F Quantitative analysis and data sources

This appendix outlines the methods, data and results of the quantitative work undertaken by the Productivity Commission to ascertain the existence and nature of disability discrimination in the Australian labour market. The results (interpreted in chapter 5 and appendix A) suggest that people with disabilities may experience some degree of discrimination in terms of both gaining employment but less so in earning equal wages.

F.1 Methods

Two analytical approaches have been used, which are referred to below as ‘decomposition techniques’ and ‘employment effects’.

Decomposition techniques are used to apportion differences in the wages received by different groups between ‘explained’ and ‘unexplained’ components. The ‘unexplained’ portion can then be interpreted as discrimination, but this interpretation relies upon a correct specification of the model.[1] These techniques have been widely used to measure sex and race discrimination (for example, Miller and Rummery 1989 and Baldwin and Johnson 1992a). Their application to disability is more difficult than to sex or race, for three main reasons:

- the potential for unobserved sources of productivity differences between the groups being compared is greater than for other types of discrimination
- the definition of disability is not as clear-cut as that of gender or race
- the potential for endogeneity between labour force status and disability status (for example, when unemployment is both caused by, and causes, a psychiatric condition).

Notwithstanding these caveats, decomposition techniques have been extended to the measurement of disability discrimination (Baldwin and Johnson 1994, 2000; and Kidd et al. 2000).

---

[1] There should be no omitted variables, for example, that might be related to the productivity of individuals, such as motivation.

The simplest and most common form of the decomposition techniques is known as the 'Oaxaca–Blinder' method and is applied to information about the wages observed for people in employment (Blinder 1973; Oaxaca 1973). An extension to this simple form incorporates the Heckman selection model, which accounts for wages that are offered but are not accepted (Reimers 1983). This is an important extension because it quantifies discrimination in offer wages (see below), which might be different from that in observed wages (Kidd et al. 2000).

Employment effects methodology focuses less on the wages of people with disabilities, and more on their employed–non-employed status. Specifically, it is used to calculate the effects on employment of:

- having a disability, which will be referred to as the 'marginal effects' technique (Wilkins 2003)
- being offered a discriminatory wage because of having a disability, which will be referred to as the 'Baldwin and Johnson' technique (Baldwin and Johnson 1992b, 1994, 2000).

Decomposition techniques

The Oaxaca–Blinder method is described below, along with its extension, the Heckman selection model. Both decomposition techniques are based on the human capital model developed by Mincer (1974), in which the central explanatory variable for wages variation is training, including formal education and on-the-job learning. In addition to these variables, firm and individual characteristics are controlled for (for example, occupation, union membership, firm size and public sector). This model is described in equation F.1 below.

$$Y_j = \sum_{i=1}^{M} \beta_i X_{ij} + \mu_j \tag{F.1}$$

where,

Y_j = logged hourly wages of individual *j*

X_{ij} = *i*th explanatory variable of individual *j* (which might be a human capital or control variable)

β_i = coefficient of the *i*th explanatory variable

μ = normally distributed error term.

Oaxaca–Blinder method

This method decomposes the difference between the average wages of people without disabilities and people with disabilities ($\overline{Y_N} - \overline{Y_D}$) into two components:

- The 'explained component', based on the difference between the averages of the explanatory variables (the first term on the right hand side of equation F.2 below). This component is so termed because it is attributable to the two groups having different characteristics.
- The 'unexplained component', based on the difference between the coefficients of the explanatory variables (the second term on the right hand side of equation F.2 below). This component is so termed because it is attributable to the same characteristics in the two groups being treated differently. It is this component which has been interpreted as measuring discrimination.

$$\overline{Y_N} - \overline{Y_D} = \sum_{i=1}^{M} \beta_i^N \left(\overline{X_i^N} - \overline{X_i^D} \right) + \sum_{i=1}^{M} \left(\beta_i^N - \beta_i^D \right) \overline{X_i^D} \qquad \text{(F.2)}$$

where,

$\overline{Y}$ = average of logged hourly wages

$\overline{X_i}$ = average of the *i*th explanatory variable

β_i = coefficient of the *i*th explanatory variable

N denotes people without disabilities

D denotes people with disabilities

This form of the model uses people without disabilities for its non-discriminatory wage benchmark; that is, it assumes in the absence of wage discrimination, people with disabilities would receive the same return on their human capital characteristics as people without disabilities (instead of, for example, returns to both groups adjusting). It can be argued that this assumption is plausible because people with disabilities only represented 11.7 per cent of the labour force in 1998 (ABS 1999b, p. 35), hence their influence on the labour market and wage rates would have been relatively small (Kidd et al. 2000, p. 970). An equivalent assumption may not hold in the case of sex or race discrimination. In any event, changing the assumptions so that wages of people without disabilities adjust downward as those of people with disabilities adjust upward when discrimination is removed does not alter the difference in reward that the same characteristic attracts for members of each group. It simply relabels part of the unexplained component as due to 'nepotism' (favouring people without disabilities) rather than disability discrimination.

Heckman selection model

As noted, the Heckman selection model is an extension of the Oaxaca–Blinder decomposition technique, run in two stages:

- a probit regression to model the decision to participate in the labour market (hence the likelihood of wages being observed)
- an ordinary least squares (OLS) regression to model the determinants of those wages that are observed.

The variables in the probit regression relate mainly to factors influencing people's trade-off between income and leisure, while those in the OLS regression relate mainly to factors influencing people's productivity.[2] There is overlap between these factors, as some characteristics influence both income and leisure trade-offs and productivity.

Wages are observed only if the wages that employers offer to people (offer wages) are greater than or equal to the wages that people are prepared to work for (reservation wages). Hence, observed wages are only a subset of offer wages, and might exhibit a particular pattern. Such a pattern is known as sample selection bias and is frequently explained in terms of the observed wages distribution being a 'truncated' version of the offer wages distribution. The Heckman selection model corrects for the non-random sampling of observed wages from offer wages, through a so-called 'inverse Mills ratio' based on the probit regression, which captures the relationship between observed and offer wages. This ratio is then included as an explanatory variable in the OLS regression, as shown in equation F.3.

$$Y_j^{OB} = \sum_{i=1}^{M} \beta_i \left(X_{ij} \right) + \gamma\lambda_j + \mu_j \qquad \text{(F.3)}$$

where,

Y_j^{OB} = logged hourly observed wages of individual j

[2] The variables in the probit regression should include those variables that relate to offer wages (which already appear in the OLS regression), as well as those relating to reservation wages (or the income-leisure trade-off), if the likelihood of wages being observed is to be comprehensively modelled. However, this is not possible since some of the variables in the OLS regression are only available for people in employment. This is a limitation of Reimer's approach (Baldwin and Johnson 1994, p. 5) and it might explain why the inverse Mills ratio is frequently found to be insignificant in research that calculates it.

X_{ij} = ith explanatory variable of individual j

β_i = coefficient of the ith explanatory variable

λ_j = inverse Mills ratio of individual j

γ = inverse Mills ratio's coefficient

μ = normally distributed error term.

The inverse Mills ratio can be used to infer the average offer wages from average observed wages in the manner depicted in equation F.4.

$$\overline{Y^{OF}} = \overline{Y^{OB}} - \gamma \overline{\lambda} = \sum_{i=1}^{M} \beta_i \left(\overline{X}_i \right) \qquad \text{(F.4)}$$

where,

$\overline{Y}$ = average of logged hourly wages (OF superscript denotes offer, OB superscript denotes observed)

$\overline{X}_i$ = average of the ith explanatory variable

β_i = coefficient of the ith explanatory variable

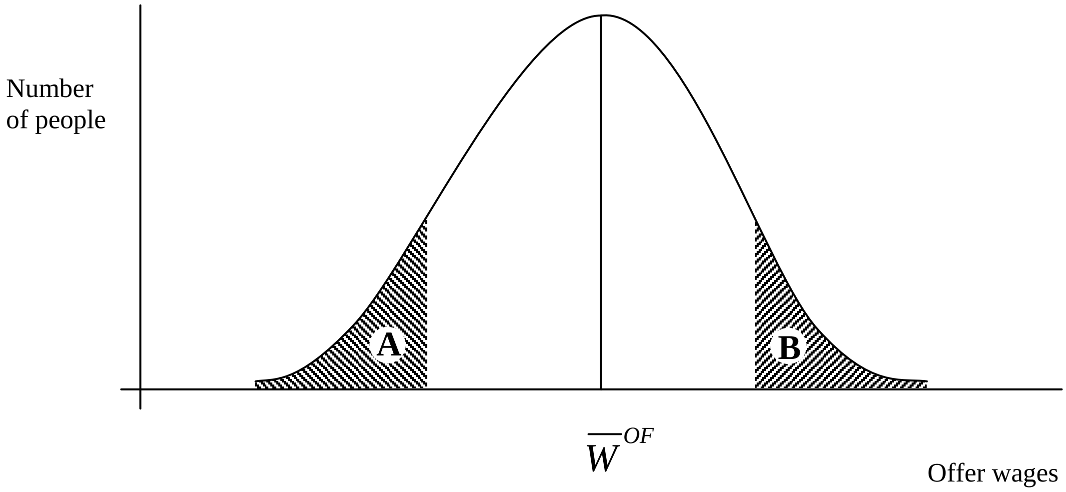
$\overline{\lambda}$ = average of the inverse Mills ratio

γ = inverse Mills ratio's coefficient.

A positive coefficient for the inverse Mills ratio in the OLS regression means that observed wages are greater on average than offer wages: that is, above average wage offers tend to be accepted and become observed wages, but below average ones are not (equation F.4). This is expressed in figure F.1 as area A missing from the offer wages distribution, leaving the distribution 'truncated'. A negative coefficient for the inverse Mills ratio means that observed wages are less than offer wages: that is, below average wage offers tend to be accepted and become observed wages, but above average ones are not (equation F.4). This is expressed in figure F.1 as area B missing from the offer wages distribution.

It is difficult to predict a priori whether the inverse Mills ratio coefficient for people with disabilities will be positive or negative. Previous research in this area for the US and UK has found both negative and positive coefficients (Baldwin and Johnson 1994 and Kidd et al. 2000). A positive coefficient is indicative that those people with disabilities with the greatest earning opportunities are relatively more likely to be observed in the labour market. A negative coefficient would indicate that those with relatively low earnings prospects tend to self-select into employment.

Figure F.1 **Truncation of offer wage distribution leading to selectivity bias**



The results from the Heckman selection model can be decomposed as shown in equation F.5, so that the gap in offer wages is broken down into explained and unexplained components (the first and second terms on the right hand side of equation of equation F.5 respectively). This approach was developed by Reimers (1983).

$$\overline{Y_N^{OF}} - \overline{Y_D^{OF}} = \left(\overline{Y_N^{OB}} - \overline{Y_D^{OB}}\right) - \left(\gamma_N\,\overline{\lambda}_N - \gamma_D\,\overline{\lambda}_D\right) = \sum_{i=1}^{M} \beta_i^N \left(\overline{X_i^N} - \overline{X_i^D}\right) + \sum_{i=1}^{M}\left(\beta_i^N - \beta_i^D\right)\overline{X_i^D} \qquad \text{(F.5)}$$

where,

$\overline{Y}$ = average of logged hourly wages (OF superscript denotes offer, OB superscript denotes observed)

$\overline{X}_i$ = average of the ith explanatory variable

β_i = coefficient of the ith explanatory variable

$\overline{\lambda}$ = average of the inverse Mills ratio

γ = inverse Mills ratio's coefficient

N denotes people without disabilities

D denotes people with disabilities

Previous research has generally shown that the gap between offer wages received by people without disabilities and people with disabilities is greater than that between observed wages (Baldwin and Johnson 1994, 2000 and Kidd et al. 2000).

Employment effects

Two methods have been used to estimate the effect of having a disability on employment, where employment is defined as spending an hour or more a week working in a job for which a person receives some kind of payment. The first method ('marginal effects') examines the marginal effect of having a disability upon the likelihood of employment (Wilkins 2003). The second method ('Baldwin and Johnson') examines the effect of discriminatory wage offers upon employment outcomes (Baldwin and Johnson 1992b, 1994, 2000). Whereas the first method collates all possible reasons for people with disabilities not being in employment, the second focuses on discriminatory wage offers as the reason for non-employment.

Marginal effects technique

The marginal effects technique measures the change that occurs in the probability of employment if a dummy variable which represents disability becomes 'active' (that is, changes from 0 to 1). It uses the employment probability model used by Wilkins for 1998 (2003, p. 39–41), and applies it to 1993 data (1998 and 1993 are the only two years for which SDAC unit-record data are available) (equation F.6). This approach controls for variables other than disability, such as age, education and family status. Unlike the decomposition techniques previously described, however, it does not constitute an explicit attempt to measure discrimination. Nonetheless, a negative impact of the disability dummy on the probability of being employed might be interpreted as being at least partly due to discrimination (subject to the usual caveats about model specification).

The probit model can be represented as follows:[3]

$$\Pr\left(E_j = 1 \mid X_{ij}\right) = F\left(\sum_{i=1}^{M} \beta_i X_{ij}\right) \qquad \text{(F.6)}$$

where,

[3] For more information about probit regression functions, see Stata Corporation 1999, *Reference Manual: Stata Release 6*, Stata Press, Texas.

$\Pr\left(E_j = 1 | X_{ij}\right)$ = the probability that individual j is employed, given the characteristics of individual j

X_{ij} = ith explanatory variable of individual j (chosen to reflect the determinants of the trade-off between income and leisure)

β_i = coefficient of the ith explanatory variable

F = normal cumulative distribution function

Baldwin and Johnson technique

This technique relies upon the Heckman selection model and uses a method outlined by Heckman (1976) but developed further by Baldwin and Johnson (1992b, 1994, 2000). It estimates the amount of employment lost through people with disabilities being offered discriminatory wages which are less than the non-discriminatory offer wages. Employment loss occurs because some people with disabilities would work for the non-discriminatory wage but will not work for the discriminatory wage. However, the total employment loss arising from discriminatory wage offers is underestimated, because employment discrimination can also manifest itself as people with disabilities working fewer hours. Baldwin and Johnson's estimate of employment loss only accounts for those hours 'lost' by people with disabilities not working at all.

There are three stages in the Baldwin and Johnson technique:

1. the probability of employment of the 'average' person with disabilities is estimated in the presence of wage discrimination, using the coefficients from the probit regression of the Heckman selection model
2. the probability of employment of the average person with disabilities in the absence of discrimination is estimated
3. the difference between discriminatory and non-discriminatory probabilities of employment is applied to the population of people with disabilities to calculate the amount of employment lost through wage discrimination.

The second stage of the Baldwin and Johnson technique is described in detail in Baldwin and Johnson (1992a and 1992b). It contains two key assumptions:

1. the larger the excess of the offer wage over a person's reservation wage, the more hours that person is likely to work
2. discrimination reduces offer wages made to people with disabilities but has no effect on the offer wages made to people without disabilities.

The second assumption is more controversial than the first and is typically applied to disability discrimination more than other types of discrimination, such as sex or race. The justification for this might be the relatively small number of people with disabilities in employment relative to people without disabilities, and the consequent lesser impact that non-discriminatory outcomes for this group are likely to have on the labour market as a whole, relative to other types of discrimination.

F.2 Data

Data from the first wave of the Department of Family and Community Services' Household, Income and Labour Dynamics in Australia (HILDA) survey conducted in 2001 are used to carry out decomposition analysis and the Baldwin and Johnson technique. These data are preferable to data from the ABS Survey of Disability, Ageing and Carers (SDAC) cited elsewhere in this report for two main reasons:

- HILDA contains detailed wages data, while SDAC only contains data about income, which is a relatively crude proxy for wages
- HILDA contains far more other variables relevant to employment, which can only be proxied (for example, work experience) or are not obtainable at all (for example, firm size and union membership) from SDAC.

However, SDAC typically covers far more people than HILDA (approximately 40 000 instead of 14 000) and contains data suitable for analysis of changes in employment outcomes since the DDA was introduced (1993 and 1998).

HILDA variables

A number of HILDA variables were used to carry out the analysis of wages of people with disabilities and people without disabilities in 2001 (table F.1).

SDAC variables

A number of SDAC variables were used to examine the effects of disability on the likelihood of employment in 1993 and 1998 (table F.2). In order for the analysis to be consistent over time, it was necessary to adjust the definition of the variables found in each of the two SDACs, including the definition of disability (table F.3).

Table F.1 **HILDA variables used in the analysis, 2001**

Analysis variable	*Description*	*HILDA variable(s)*
Demography		
MARRIED	Married	AHGMS
AGE1	Age	AHHFAG
AGE1SQ	Age squared	AHHFAG
NC04	Number of children aged 0 to 4	ATCN04
		ATCR04
NC514	Number of children aged 5 to 14	ATCN514
		ATCR514
CITY	City dweller	AHHRA
NESC	Born in non-English speaking country[a]	AANCOB
Health		
DISABIL	Disability status	AHELTH
HEALTH	Current health status	AHESA
Education		
POSTDEG	Doctorate, masters, graduate diploma or graduate certificate	AEDHIGH
BACHELOR	Bachelor degree	AEDHIGH
DIPLOMA	Advanced diploma or diploma	AEDHIGH
CERTY	Certificate III, IV, II, I or not defined	AEDHIGH
YR12	Completed Year 12	AEDHIGH
YR11	Completed Year 11 or below	AEDHIGH
Employment and income		
EMPLOY	Employed[b]	AESBRD
LWPH	Log of gross wages per hour in main job	AWSCME
		AJBHRU
		AJBMHRU
TRDUN	Trade union member	AJBMUABS
NWAGEINC	Non wage income[c]	ABNC
		ABIFP
		ABIFN
		AOIFINVP
		AOIFINVN
		AOIFOTHT
YRSOCC	Years spent in current occupation	AJBOCCT
YRSU	Years spent in unemployment	AEHTUJMT
		AEHTUJYR
FIRMLTW	Business has less than 20 employees	AJBMEMSZ
		AJBMWPSZ
PTM	Works less than 35 hours a week	AJBHRU
		AJBMHRU
		AJBPTREA
GOVTBUSI	Government-owned business	AJBMMPLR
Occupation		
MANAGER	Managers and administrators	AJBMOCC2
PROF	Professionals	AJBMOCC2

(continued next page)

Table F.1 (continued)

APROF	Associate professionals	AJBMOCC2
TRADE	Tradespersons and related workers	AJBMOCC2
CLERKY	Advanced, intermediate, or elementary clerical, sales and service workers	AJBMOCC2
IPROD	Intermediate production and transport workers	AJBMOCC2
LABOUR	Labourers and related workers	AJBMOCC2
Industry		
INDUST1	Primary[d]	AJBMIND2
INDUST2	Manufacturing	AJBMIND2
INDUST3	Utilities and resource movement[e]	AJBMIND2
INDUST4	Government[f]	AJBMIND2
INDUST5	Recreational and personal services[g]	AJBMIND2
INDUST6	Business services[h]	AJBMIND2

[a] Excludes: Australia, New Zealand, UK, Channel Islands, Ireland and Eire, Canada, America and South Africa. [b] A person is determined to be employed if they have responded affirmatively to receiving a wage or salary and spending weekly hours in a job. [c] This year's pensions and benefits and other income and last year's business income and investment income. [d] Agriculture, forestry and fishery; Mining. [e] Construction; Transport and storage; Electricity, gas and water. [f] Government administration and defence; Education; Health and community services. [g] Wholesale trade; Retail trade; Accommodation, cafes and restaurants; Cultural and recreational services; Personal and other services. [h] Communication services; Finance and insurance; Property and business services.

In both SDACs, the existence of disability was determined by reference to screening questions about whether a list of limitations, restrictions or impairments had lasted, or was likely to last, for a period of 6 months or more. It is necessary to 'match' the list of screening questions from SDAC in 1993 and 1998 in order to create a similar definition of disability for analysis over time (table F.3). This matching exercise fails to remove all differences, in particular, some wording changes in the screening questions and the effect of new screening questions upon the responses to pre-existing questions.

F.3 Results

Oaxaca–Blinder decomposition

Equation F.1 is estimated for people without disabilities and people with disabilities who are of working age (15 to 64 years old) (table F.4). The dependent variable is the log of wages per hour and the explanatory variables are a combination of human capital and control variables, the choice of which is influenced by previous research in this area, such as Baldwin and Johnson (1994, 2000), Kidd et al. (2000) and Kidd and Viney (1989). It is common practice to run these regressions separately for men and women, because of the different employment experiences of these two groups, and frequently only the results for men are presented in overseas research (Baldwin and Johnson 1994, 2000 and Kidd et al. 2000).

Table F.2 **SDAC variables, 1993 and 1998**

		SDAC variable(s)	
Analysis variable	*Description*	1993	1998
Employment			
EMPLOY	Employed	lbfstatus	lbf400
Health			
DISABIL	Has a disability	disabled	dis401 with adjustments[a]
Age			
AGE1524	Between 15 and 24	Ageurf	psn401
AGE2534	Between 25 and 34	Ageurf	psn401
AGE3544	Between 35 and 44	Ageurf	psn401
AGE4554	Between 45 and 54	Ageurf	psn401
AGE5564	Between 55 and 64	Ageurf	psn401
Education			
DEGREE	Bachelor or higher degree	hiedulev	edn412
OTHPS	Post-school education other than bachelor or higher degree	hiedulev	edn412
CHS[b]	Completed high school	hiedulev	edn412
		alsurf	edn411
		typcours	
NCHS	Did not complete high school	hiedulev	edn412
		alsurf	edn411
		typcours	
Country of birth			
AUSB	Australia	Coburf	psn403
ESB[c]	Main English-speaking country	Coburf	psn403
NESB	Non-English speaking country	Coburf	psn403
Family circumstances			
S_NODEP	Single with no dependents	famtype	fam200
C_NODEP	Couple with no dependents	famtype	fam200
S_DEP	Single with dependents	famtype	fam200
C_DEP	Couple with dependents	famtype	fam200

a The adjustments carried out to the disability definition in 1998 to make it similar to 1993 are described in table F.3. **b** There is no equivalent variable in 1993 for the 1998 variable edn411 ('completed year 12'). Instead, this variable is proxied using information about whether people: completed a high school certificate post-school (hiedulev), left school aged 18 or older (alsurf) and whether a current course of study is for an associate diploma, undergraduate diploma or bachelor degree (typcours – because these people are assumed to have completed Year 12 or equivalent). **c** 1998 includes: New Zealand, UK and Ireland, balance of Oceania and Antarctica and Oceania and Antarctica nfd. 1993 includes: New Zealand, UK and Ireland and other Oceania.

Table F.3 **Matched definition of disability in SDAC, 1993 and 1998[a]**

Description of screening question[b]	*CURF variable name* 1993	1998
Screening questions included in the 'matched' definition of disability		
Loss of sight not corrected by (*even when wearing*) glasses or contact lenses	sight	cnd4156a
Loss of hearing	hearing	cnd4021
A speech difficulty (*in native languages*)	speech	cnd4156c
Difficulty (*slowness at*) learning or understanding	slowund	cnd4156g
Needs help or supervision due to mental illness	mental	cnd4156n
Blackouts, fits or loss of consciousness	fitsec	cnd4156f
Incomplete use of arms or fingers	usearms	cnd4156h
Difficulty gripping or holding things (*small objects*)	diffgrip	cnd4156i
Incomplete use of legs or feet	uselegs	cnd4156j
Restricted in physical activities or in doing physical work	resphys	cnd4156l
Nervous or emotional condition, for which receives treatment	nerves	cnd4112
Disfigurement or deformity	deform	cnd4156m
Long-term effects of head injury, stroke or other brain damage causing restriction	brdam	cnd4156o
Receiving treatment or medication for other long-term condition or ailments and still restricted	longterm	cnd4156p
Any other additional long-term condition resulting in a restriction	othcnd	cnd4156p
Screening questions omitted from the 'matched' definition of disability		
Loss of hearing where communication is restricted, or an aid to assist with, or substitute for, hearing is used		cnd4156b
Shortness of breath or breathing difficulties causing restriction		cnd4156d
Chronic or recurrent pain or discomfort causing restriction		cnd4156e
Nervous or emotional condition causing restriction		cnd4156k

[a] This matched definition is the same as the ABS' original definition of disability for 1993 but is altered for 1998 to be the same as Wilkins' 'inclusive' match (2003, p. 75), which is different from the ABS' criteria-adjusted definition of disability for 1998 (1999d, p. 18), because it contains one less 1998 criterion and only people that are receiving treatment from a nervous or emotional condition. [b] Where there are differences in the wording of the screening questions between 1993 and 1998, the 1998 wording is used but the 1993 wording is shown in brackets, for example, 'difficulty' is used but 'slowness at' is shown in brackets in relation to learning or understanding.

The base case in table F.4 is: non-city dweller, English-speaking background, firm size greater than 20 employers, full-time worker, non-trade union member, non-government business, no formal qualifications, labourer or related worker and primary industry. The results in table F.4 accord well with prior expectations: work experience and education have a positive effect upon wages, but a non-English-speaking background, spells of unemployment, small firm size and poor health (except for women with disabilities) have a negative effect upon wages. Most variables are significant in regressions for people without disabilities. For people with disabilities, variables are generally less significant, possibly as a result of the smaller sample size. The relatively low explanatory power of the four regressions, as measured by their R-squared values, is expected for Mincer earnings equations. Nonetheless, the proportion of the variation in log hourly wages explained by the model is comparable to that obtained by similar overseas studies (for example, Baldwin and Johnson 1994, 2000).

The decomposition results based on equation F.2 and table F.4 are presented in table F.5. The unexplained component for men is a higher percentage of the total wage gap than that for women.

Table F.4 **Results of regression of log hourly wages**

Explanatory variables	*People without disabilities*				*People with disabilities*			
	Men		Women		Men		Women	
	Mean	Coef[a]	Mean	Coef	Mean	Coef	Mean	Coef
YRSWK	18.7	0.03***	16.1	0.03***	22.0	0.02**	17.9	0.00
YRSWKSQ	488	0.00***	361	0.00***	642	0.00*	437	0.00
YRSOCC	9.19	0.01***	8.07	0.00***	9.82	0.00	8.99	0.00
YRSU	0.45	-0.02**	0.29	-0.01*	0.66	-0.01	0.32	-0.04
CITY	0.64	0.08***	0.63	0.04**	0.57	0.05	0.60	0.11**
NESC	0.13	-0.04	0.13	-0.01	0.10	0.01	0.09	0.02
FIRMLTW	0.23	-0.17***	0.24	-0.07***	0.22	-0.16***	0.24	-0.09
PTM	0.12	0.05	0.46	0.06***	0.15	-0.02	0.50	0.12**
TRDUN	0.33	0.07***	0.31	0.03	0.37	0.15***	0.32	0.11*
GOVTBUSI	0.22	0.02	0.31	0.03	0.27	0.06	0.37	0.14**
POSTDEG	0.09	0.30***	0.11	0.25***	0.08	0.43***	0.11	0.20**
BACHELOR	0.15	0.26***	0.21	0.18***	0.11	0.26***	0.18	0.16
DIPLOMA	0.09	0.11***	0.10	0.12***	0.09	0.18**	0.10	0.03
CERTY	0.34	0.05**	0.24	0.04	0.40	0.11*	0.28	0.10
YR12	0.14	0.06**	0.14	0.05*	0.10	0.24***	0.13	0.12
MANAGER	0.09	0.34***	0.04	0.43***	0.08	0.27**	0.04	0.51***
PROF	0.21	0.22***	0.29	0.28***	0.16	0.17	0.26	0.39**
APROF	0.11	0.23***	0.12	0.19***	0.12	0.23**	0.12	0.33**
TRADE	0.18	0.11***	0.03	-0.01	0.18	0.16	0.04	0.08
CLERKY	0.17	0.07**	0.44	0.09***	0.21	0.06	0.43	0.19
IPROD	0.14	0.05	0.02	-0.06	0.14	0.12	0.03	0.07
HEALTH	2.16	-0.01	2.09	-0.03***	2.98	-0.04*	2.93	0.02
INDUST2	0.17	0.00	0.06	-0.03	0.15	-0.03	0.07	0.16
INDUST3	0.17	0.01	0.04	0.04	0.17	0.02	0.03	-0.27
INDUST4	0.17	-0.12***	0.43	-0.10	0.22	-0.13	0.47	0.04
INDUST5	0.26	-0.15***	0.27	-0.10	0.26	-0.14*	0.30	0.00
INDUST6	0.17	0.10**	0.18	0.03	0.13	0.03	0.12	0.12
Constant		2.36***		2.38***		2.47***		2.08***
Number of observations		2618		2590		423		352
R-squared		0.37		0.28		0.32		0.27

[a] *** = statistically significant at the 1 per cent level, ** = statistically significant at the 5 per cent level and * = statistically significant at the 10 per cent level.

Table F.5 **Results of the Oaxaca–Blinder decomposition**

Components	*Men*	*Women*
Total gap in logged hourly wages	0.060	0.052
Of which		
Explained	0.031	0.041
Unexplained	0.030	0.010
Unexplained percentage of total wage gap	49.2	20.0

Heckman selection model

Equation F.3 is estimated for people without disabilities and people with disabilities who are of working age (15 to 64 years old) (table F.7). The choice of variables for the probit regression in table F.6 and the OLS regression in table F.7 was influenced by previous research in this area (Baldwin and Johnson 1994; 2000 and Kidd et al. 2000).

The first stage of the Heckman model is the probit regression (table F.6). The base case in table F.6 is: English-speaking background, no formal qualifications, no children, not married and non-city dweller. Again, the results in table F.6 accord well with prior expectations: a non-English speaking background, poor health and children generally reduce the likelihood of employment, but education and living in a city generally increase it. The effect of marriage depends on gender. Surprisingly, non-wage income increases labour market participation, but this might reflect synergies between participation in the labour market and opportunities to accumulate income-producing assets.

In general, the sign and significance of the coefficients in table F.7 are broadly similar to those in table F.4 for corresponding variables. The inverse Mills ratio only appears in the Heckman OLS regression; its coefficient is negative for all groups in table F.7 except for women with disabilities. Baldwin and Johnson (1994) found a negative inverse Mills ratio for all men while Kidd et al. (2000) found a positive inverse Mills ratio for all men. However, the inverse Mills ratio is not statistically significant in any of the regressions in table F.7. This accords with Kidd et al. (2000) but goes against Baldwin and Johnson (1994), who found the inverse Mills ratio for men without disabilities to be significant at the 1 per cent level.

Converting the information contained in table F.7 into decomposition results based on equation F.5 yields a wage offer gap greater than the observed wage gap for women, but smaller for men (table F.8). The negative result for men is the reverse of that found overseas by Baldwin and Johnson (1994, 2000) and Kidd et al. (2000). However, the lack of statistical significance of the inverse Mills ratio in table F.7

implies that the offer wages gap cannot be said to be statistically different from the observed wages gap.

Table F.6 **Results from labour market participation probit regression in Heckman selection model**[a]

Explanatory variables	*People without disabilities*				*People with disabilities*			
	Men		Women		Men		Women	
	Mean	Coef	Mean	Coef	Mean	Coef	Mean	Coef
AGE1	36.3	0.19***	36.2	0.22***	43.2	0.11***	43.1	0.13***
AGE1SQ	1491	0.00***	1475	0.00***	2049	0.00***	2036	0.00***
NESC	0.14	-0.45***	0.15	-0.51***	0.14	-0.57***	0.14	-0.36**
POSTDEG	0.08	0.92***	0.07	1.08***	0.05	0.65***	0.05	1.60***
BACHELOR	0.13	0.62***	0.16	0.87***	0.07	0.99***	0.11	1.16***
DIPLOMA	0.08	0.56***	0.09	0.64***	0.08	0.34*	0.07	1.00***
CERTY	0.31	0.42***	0.22	0.52***	0.36	0.5***	0.24	0.51***
YR12	0.13	0.54***	0.13	0.60***	0.08	0.52***	0.10	0.76***
HEALTH	2.13	0.04	2.13	-0.04	3.30	-0.29***	3.36	-0.39***
NC04	0.22	0.00	0.25	-0.68***	0.19	-0.04	0.14	-0.33***
NC514	0.47	-0.12***	0.52	-0.25***	0.42	-0.09	0.39	-0.04
NWAGEINC	7659	0.00***	4587	0.00***	10572	0.00***	7969	0.00***
MARRIED	0.50	0.3***	0.51	-0.24***	0.52	0.46***	0.50	-0.09
CITY	0.62	0.11**	0.61	0.01	0.51	0.24**	0.54	0.17*
Constant		-2.97***		-3.27***		-1.28***		-1.40***
Pseudo R-squared		0.22		0.24		0.26		0.29
Number of observations		3463		4057		992		887

a *** = statistically significant at the 1 per cent level, ** = statistically significant at the 5 per cent level and * = statistically significant at the 10 per cent level.

It is an inherent problem of the Heckman approach that the existence of sample selection bias can be very difficult to detect when the regression for the minority group relies on a very small sample, as is the case here (386 men and 312 women with disabilities). Despite the lack of significance of the inverse Mills ratio in table F.7, the presumption remains that sample selection bias is present in observed wages for people with disabilities. Some inquiry participants argued that the additional costs (including the loss of government income support) that having a disability impose on a person create a barrier to involvement in the labour force (Disability Services Commission, Western Australia, sub. 44; Australian Federation of Deaf Societies, sub. DR363). This suggests that some people with disabilities who receive wage offers choose not to accept them, because it would be unprofitable for them to do so. It can be expected, therefore, that the distribution of observed wages for people with disabilities will be significantly different from that of the offer wages they receive. In that case, offer wages should be used in preference to observed wages in measuring discrimination.

Table F.7 **Results from wage determination OLS regression in Heckman selection model**[a, b]

Explanatory variables	*People without disabilities*				*People with disabilities*			
	Men		Women		Men		Women	
	Mean	Coef	Mean	Coef	Mean	Coef	Mean	Coef
YRSWK	18.4	0.03***	15.7	0.03***	21.6	0.02**	17.6	0.00
YRSWKSQ	478	0.00***	347	0.00***	623	0.00*	430	0.00
YRSOCC	9.03	0.01***	7.81	0.00***	9.50	0.00	9.18	0.00
YRSU	0.46	-0.02**	0.30	-0.01**	0.66	-0.01	0.33	-0.04
CITY	0.64	0.07***	0.64	0.04**	0.56	0.04	0.60	0.12**
NESC	0.13	-0.04	0.13	0.00	0.10	0.00	0.08	0.05
FIRMLTW	0.23	-0.16***	0.24	-0.07***	0.23	-0.12**	0.24	-0.10
PTM	0.12	0.04	0.46	0.06***	0.15	-0.03	0.50	0.14**
TRDUN	0.32	0.08***	0.30	0.02	0.36	0.16***	0.31	0.12*
GOVTBUSI	0.22	0.02	0.31	0.04*	0.26	0.06	0.38	0.09
POSTDEG	0.09	0.29***	0.11	0.23***	0.08	0.43***	0.10	0.26**
BACHELOR	0.15	0.25***	0.21	0.16***	0.11	0.28***	0.18	0.19
DIPLOMA	0.08	0.12***	0.10	0.09**	0.08	0.17*	0.11	0.06
CERTY	0.34	0.05**	0.24	0.01	0.40	0.10	0.26	0.15*
YR12	0.14	0.06*	0.14	0.04	0.10	0.24**	0.14	0.15
MANAGER	0.09	0.33***	0.03	0.4***	0.07	0.20	0.04	0.49***
PROF	0.21	0.23***	0.28	0.28***	0.16	0.17	0.25	0.38**
APROF	0.11	0.24***	0.12	0.19***	0.12	0.18	0.12	0.35**
TRADE	0.19	0.1***	0.03	0.00	0.18	0.17	0.04	0.01
CLERKY	0.17	0.07**	0.44	0.09***	0.22	0.05	0.44	0.15
IPROD	0.14	0.06*	0.02	-0.05	0.14	0.13	0.04	0.03
HEALTH	2.16	-0.01	2.09	-0.04***	2.95	-0.03	2.93	0.02
INDUST2	0.17	0.00	0.06	-0.06	0.16	-0.03	0.06	0.19
INDUST3	0.17	0.01	0.04	0.01	0.17	0.01	0.03	-0.31
INDUST4	0.16	-0.11**	0.43	-0.13	0.22	-0.12	0.47	0.07
INDUST5	0.26	-0.14***	0.28	-0.13	0.26	-0.14*	0.30	0.01
INDUST6	0.16	0.12**	0.18	0.01	0.12	0.03	0.12	0.13
Constant		2.35***		2.45***		2.46***		2.06***
Inverse Mills ratio	0.38	-0.01	0.50	-0.03	0.72	-0.04	0.73	0.03
R-squared		0.38		0.28		0.32		0.27
Number of observations		2417		2353		386		312

a *** = statistically significant at the 1 per cent level, ** = statistically significant at the 5 per cent level and * = statistically significant at the 10 per cent level. **b** Following Kidd and Viney (1989), White's heteroscedasticity corrected standard errors have been used to correct for heteroscedasticity introduced into the model by the inclusion of the inverse Mills ratio.

Because of the gap between offer and observed wages, the percentage of the total wage gap that is unexplained by the Heckman model (table F.8) differs from that obtained from the Oaxaca–Blinder decomposition (table F.5). Relative to the results presented in table F.5, table F.8 contains an unexplained component which is a greater percentage of the total wage gap for women but a smaller percentage for men. Moreover, the unexplained percentages by gender are almost the reverse of those obtained from the Oaxaca–Blinder decomposition. This underlines the importance of undertaking a decomposition based on offer wages (the Heckman selection model) when there is a likelihood of sample selection bias affecting observed wages.

Table F.8 **Results of the Heckman selection model decomposition**

Components	*Men*	*Women*
Total gap in logged hourly wages		
Observed	0.065	0.043
Offer	0.043	0.079
Of which		
Explained	0.032	0.044
Unexplained	0.011	0.035
Unexplained percentage of total offer wage gap	26.6	44.0

Employment effects

The loss in employment of people with disabilities arising from discriminatory wage offers is estimated using the Baldwin and Johnson extension of the Heckman model (1992b, 1994, 2000) (table F.9). Following these studies, this exercise was undertaken for men only.

Table F.9 **Employment effects for men with disabilities, 2001**

	Men with disabilities
Population with a disability[a]	1 357 367
Estimated discriminatory probability of employment (%)	34.36
Estimated probability of employment in absence of discrimination (%)	35.40
Differences in probabilities of employment	1.04
Jobs not taken as a result of the disincentives of discrimination	14 128

a This includes people in employment and people not in employment.

The Baldwin and Johnson technique suggests that 14 128 men with disabilities in Australia in 2001 were not in employment because of discriminatory wages offered to them. That is, the wage offers they received were below their reservation wage. Compared to Baldwin and Johnson’s estimate for the United States (2000, p. 559), this result represents a larger proportion of men in the general population (0.2 per cent compared to 0.01 per cent), but it is a similar proportion of the population of men with disabilities (1.0 per cent compared to 1.3 per cent). This difference might partly reflect the broader definition of disability in Australian data compared to that overseas.

Equation F.6 is used to analyse the effects of disability on the employment status of working-age people (15 to 64 years old) living in households, using SDAC data from 1993 and 1998 and a model developed by Wilkins (2003) (table F.10). The explanatory power of the probit model, as measured by the Pseudo R-Squared statistic, is low, although most coefficients are significant. The coefficient of the disability variable is negative and highly significant, for men and women in both years. Comparison of the marginal effects between 1993 and 1998 suggests that having a disability became less detrimental to the employment prospects of men over time. It is not possible to ascertain whether this was due to, for example, men with disabilities undergoing a change in their preferences for work or being less subject to discrimination in employment. For women, the adverse impact on employment of having a disability increased over time. This might be due in part to women's employment patterns becoming more similar to men's over time. In both 1993 and 1998, disability had a more detrimental effect on the likelihood of employment for men than for women.

Table F.10 **Probit analysis of employment outcomes, 1993 and 1998**

Explanatory variables	*1993*				*1998*			
	Men		Women		Men		Women	
	Mean	dF/dx[a,b]	Mean	dF/dx	Mean	dF/dx	Mean	dF/dx
AGE2534	0.23	0.189***	0.24	0.141***	0.22	0.155***	0.22	0.059***
AGE3544	0.23	0.243***	0.24	0.246***	0.23	0.191***	0.24	0.164***
AGE4554	0.18	0.223***	0.18	0.118***	0.21	0.167***	0.20	0.073***
AGE5564	0.13	0.045***	0.12	-0.280***	0.13	0.000	0.13	-0.296***
DEGREE	0.13	0.150***	0.11	0.313***	0.14	0.147***	0.14	0.283***
OTHPS	0.37	0.149***	0.29	0.224***	0.36	0.126***	0.27	0.181***
CHS[c]	0.10	0.070***	0.09	0.148***	0.15	0.095***	0.17	0.142***
ESB	0.11	0.000	0.11	-0.010	0.11	0.000	0.10	0.000
NESB	0.15	-0.139***	0.15	-0.178***	0.15	-0.127***	0.15	-0.195***
C_NODEP	0.32	0.088***	0.32	0.026*	0.32	0.093***	0.32	0.040***
S_DEP	0.03	-0.202***	0.09	-0.328***	0.03	-0.168***	0.09	-0.232***
C_DEP	0.45	0.021*	0.43	-0.213***	0.40	0.056***	0.39	-0.147***
DISABIL[d]	0.16	-0.306***	0.14	-0.199***	0.19	-0.278***	0.17	-0.220***
Number of observations		13 995		14 167		12 087		12 472
Pseudo R-squared		0.19		0.15		0.19		0.14

a dF/dx = marginal effect measuring the change in the probability of employment if the explanatory dummy variables change from 0 to 1. **b** *** = statistically significant at the 1 per cent level, ** = statistically significant at the 5 per cent level and * = statistically significant at the 10 per cent level, where significance relates to underlying coefficient estimate **c** This group was identified in 1998 as those people who completed Year 12. This information is not available for 1993 and has been proxied by the age at which people left school and current course type. **d** This is the definition of disability which matches 1993 and 1998 (table F.3), instead of the 1998 SDAC definition used by Wilkins (2003, p. 40).