-6.8 **	2.9	0.4	0.4	3.0 *	1.7	3.4	2.6	–
Participation in social cultural events	binary	9.7 ***	2.1	0.1	0.2	-1.5	1.0	-8.3 ***	1.8	–
Does not provide support outside household	binary	-9.9 ***	2.6	0.2	0.3	0.1	1.1	9.7 ***	2.4	–
Provides unpaid childcare outside household	binary	-7.0 ***	2.4	-0.1	0.2	3.4 **	1.4	3.7 *	2.2	–

M.E. = marginal effect. S.E. = standard error. ppt = percentage point change. *** = significant at 1 per cent level (a 1 in 100 possibility that the result is due to chance); ** = significant at 5 per cent level (a 5 in 100 possibility that the result is due to chance); * = significant at 10 per cent level (a 10 in 100 possibility that the result is due to chance). No stars indicate that the variable is not statistically significant (chapter 3). no. = number. CDEP = Community Development Employment Projects. NILF = Not in the Labour Force. – Nil or rounded to zero.

a See table A.1 for definitions of variables. **b** The results for this study were derived using the Stata statistical package and the ABS' Remote Access Data Laboratory. A multinomial model is used because there are four possible LMOs. There are other discrete choice modelling options, such as binary and nested models. There are two types of discrete choice models — logit and probit models. The difference between the two is largely theoretical. A probit model is used because there is reason to believe that the LMOs are not independently determined, based on there being reasonable doubt that the data would pass a logic test for the Independence of Irrelevant Alternatives assumption, which is necessary for a logit specification (see Fry and Harris (1998)). **c** Marginal effects are derived from the coefficients of the regression and the predicted probabilities. The coefficients are reported in tables A.5 and A.6, and the predicted probabilities in table A.7. Formulae for the marginal effects can be found in the Stata help manuals. **d** Marginal effects are calculated with continuous variables held at mean values and binary variables held at the mode (most common) values (with the exception of marital status in the model for women which is held at the value '1' to make it consistent with the model for men). For continuous variables, the marginal effects represent the percentage point change in the probability of a labour market outcome associated with a one unit increase in the variable from its mean value, holding the value of all other explanatory variables constant. For binary variables other than the education variables, the marginal effects represent the percentage point change in probability of a labour market outcome associated with a change from '0' to '1', holding the value of all other explanatory variables constant. For the education variables, the marginal effects represent the percentage point change in the probability of a labour market outcome associated with an increase in education compared to having year 10 or 11 but no non-school qualification, holding the value of all other explanatory variables constant. **e** Economic theory indicates the factors that might influence or determine LMOs. However, regression analysis only reflect associations between the explanatory and dependent variables, which are the net effect after accounting for endogeneity and other bias (chapter 4).

Source: Productivity Commission estimates based on ABS NATSISS 2008.

Table A.5 **Estimated coefficients in the model of labour market outcomes, Indigenous men aged 15 to 64 years, 2008[a,b,c]**

Variable name	*Employed*		*CDEP*		*Unemployed*	
	Coef.	*S.E.*	*Coef.*	*S.E.*	*Coef.*	*S.E.*
Explanatory variables						
Population and demographic characteristics						
Age	-0.16	0.03	-0.12	0.04	-0.02	0.03
Age squared	0.00	0.00	0.00	0.00	-0.00	0.00
Married	-0.80	0.11	-0.51	0.16	-0.40	0.14
Dependent children	-0.16	0.05	-0.24	0.06	0.00	0.05
Lives in multifamily household	-0.52	0.15	-0.20	0.17	-0.26	0.16
Remote	0.11	0.12	1.94	0.18	0.11	0.13
Lives in homelands	-0.18	0.11	0.25	0.14	-0.06	0.13
SEIFA	0.09	0.02	-0.31	0.06	-0.03	0.03
Difficulty with English	-0.56	0.23	-0.24	0.22	-0.74	0.28
Self assessed health						
Poor general health	-1.08	0.11	-0.87	0.17	-0.56	0.14
Psychological distress	-0.34	0.11	-0.23	0.15	0.13	0.12
Disability	-1.03	0.18	-0.53	0.27	-1.20	0.25
Education						
Degree or higher	0.33	0.28	-1.21	0.77	-0.05	0.40
Year 12 plus non-school qualification	0.94	0.30	0.09	0.41	0.64	0.34
Year 12, no non-school qualification	0.20	0.19	0.00	0.25	0.10	0.21
Year 10 or 11 plus non-school qualification	0.34	0.16	-0.02	0.24	0.18	0.19
No year 10 plus non-school qualification	0.20	0.20	-0.31	0.30	0.13	0.24
No year 10, no non-school qualification	-0.44	0.12	-0.37	0.16	-0.33	0.14
Work experience						
Years in paid employment	0.28	0.03	0.27	0.04	0.05	0.03
Years in paid employment squared	-0.00	0.00	-0.00	0.00	-0.00	0.00
Crime						
Arrested	-0.36	0.12	0.13	0.15	0.13	0.13
Imprisoned	-0.45	0.12	-0.28	0.16	0.06	0.14

(Continued next page)

Table A.5 (Continued)

Variable name	Employed		CDEP		Unemployed	
	Coef.	*S.E.*	*Coef.*	*S.E.*	*Coef.*	*S.E.*
Social and cultural factors						
Participation in traditional cultural activities	-0.06	0.13	-0.03	0.16	-0.13	0.16
Participation in social cultural events	0.29	0.13	0.56	0.16	0.27	0.15
Does not provide support outside household	-0.24	0.11	-0.10	0.15	-0.09	0.13
Provides unpaid childcare outside household	0.01	0.16	0.02	0.21	0.21	0.18
Constant	4.12	0.51	1.53	0.70	1.31	0.57

Coef. = Coefficient. S.E. = Standard error.

a See table A.1 for definitions of variables. **b** The coefficients are estimated with reference to the base category of 'not in the labour force'. The coefficients cannot be interpreted as slope coefficients because the model is non-linear. Given the complexity of interpreting coefficient estimates, it is more informative to interpret the marginal effects of the explanatory variables (tables A.3 and A.4). The coefficients are published here for the convenience of researchers who may wish to undertake more advanced analysis. **c** The standard error is a measure of the extent to which an estimate is likely to deviate from its true value and is related to the standard deviation of the sample. The larger the sample size, the smaller the standard error.

Source: Productivity Commission estimates based on NATSISS 2008.

Table A.6 **Estimated coefficients in the model of labour market outcomes, Indigenous women aged 15 to 64 years, 2008[a,b,c]**

Variable name	*Employed*		*CDEP*		*Unemployed*	
	Coef.	*S.E.*	*Coef.*	*S.E.*	*Coef.*	*S.E.*
Explanatory variables						
Population and demographic characteristics						
Age	0.04	0.02	0.01	0.03	0.07	0.03
Age squared	-0.00	0.00	-0.00	0.00	-0.00	0.00
Married	0.00	0.08	-0.08	0.12	0.21	0.10
Dependent children	-0.34	0.03	-0.29	0.05	-0.32	0.04
Lives in multifamily household	-0.25	0.11	0.08	0.15	-0.13	0.13
Remote	0.11	0.09	1.30	0.15	-0.10	0.11
Lives in homelands	-0.25	0.09	0.26	0.13	0.03	0.11
SEIFA	0.09	0.02	-0.35	0.07	-0.01	0.02
Difficulty with English	-0.29	0.25	0.06	0.24	-0.25	0.29
Self assessed health						
Poor general health	-0.46	0.10	-0.25	0.16	-0.24	0.12
Psychological distress	-0.32	0.08	-0.11	0.13	0.20	0.10
Disability	-0.62	0.15	-0.61	0.25	-0.39	0.18
Education						
Degree or higher	0.92	0.19	-0.04	0.42	0.28	0.28
Year 12 plus non-school qualification	0.74	0.15	0.40	0.26	0.08	0.21
Year 12, no non-school qualification	0.30	0.14	0.04	0.21	-0.21	0.18
Year 10 or 11 plus non-school qualification	0.40	0.12	0.20	0.20	0.13	0.15
No year 10 plus non-school qualification	0.36	0.17	-0.07	0.33	0.30	0.21
No year 10, no non-school qualification	-0.42	0.11	-0.11	0.16	-0.24	0.12
Work experience						
Years in paid employment	0.26	0.02	0.27	0.03	0.08	0.02
Years in paid employment squared	-0.01	0.00	-0.01	0.00	-0.00	0.00
Crime						
Arrested	-0.50	0.14	0.34	0.18	0.19	0.13
Imprisoned	-0.09	0.22	-0.22	0.34	-0.26	0.24

(Continued next page)

Table A.6 (Continued)

Variable name	Employed		CDEP		Unemployed	
	Coef.	*S.E.*	*Coef.*	*S.E.*	*Coef.*	*S.E.*
Social and cultural factors						
Participation in traditional cultural activities	-0.20	0.11	0.14	0.14	0.13	0.13
Participation in social cultural events	0.42	0.09	0.31	0.13	0.13	0.12
Does not provide support outside household	-0.40	0.10	-0.14	0.15	-0.24	0.12
Provides unpaid childcare outside household	-0.21	0.09	-0.13	0.15	0.15	0.11
Constant	-0.91	0.37	-2.17	0.61	-1.70	0.45

Coef. = Coefficient. S.E. = Standard error.

a See table A.1 for definitions of variables. **b** The coefficients are estimated with reference to the base category of 'not in the labour force'. The coefficients cannot be interpreted as slope coefficients because the model is non-linear. Given the complexity of interpreting coefficient estimates, it is more informative to interpret the marginal effects of the explanatory variables (tables A.3 and A.4). The coefficients are published here for the convenience of researchers who may wish to undertake more advanced analysis. **c** The standard error is a measure of the extent to which an estimate is likely to deviate from its true value and is related to the standard deviation of the sample. The larger the sample size, the smaller the standard error.

Source: Productivity Commission estimates based on NATSISS 2008.

Table A.7 **Sample sizes and diagnostic statistics, model of labour market outcomes, Indigenous people aged 15 to 64 years, 2008**

Diagnostic statistics	*Men*			*Women*		
	Unweighted		*Weighted*	*Unweighted*		*Weighted*
	no.	*%*	*%*	*no.*	*%*	*%*
Sample and sub-sample sizes[a]						
Labour market outcome						
Mainstream (non-CDEP) employment	1 508	56.6	59.1	1 478	42.0	42.5
CDEP participation	275	10.3	8.4	184	5.2	4.6
Unemployed	299	11.2	12.1	282	8.0	8.7
Not in the labour force	580	21.8	20.4	1 576	44.8	44.2
Total	2 662	100.0	100.0	3 520	100.0	100.0
Based predicted probabilities[b]						
Labour market outcome						
Mainstream (non-CDEP) employment		89.4			66.7	
CDEP participation		0.6			0.6	
Unemployed		5.8			6.6	
Not in the labour force		4.2			26.0	
Total		100.0			100.0	
Indicators of model fit						
Log likelihood	- 1 946			- 2 629		
Likelihood ratio tests[c]	*test statistic*	*significance*		*test statistic*	*significance*	
Age	285.8	***		279.8	***	
Married	50.0	***		6.2		
Dependent children	25.4	***		141.5	***	
Lives in multifamily household	13.2	***		6.9	*	
Remote	181.1	***		92.4	***	
Lives in homelands	11.7	***		18.4	***	
SEIFA	90.8	***		95.1	***	
Difficulty with English	10.4	**		2.5		
Poor general health	95.2	***		23.1	***	
Psychological distress	20.4	***		30.4	***	
Disability	47.1	***		21.0	***	
All health	219.5	***		105.6	***	

(Continued next page)

Table A.7 (Continued)

Likelihood ratio tests[c]	test statistic	significance	test statistic	significance
Education	60.0	***	106.3	***
Work experience	336.7	***	508.7	***
Arrested	22.3	***	29.5	***
Imprisoned	19.4	***	1.4	
All crime	58.2	***	31.3	***
Participation in traditional cultural activities	0.8		8.9	**
Participation in social cultural events	12.9	***	22.8	***
Does not provide support outside household	4.9		18.1	***
Provides unpaid childcare outside	1.7		11.4	***

*** = significant at 1 per cent level (a 1 in 100 possibility that the result is due to chance); ** = significant at 5 per cent level (a 5 in 100 possibility that the result is due to chance); * = significant at 10 per cent level (a 10 in 100 possibility that the result is due to chance). No stars indicate that the variable is not statistically significant (chapter 3).

[a] The modelled sample size differs from the full sample of people aged 15-64 because some respondents did not answer all questions. Non-response mostly related to level of education. [b] The base predicted probability (table A.7) is the probability associated with the LMO of a 'base person'. The base person in this study is someone who is married, lives in a non-remote area, has no difficulty communicating in English, is in good health, has low levels of psychological distress, does not have a severe or profound disability, has a year 10 or 11 education and no non-school qualifications, has not been arrested in the last five years and has never been imprisoned. [c] A likelihood ratio test compares the log likelihoods of a full model and an identical model that excludes one or a group of explanatory variables. The row label indicates the variable or group of variables excluded. A result significantly different from zero indicates that the explanatory variable improves the fit of the model. The stars indicate the level of significance based on the critical values of the test statistic. One, two and three stars represent significance at the 10, 5 and 1 per cent levels, respectively — the more stars, the more confidence in the test result.

Source: Productivity Commission estimates based on NATSISS 2008.

References

ABS (Australian Bureau of Statistics) 2010a, *Australian Standard Geographical Classification (ASGC)*, Cat. no. 1216.0, Canberra.

—— 2010b, *National Aboriginal and Torres Strait Islander Social Survey: Users' Guide*, Cat. no. 4720.0, Canberra.

—— 2006, *Socio-Economic Indexes for Areas (SEIFA) — Technical Paper*, Cat. no. 2039.0.55.001, Canberra.

AIHW (Australian Institute of Health and Welfare) 2004, *Australia's Health 2004*, Cat. no. AUS 44, Canberra.

Bartus, T. 2005, 'Estimation of Marginal Effects Using Margeff', *The Stata Journal*, vol. 5, no. 3, pp. 309–29.

Biddle, N. and Webster, A. 2007, 'Modelling the labour force status of Aboriginal and Torres Strait Islander Australians', Australian Labour Market Research Workshop, February, Melbourne.

Biddle, N. and Yap, M. 2010, *Demographic and Socioeconomic Outcomes Across the Indigenous Australian Lifecourse: Evidence from the 2006 Census*, Research Monograph no. 31, Centre for Aboriginal Economic and Policy Research, Australian National University, Canberra.

Booth, A. and Carroll. N. 2005, *The Health Status of Indigenous and Non-Indigenous Australians*, IZA Discussion paper no. 1534, Institute for the Study of Labor, Bonn, Germany.

Borland, J. and Hunter, B. H. 2000, 'Does Crime Affect Employment Status? — the case of Indigenous Australians', *Economica*, vol. 67, no. 1, pp. 123–44.

Boushey, H. 2005, *Are Women Opting Out? Debunking the Myth, Center for Economic and Policy Research*, Briefing paper, November.

Brough, M. and Bond, C. 2009, *Social Capital and Social Justice: Critical Australian Perspectives*, editors Geoff Woolcock and Lenore Manderson, Charles Darwin University Press.

Cai, L. 2010, 'The Relationship Between Health and Labour Force Participation: Evidence From a Panel Data Simultaneous Equation Model', *Labour Economics*, vol. 17, no. 1, pp. 77–90.

—— and Kalb, G. 2005, *Health Status and Labour Force Status of Older Working-Age Australian Men*, Melbourne Institute Working Paper no. 9, Melbourne.

COAG (Council of Australian Governments) 2011, *National Indigenous Reform Agreement (Closing the Gap)*, Canberra, www.federalfinancialrelations.gov.au/content/national_agreements.aspx (accessed 14 May 2011).

—— 2008, *National Partnership Agreement on Indigenous Economic Participation*, www.federalfinancialrelations.gov.au/content/national_partnership_agreements/OT002/indigenous_economic_participation_np.rtf (accessed 4 August 2010).

CRC (COAG Reform Council) 2010, *National Indigenous Reform Agreement: Baseline Performance Report for 2008-09*, Report to the Council of Australian Governments, Sydney.

Cox, E. 1997, 'Building Social Capital', *Health Promotion Matters*, vol. 4, pp. 1–4.

Daly, A. 1995, *Aboriginal and Torres Strait Islander People in the Australian Labour Market*, Occasional paper, Cat. no. 6253, ABS, Canberra.

Freeman, R.B. 1999, 'The Economics of Crime', *Handbook of Labour Economics*, vol. 3c, part 13, ch 52, Elsevier Science.

Forbes, M., Barker, A. and Turner, S. 2010, *The Effects of Education and Health on Wages and Productivity*, Productivity Commission Staff Working Paper, Melbourne, March.

Fry, T.R.L. and Harris, M.N. 1998, 'Testing for Independence of Irrelevant Alternatives', *Sociological Methods and Research*, vol. 26, no. 3, pp. 401–23.

Gelman, A. 2007, 'Struggles with Survey Weighting and Regression Modelling', *Statistical Science*, vol. 22, no. 2, pp. 153–64.

Greene, W. 2003, *Econometric Analysis*, fifth edition, Prentice Hall, New Jersey.

Goldin, C. 2006, 'The Quiet Revolution that Transformed Women's Employment, Education, and Family', *AEA Papers and Proceedings*, vol. 96, no. 2, May.

Hunter B. and Borland, J. 1997, *The Interrelationships Between Arrest and Employment: More Evidence on the Social Determinants of Indigenous Employment*, Discussion Paper No. 136/1997, Centre for Aboriginal Economic and Policy Research, The Australian National University, Canberra.

Hunter B. and Daly A. 2008, *Interactions Between Crime and Fertility in the Labour Supply of Indigenous Australian Women*, Working Paper no. 40/2008, Centre for Aboriginal Economic and Policy Research, Australian National University, Canberra.

Hunter B. and Gray, M. 2001, Indigenous Labour Force Status Re-visited: Factors Associated with the Discouraged Worker Phenomenon, *Australian Journal of Labour Economics*, vol. 4, no. 2, pp. 111–33.

—— 1999, *Further Investigations into Indigenous Labour Supply: what Discourages Discouraged Workers?*, Working Paper no. 2/1999, Centre for Aboriginal Economic and Policy Research, Australian National University, Canberra.

Laplagne, P., Glover, M. and Shomos, A. 2007, *Effects of Health and Education on Labour Force Participation*, Productivity Commission Staff Working Paper, Melbourne, May.

Lattimore, R. 2007, *Men Not at Work: An Analysis of Men Outside the Labour Force*, Productivity Commission Staff Working Paper, Canberra, January.

Leigh, A and Gong, X. 2008, *Estimating Cognitive Gaps Between Indigenous and Non-Indigenous Australians*, Working Paper, Research School of Social Sciences, Australian National University, Canberra.

McCoull, F. and Pech, J. 2000, 'Trans-generational income support dependence in Australia: Early evidence', in Saunders, P. (ed.) *Reforming the Australian Welfare State*, Australian Institute of Family Studies, Melbourne.

Mincer, J. 1974, *Schooling, Experience and Earnings*, National Bureau of Economic Research, New York.

Murray, C. 1994, *Losing Ground: American Social Policy 1950–1980*, 2nd edn, Basic Books, New York.

PC (Productivity Commission) 2011, *Disability Care and Support,* Inquiry Report no. 54, Canberra.

—— 2003, *Social Capital: Reviewing the Concept and its Policy Implications*, Research Paper, AusInfo, Canberra.

Putman, R.. 2000, *Bowling Alone: The Collapse and Revival of American Community*, Touchstone, New York.

OECD (Organisation for Economic Cooperation and Development) 2001, *The Wellbeing of Nations: the Role of Human and Social Capital*, OECD, Paris.

Orzechowska-Fischer E. 2004, *Labour Force Ageing: Its Impact on Employment Level and Structure. The Cases of Japan and Australia*, 12th Biennial Conference, Population and Society: Issues, Research, Policy, Australian Population Association, 15–17 September, Canberra.

Portes, A. 1998, ‘Social capital: Its Origins and Applications in Modern Sociology’, *Annual Review of Sociology*, vol. 24, pp. 1–24.

Shomos, A. 2010, *Links Between Literacy and Numeracy Skills and Labour Market Outcomes*, Productivity Commission Staff Working Paper, Melbourne, August.

SCRGSP (Steering Committee for the Review of Government Service Provision) 2011, *Overcoming Indigenous Disadvantage: Key Indicators 2011*, Productivity Commission, Canberra.

—— 2009, *Overcoming Indigenous Disadvantage: Key Indicators 2009*, Productivity Commission, Canberra.

Stephens, B. J. 2010, ‘The Determinants of Labour Force Status Among Indigenous Australians’, *Australian Journal of Labour Economics*, vol. 13, no. 3, pp. 287–312.

Stone, W., Grey M., and Hughes, J. 2003, *Social Capital at Work: How Family, Friends and Civic Ties Relate to Labour Market Outcomes*, Research Paper no. 31, Australian Institute of Family Studies, Melbourne.

Vos, T., Barker, B., Stanley, L., Lopez, A.D. 2007, *The Burden of Disease and Injury in Aboriginal and Torres Strait Islander Peoples 2003*, School of Population Health, The University of Queensland, Brisbane.