Estimation of Labour Supply Models for Four Separate Groups in the Australian Population *

Guyonne Kalb
Melbourne Institute of Applied Economic and Social Research
The University of Melbourne

Melbourne Institute Working Paper No. 24/02

ISSN 1328-4991 (Print)

ISSN 1447-5863 (Online)

ISBN 0 7340 3107 6

November 2002

* I am grateful to the Department of Family and Community Services for funding this research. The views expressed in this paper are those of the author and do not represent the views of the Minister for Family and Community Services, the Department of Family and Community Services or the Commonwealth Government. I should also like to thank John Creedy, Alan Duncan, Rosanna Scutella and members from the Department of Family and Community Services for their comments and suggestions.

Melbourne Institute of Applied Economic and Social Research
The University of Melbourne
Victoria 3010 Australia
***Telephone* (03) 8344 3701**
***Fax* (03) 8344 5630**
***Email* melb-inst@unimelb.edu.au**
***WWW Address* http://www.melbourneinstitute.com**

Abstract

This paper estimates discrete choice models of labour supply for couples, single men, single women and sole parents in Australia using the Income and Housing Costs Survey of 1994/1995, 1995/1996, 1996/1997 and 1997/1998. These models are estimated to serve as input in a microsimulation model, where they generate the behavioural responses to policy changes. The results are according to expectations, with preferences for work being higher for people with higher education, who are in their thirties. Furthermore, for women the presence of young children decreases the preference for work. Expected labour supply, predicted by using the estimated models, results in values close to the observed averages.

1. Introduction

This paper describes in detail the estimation of labour supply models for four subgroups of the Australian population. The groups are the following: couples with and without children, single men, single women, and sole parents. Each of these groups is relatively homogenous and we specify one labour supply model for each group. The four groups together add up to a sample representing the Australian population over 15 years of age. The estimation of these models allows us to update the parameters of labour supply behaviour in the Melbourne Institute Tax and Transfer Simulator (MITTS). With these parameters, labour supply responses resulting from policy changes in the tax and social security area can be predicted. Therefore, it is important to have reliable and up-to-date parameters available for the relevant population including single men and women for whom few labour supply models have been estimated so far. In addition, the estimation of similar models for the different groups allows us to compare the effect of characteristics on labour supply for the different demographic groups.

Until recently, the parameters in MITTS for single men and women and for couples were not dependent on personal characteristics such as age and education and some of the quadratic terms were set to zero rather than estimated. Except for the simple specification in Creedy et al. (2002) there is currently no labour supply model available for Australian singles which takes the details of the benefit and tax system into account. The model for sole parents already included more personal characteristics than the models for the other groups but, in the specification presented in this paper, more detailed information on age and education has been included, in addition to an indicator variable for gender[1].

The model allows for the presence of fixed costs associated with working and for heterogeneity in preferences for labour supply and income. The emphasis of the basic framework is on the separation of income into different categories and on a correct representation of net income at all levels of gross income, taking taxes and benefit withdrawal rates into account. This results in a highly nonlinear and non-convex budget set. Estimation of a continuous labour supply model for two persons, using this budget constraint, is complicated and computationally intensive, therefore labour supply is discretized for all groups. Following Van Soest (1995), we use a multinomial logit specification in the discrete choice model, which allows us to choose a relatively large number of labour supply points for both adults in the household.

---

1 See Duncan and Harris (2002) for the previous version of the sole parent labour supply model or Creedy et al. (2002) for previous estimates of the labour supply models for all groups.

Section 2 briefly discusses the economic model. Section 3 describes the data. Section 4 contains the econometric details. The results from the models for the different groups are discussed in Section 5. First the estimated parameters are discussed and then predicted labour supply using the estimated parameters is presented. Finally, in Section 6 some conclusions are presented.

2. The Economic Model

2.1. Choice of population subgroups

The groups in which the population is subdivided are couples with and without children, single men, single women, and sole parents. Each of these groups is relatively homogenous, which allows us to specify one labour supply model for each group. The four groups together add up to a sample representing the Australian population over 15 years of age. Compared to groupings according to other criteria (such as education or age), it seems reasonable to assume that there is more difference between a single parent and a single man than there is between someone with a vocational qualification and someone with a degree or between someone aged 20 to 30 and someone aged 30 to 40. However, the choice for these groups because of their relative homogeneity does not imply that within each group there is no longer considerable heterogeneity. The model allows for this heterogeneity by including individual and household characteristics, such as age, education, and age and number of children in the model. In addition the model can allow for unobserved heterogeneity.

Further subdivision of the above groups could make the sample size of individual subgroups too small to consider separately in a model. For modelling reasons, single person households and couples need to be in separate groups, since the model for couples includes several parameters that are not relevant for a single person. The further subdivision into men and women is one that is commonly followed in the literature, given the observed differences in labour supply behaviour and in wage levels. Finally, a large part of the applied labour economic literature focuses on sole parents, as a particular group of interest. In comparison to other groups they are often found to be more responsive to financial incentives than other groups (Eissa and Hoynes, 1999; Blundell and Hoynes, 2000). Therefore, it seems sensible to estimate a separate model for this group as well. Few researchers have aimed to estimate labour supply for the complete population. Most articles dealing with labour supply thus focus on a subgroup and in the choice of subgroup similar groups to the four groups distinguished here are often selected. For example, couple households in Van Soest (1995) or Hoynes (1996), and sole parent households in Bingley and Walker (1997) or Duncan, Giles and MacCrae (1999). In contrast, single men and women have received relatively little attention in labour supply modelling[2].

2.2. Utility Maximization

In the model chosen in this paper, the household is assumed to be the decision-making unit on labour supply and consumption. Thus, we use a household utility function or a unitary utility function, which does not explicitly take into account individual consumption or utility, but assumes there is one common utility function for the whole household. Although alternative models are available, which incorporate more realistic assumptions on utility maximization in the household or allow for home production to enter the model independently[3], these models would introduce additional complications. To estimate a model where each household member has their own utility function, information is needed on the private consumption of individuals or on the amount of income allocated to them. No data set combines all necessary information on consumption or home production, income sources, and labour supply. Therefore strong assumptions are often needed on how income is shared to allow estimation of collective utility models or on the value and amount of home produced goods to estimate models that explicitly allow for home production, instead of implicitly as in the unitary utility models. To deal with these additional complications other parts of the model need to be simplified and as a result keeping all the current detail of the tax and transfer system would be very difficult.

Given the aim of MITTS of simulating policy changes with regard to the tax and transfer system and assessing its effect on labour supply, priority is given to incorporating all possible detail on taxes and transfers. The literature that studies the effect of policy changes in taxation or social security systems mostly favours the neoclassical approach for its suitability to incorporate detailed budget constraints.

By setting up the model in the familiar neoclassical way, starting from utility maximization under a budget constraint, a logical and consistent framework can be built to analyse labour supply (see for example Deaton and Muellbauer, 1980; or Killingsworth, 1983). For example, take a two-adult household (with or without dependent children), where the adults choose their labour supply to optimise its utility. Their utility depends on household consumption

2 Euwals and Van Soest (1999) estimate one labour supply model for unmarried men and women with some separate parameters for men and women. However, their model also includes sole parents.

3 See for example, Bourguignon and Chiappori (1994), Browning et al. (1994), and Apps and Rees (1996, 1997, 2000).

(which is assumed to be equal to net household income x[4]), on the amount of leisure time[5] of adult 1, and the amount of leisure time of adult 2. Depending on the utility function chosen, this approach allows for direct interdependencies between the two adults' labour supply or one adult's labour supply and household income. This utility is maximized conditional on the restricted total amount of time available to each adult and the restricted amount of total household income. It is expected that utility increases with an increase in leisure and income. Usually more income means less leisure time for one of the adults, except when more income is obtained through social security benefits[6]. In short, maximizing a household's utility involves balancing the amount of leisure and income.

A simple utility maximizing model would look as follows:

$$\max U(x,l_1,l_2) \tag{1}$$

subject to:

$$T = l_1 + h_1$$

$$T = l_2 + h_2$$

$$x = \int_0^{h_1} g_1(t_1,h_2)dt_1 + \int_0^{h_2} g_2(h_1,t_2)dt_2 + n(y_1) + n(y_2) + n(B(c))$$

where:

U( ) is the utility function of a two-adult household,

l_1 and l_2 indicate the aggregate of leisure time and home production time per week of the husband and wife (married or de facto) respectively,

x indicates net income per week,

T is the total available time for each person in the household,

h_1 and h_2 are the hours of work of husband and wife,

$g_1(,)$ and $g_2(,)$ are the marginal net wages of husband and wife at the different hours of work h_1 and h_2 taking into account taxation and withdrawal of benefits,

y_1 and y_2 are the non-labour incomes of husband and wife,

c is household composition,

---

4 There is no provision in the model for intertemporal transfers of money. However, the payout of dividends on investments or the payout of interest on savings in the current period are included in the "other income" variable.

5 This leisure time is not pure leisure time but it also includes home production time.

6 In the current specification of the model it is assumed that everyone who is eligible for benefits takes them up.

B(c) is the amount of benefit a household is eligible for, given their household composition c,

n( ) is the amount of income after the deduction of taxes.

The first two restrictions are time restrictions for the two adults. The third restriction, the budget constraint, denotes the level of available income in the household. If the three restrictions are taken together, the budget constraint may be written:

$$\begin{aligned} & x + \int_{T-l_1}^{T} g_1(t_1, T-l_2)dt_1 + \int_{T-l_2}^{T} g_2(T-l_1, t_2)dt_2 = \\ & \int_{0}^{T} g_1(t_1, T)dt_1 + \int_{0}^{T} g_2(T, t_2)dt_2 + n(y_1) + n(y_2) + n(B(c)) \end{aligned} \quad (2)$$

In this paper, the term 'leisure' is used to indicate both pure leisure time and home production time. The combination of leisure and income that delivers the highest utility to the household is regarded as the optimal choice.

The choice of labour supply is simultaneously determined for both adult members of the household. Depending on the choice of utility function, different interactions between household income and labour supply of both adults can be modelled.

For households with only one adult, the model can be simplified by leaving out everything relating to the second adult:

$\max U(x,l_1)$(3)

subject to:

$T= l_1 + h_1$

$$x = \int_{0}^{h_1} g_1(t_1)dt_1 + n(y_1) + n(B(c))$$

Or combining the two restrictions:

$$x + \int_{T-l_1}^{T} g_1(t_1)dt_1 = \int_{0}^{T} g_1(t_1)dt_1 + n(y_1) + n(B(c)) \quad (4)$$

With regard to the assumption of free choice underlying this economic model; in practice, it is often not known whether the observed labour supply is the optimal labour supply or, alternatively, whether people are restricted in their labour supply choice by demand side factors[7]. It would be interesting to analyse desired hours of work instead of actual hours of work or to allow for the restrictions in actual hours caused by the demand for labour (see for

7 See for example, Laisney et al. (1992), Bingley and Walker (1997) or Duncan, Giles and MacCrae (1999).

example Euwals and Van Soest (1999) or Euwals (2001)). However, if a person works, it is assumed that the preferred hours equal their actual working hours, because no information on the preferences of working respondents is available. For the moment it is also assumed that all non-participants are voluntarily not working.

2.3. Unobserved Wages

Like other researchers in this area, we have to deal with unobserved market wages for people who are not working. In this paper, we use the popular approach of estimating the wage equation separately and using estimated wages as if they represented the true values of the unobserved wages[8]. To correct for a possible selection bias as a result of only observing wage rates for those gainfully employed the Heckman correction term for participation is included in the wage equation (Heckman, 1979). In future research, the possibility of incorporating unobserved wages within the likelihood function and estimate wages and labour supply simultaneously will be explored. However, this is computationally more demanding and it is not attempted very often[9].

Separate wage equations have been estimated for the five demographic groups. The specification of the wage equation is discussed in a separate paper (Kalb and Scutella, 2002). For each non-participant we impute an expected value for the wage rate in the labour supply model.

3. The Data

The Survey of Income and Housing Costs 1994-95, 1995-96, 1996-97 and 1997-98, all released by the Australian Bureau of Statistics (ABS), have been used for the analyses. They contain detailed income information for each person separately and for the household as a whole. This allows the budget constraint to keep its full complexity. In order to combine the four years into input for one model, the monetary variables from 1994/1995 to 1996/1997 are converted to the March 1998 level[10]. Furthermore, the observed nominal wages in these survey years are adjusted by the average wage increases for men or women as relevant.

---

8 Van Soest (1995) uses this approach and points out that most of the papers in a special issue on Taxation and Labor Supply in Industrial Countries of the Journal of Human Resources (Moffitt, 1990) follow this approach as well.

9 Exceptions are for example Fraker and Moffitt (1988), Gerfin (1993) and Murray (1996).

10 For this the Consumer Price Index as published by the Australian Bureau of Statistics (1998) is used.

3.1. Selection Criteria for the Four Groups

In this section, the selection criteria for each of the groups are discussed.

The criteria for the *first* group, work-age couples, are the following:

- Only income units that consist of a head and a partner with or without dependants.
- Self-employed are excluded from the analyses. The surveys used for the analyses do not report the number of hours worked for people in self-employment. In addition, for self-employed the relationship between total earned income and labour supply is not as simple as for many wage and salary earners, where total earned income equals labour supply multiplied by the wage rate.
- People of an age to be eligible for government paid age pensions are excluded. They are expected to behave differently from younger people.
- All people temporarily or permanently unable to work because of illness or disability are excluded from the analysis.
- People receiving a (military) service pension are not included, since these pensions are paid instead of age pension or in cases of disability.
- All full-time students are excluded.

The criteria for the *second* group, work-age single men, are the same as above with the first criterion replaced by income units that consist of one adult man without dependants. The criteria for the *third* group, work-age single women, are also the same as above but with the first criterion now replaced by income units that consist of one adult woman without dependants. Finally, the criteria for the *fourth* group, work-age sole parents, are the same as above with the first criterion replaced by income units that consist of one adult man or woman with dependants.

Missing values or outliers (which may be measurement errors) result in the deletion of a few additional households. First, some observed values for wage income seem unrealistically small when compared to the corresponding hours worked. In Australia there is no Federal or state minimum wage covering all employees. Each award has its own minimum wage. Therefore, across states, occupations and industries, minimum wage levels vary. In addition, some workers, such as trainees, apprentices and supported workers, are not covered by an award and some employees may work in unpaid overtime. This makes it difficult to decide on a wage level, which distinguishes realistic from unrealistic wage levels. In the estimation of the labour equation in this paper, all persons earning less than $4 per hour[11] or more than

11 4 sole parents, 41 single men, 33 single women, 68 married men and 56 married women (where for 13 couples both partners are on an extremely low wage) fall into this group.

$100 per hour are excluded[12] as such low and high values seem likely to be due to measurement error (the same selection is used to estimate the wage equation in Kalb and Scutella (2002))[13]. Second, all households who have zero net income at zero hours of work are excluded[14]. After these selection processes, a data set of 10250 income units is left for the labour supply analysis in group 1; 5671 income units in group 2; 4596 income units in group 3; and 1822 income units remain in group 4.

3.2. Variables used in the Analyses

Figures 1 and 2 give an overview of the sample frequency distribution of (categorized) male and female working hours in the samples for the different groups. The difference between men and women is obvious and as expected. Relatively more women work part time and more men work full time (especially over 45 hours per week) in the different subsamples. There is also a clear difference between singles and couples. Single men are more likely to be nonparticipants or work part time than men in couples. They are also less likely to work more than 42.5 hours and in particular to work more than 47.5 hours per week.

Figure 1 Labour supply of men

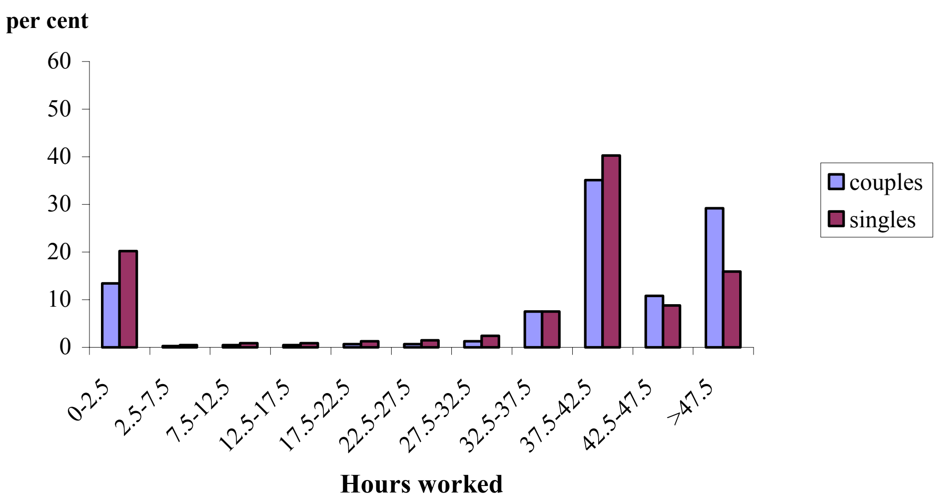

Figure 2 shows that single and married women behave differently as well. Single women work more hours and are less likely to work part time or be out of the labour force. The sole parents in this figure also contain the sole fathers, which is a rather small group. Sole parents

---

12 1 sole parent, 4 single men, 3 single women, 34 married men and 17 married women fall into this group.

13 None of the imputed wage rates fall into this category of wages that seem too low or too high.

14 2 sole parents, 62 single men, 55 single women and 0 couples fall into this group.

are by far the least likely to participate in the labour force and if they participate they are more likely than the other groups to work in the lowest hours categories.

Figure 2 Labour supply of women and sole parents

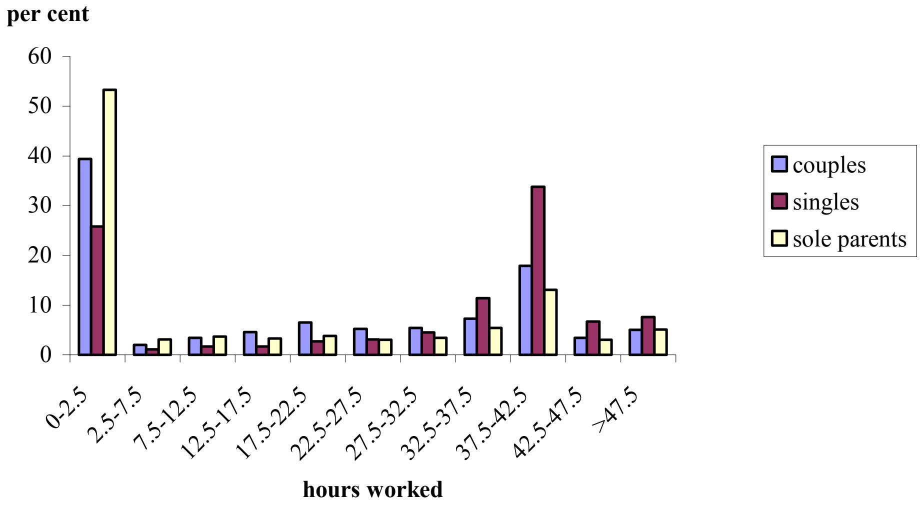

Table 1 gives summary statistics of the variables, which are used in the analyses. The background characteristics used to specify preferences in the utility function are listed below.

Age is known exactly for those under 25 and those over 54 years of age, while the ages between 25 and 54 are known in five-year intervals. The midpoint values of each category are used in the analyses and to calculate the average age. Younger and older persons are expected to have a higher preference for leisure. Many studies include age and age squared to allow for a non-linear relation between age and the preference for leisure. Van Soest (1995), Aaberge, Colombino and Strøm (1999), Duncan and MacCrae (1999), Euwals and Van Soest (1999), and Van Soest, Das and Gong (2002) either find that age reduces the preference for leisure or they find a reduction of the preference for leisure at first, followed by an increase in the preference for leisure after a certain age.

Education is divided into the following categories:

- no qualifications
- vocational qualifications
- associate or undergraduate diploma
- higher or bachelor degree or postgraduate diploma

Education is expected to increase the preference for work, because time and money have been invested in human capital. Apart from the financial rewards, one would also expect a high-skill job to be more interesting than a low-skill job and hence more desirable. In accordance with the above expectation, Duncan and Harris (2002) find that having some qualifications increases the preference for labour supply and Duncan and MacCrae (1999) find that leaving school at 16 years of age decreases the preference for labour supply. Murray (1996) similarly finds that sole parents with some form of post-secondary school qualifications have a higher preference for work.

Table 1: Summary Statistics for the combined Income and Housing Cost Survey 1994/1995, 1995/1996, 1996/1997 and 1997/1998

	Couples N=10250		Single men N=5671		Single women N=4596		Sole parents N=1822	
Continuous Variables	**mean**	**st.dev**	**mean**	**st.dev**	**mean**	**st.dev**	**mean**	**st.dev**
Average hours worked by head	36.807	15.939	31.958	17.533	27.318	18.015	14.469	17.975
Average hours worked by spouse	19.043	17.984						
Age head	42.363	10.451	31.010	11.811	34.648	14.822	36.743	8.946
Age spouse	39.936	10.211						
Number of children in income unit	1.206	1.208					1.714	0.879
Percentage of households without children	0.397	0.489						
Wage rate head	18.480	8.953	14.662	6.757	13.584	5.575	12.428	6.250
Wage rate spouse	13.991	6.549						
Dummy Variables								
Living in New South Wales	0.229	0.420	0.232	0.422	0.234	0.424	0.203	0.402
Residence of income unit in capital city	0.616	0.486	0.634	0.482	0.674	0.469	0.603	0.489
Gender(woman)							0.889	0.315
Education of head								
• No qualifications	0.429	0.495	0.558	0.497	0.569	0.495	0.648	0.478
• Vocational qualification	0.287	0.452	0.236	0.425	0.173	0.378	0.184	0.388
• Diploma	0.114	0.318	0.080	0.271	0.091	0.288	0.074	0.262
• University degree	0.171	0.376	0.126	0.332	0.167	0.373	0.094	0.292
Education of spouse								
• No qualifications	0.597	0.490						
• Vocational qualification	0.174	0.379						
• Diploma	0.091	0.287						
• University degree	0.138	0.345						
Youngest child in income unit is								
• between 0 and 2	0.157	0.364					0.203	0.402
• between 3 and 4	0.067	0.250					0.133	0.339
• between 5 and 9	0.108	0.311					0.211	0.408
• between 10 and 15	0.076	0.265					0.153	0.360
Employment status head								
Non participation	0.072	0.258	0.064	0.245	0.160	0.367	0.440	0.497
Unemployed	0.061	0.239	0.136	0.343	0.096	0.295	0.086	0.281
Employed	0.867	0.339	0.800	0.400	0.744	0.437	0.474	0.499
Employment status spouse								
Non participation	0.353	0.478						
Unemployed	0.038	0.190						
Employed	0.609	0.488						

The *number of dependent children* in each income unit is calculated by adding the number of dependent children from 0 to 24 years old. This variable is expected to be especially important for the female adult in the income units. Children are likely to increase the value of time at home, which is reflected in a higher preference for leisure in the model.

The survey records the *age of the youngest dependent child under 15 years of age* in the income unit. The effect of dependent children in the income unit on the preference for time spent in working is likely to be bigger when young children are present.

The expected effects with regard to children are found in several studies. The effects are particularly strong for women. Van Soest (1995) finds effects for both men and women, where the female effects are somewhat larger. Van Soest, Das and Gong (2002), Aaberge, Colombino and Strøm (1999), Fraker and Moffitt (1988), Hagstrom (1996) and Hoynes (1996) find effects for married women. Duncan and MacCrae (1999) find strong effects for sole parents (mostly women) and married women of both the age of the youngest child and the number of preschool children. Much lower (and often no) effects are found for men. Similar effects are found for sole parents in Australia (Murray, 1996).

Residence of income unit in capital city and *Living in New South Wales* are location variables for where the income unit lives in Australia. It is expected that the fixed cost of working is different for people in or outside the capital cities and in or out of New South Wales, in particular for people with children who may need childcare services (Duncan and Harris, 2002).

Finally, *men* and *women* are expected to have different preferences for "leisure" time. In the models for two-adult income units, person 1 is male and person 2 is female. None of the two-adult income units contain two adult men or two adult women. In the single-adult income units, models are estimated separately by gender. For sole parents, the male group is too small to estimate separate models, so therefore a dummy variable for gender is included in the preference for labour supply and income, and in the fixed cost parameter to explore whether gender affects the preferences in this group.

4. Econometric Specification of a Labour Supply Model

In Section 2 an economic model was introduced that serves as a starting point for the specification of an econometric model. In the following sections, the econometric specification is discussed.

4.1. Allowing for a Nonlinear and Non-convex Budget Set

Including taxes and benefits for two persons in the budget constraint produces a highly nonlinear constraint. Looking at the benefit and tax regimes of 1994/1995, 1995/1996, 1996/1997 and 1997/1998[15] leads us to expect many kinks in the budget constraint. Since we prefer to keep the representation of taxes and benefits as close to reality as possible, a complex budget constraint cannot be avoided. In the case where one only considers one potential worker at a time, the labour supply estimation can already be quite complex[16]. The complexity is even greater in the case where income units with two potential workers are analysed, subject to their joint budget constraint.

Restricting the number of possible working hours to a limited set of discrete values, as is done by other authors (for example Van Soest, 1995; Duncan, Giles and MacCrae, 1999; Keane and Moffitt, 1998) facing the same problem, appears an attractive solution. For this limited set of hours, one can calculate the level of utility that each possible combination of hours would generate, according to the specified utility function. An additional (computational) advantage of the discrete approach is that quasi-concavity does not have to be imposed before using maximum likelihood methods to estimate the model, as is necessary in the case of continuous labour supply for some utility functions (see Van Soest, Kapteyn and Kooreman, 1993), but can be checked after estimation.

Instead of being defined on a continuous set of working hours [0,T], in the discrete choice case the budget constraint is defined on a discrete set of points $h_1 \in \mathcal{A} = \{0, h_{11}, h_{12}, ..., h_{1m}\}$ and $h_2 \in \mathcal{B} = \{0, h_{21}, h_{22}, ..., h_{2k}\}$ on the interval [0,T][17]. Using these sets, the net income $x(h_1, h_2)$ is calculated for all (m+1)×(k+1) combinations of h_1 and h_2 (where m+1 is the number of discrete points for h_1 and k+1 is the number of discrete points for h_2). The Melbourne Institute Tax and Transfer Simulator can calculate net income at all chosen discrete labour supply points. By increasing the number of different hours in the choice set, the quality of the representation improves. However, the computational load also increases, so a compromise between quality and computational feasibility is necessary. Furthermore, some of the theoretically possible hours ranges may not be observed in the data such as low part-time hours for men, which may mean fewer discrete points are necessary in that range.

Net income x is dependent on labour supply and wage rates of both adults, on non-labour income, on household composition and on eligibility for benefits. Net income for the records originating from the 1994/1995, 1995/1996, and 1996/1997 data sets are inflated up to the

---

15 The Melbourne Institute Tax and Transfer Simulator (MITTS) contains all the necessary information to calculate net income from gross income for these years.

16 See e.g. Burtless and Hausman (1978), Hausman (1979), Hausman (1985) or Moffitt (1986) for a continuous labour supply approach with a nonlinear (non-convex) budget constraint.

17 0, h11, h12, etc represent the discrete values that labour supply can take.

1997/1998 level by multiplying the amount by the relevant CPI. In this way, net incomes in the different years are comparable. Wage rates, non-labour income and household composition are exogenous in this model. The model becomes:

$$\max U(x,l_1,l_2) \tag{5}$$

subject to:

$$\begin{aligned} & l_1 + h_1 = T \\ & l_2 + h_2 = T \\ & \text{where } (h_1, h_2) \in \mathcal{A} \times \mathcal{B} \end{aligned} \tag{6}$$

$$x = w_1h_1 + w_2h_2 + y_1 + y_2 + B(c, w_1h_1 + w_2h_2 + y_1 + y_2) - \tau(B, w_1h_1 + y_1, w_2h_2 + y_2, c)\}$$

w_1 and w_2 are the gross wage rates of husband and wife respectively,

$\mathcal{A}$ and $\mathcal{B}$ are the sets of discrete points from which values can be chosen for h_1 and h_2,

B is the amount of benefit, for which the household is eligible, given household composition c and income,

τ is the tax function that indicates the amount of tax to be paid.

A likelihood function can be formed using the above utility function. Based on the assumption of utility maximization for each household the following can be stated. The contribution of each household to the likelihood function is the probability that its observed hours result in an optimal utility for the household of interest when compared with all other possible choices for hours. This probability looks as follows:

$$\Pr(U(x((h_1,h_2)_r),(h_1,h_2)_r,\varepsilon_r) \geq U(x((h_1,h_2)_s),(h_1,h_2)_s,\varepsilon_s) \text{ for all } s) \tag{7}$$

where:

r stands for the combination h_1 and h_2 that is preferred,

s stands for all (k+1)×(m+1) possible combinations that can be made, given the discrete choice sets for hours worked,

ε_r and ε_s represent error terms.

Adding an error term to the utility function prevents contributions to the likelihood in any data point from becoming zero. It allows for optimization errors made by the household. Choosing an extreme value specification for the error term in (7) results in a multinomial logit model (see Maddala, 1983). If we can calculate utility levels for each of the possible combinations of leisure and income, and the error terms are specified, then for each possible combination we can calculate the probability of that combination being preferred according to the estimated model.

The log likelihood contribution for couples looks as follows:

$$\ln L = U_{i'j'} - \ln\left(\sum_{i,j} \exp(U_{ij})\right) \tag{8}$$

where:

i indicates the husband's labour supply;

j indicates the wife's labour supply;

i', j' are the preferred (observed) states of labour supply (combination r in equation 7);

U_{ij} is the level of utility derived from the state where the husband has labour supply i and the wife has labour supply j.

Expression (8) denotes the probability that the utility in the observed combination of hours is higher than the utility in any other situation. The aim is to choose parameter values for the utility function that maximize the log likelihood function in the observed data points.

For single adult households equation (8) simplifies to:

$$\ln L = U_{i'} - \ln\left(\sum_{i} \exp(U_i)\right) \tag{9}$$

4.2. Specification of the Utility Function

For the sake of convenience the utility function used here is a quadratic specification (following Keane and Moffitt, 1998). The quadratic specification is simple but quite flexible in that it allows for the leisure of each person and income (or consumption) to be substitutes or complements. This means the model can represent complex interactions. Furthermore, the quadratic utility function can be expressed as a function of labour supply rather than leisure without the need to choose a value for total endowment of time (T). T is not important in this specification, as it is a constant, which can be incorporated in the parameters to be estimated.

The above advantages make the quadratic utility function a good choice, even though this utility function is not automatically quasi-concave. However, the latter is not a problem in a discrete labour supply model, because if the two conditions outlined in Van Soest (1995) are fulfilled at a data point, then U is quasi-concave at that point. In the discrete approach taken here, these two conditions can be tested at all data points after estimation of the parameters. In a model with continuous hours of labour supply, these conditions would have had to be imposed a priori to guarantee coherency, as has been mentioned earlier.

Many models have had the problem of overpredicting part-time hours and underpredicting non-participation. An intuitively appealing approach is to include a fixed cost of working parameter in the income variable x to indicate the cost of working versus nonparticipation (Callan and Van Soest, 1996). As a result of the inclusion in x, this cost of working parameter is measured in dollars per week. The utility derived from leisure and income can be written as:

$$U(x,h_1,h_2)=\beta_x(x-\gamma_1-\gamma_2)+\beta_1 h_1+\beta_2 h_2+\alpha_{xx}(x-\gamma_1-\gamma_2)^2+\alpha_{11}(h_1)^2+\alpha_{22}(h_2)^2+ \alpha_{x1}(x-\gamma_1-\gamma_2)h_1+\alpha_{x2}(x-\gamma_1-\gamma_2)h_2+\alpha_{12}h_1 h_2 \quad (10)$$

where $\alpha_{..}$, $\beta_{.}$, and φ are preference parameters that have to be estimated; and γ_1 and γ_2 are the fixed cost of working parameters to be estimated for husband and wife, they are zero when the relevant person is not working.

This quadratic utility function has a simple form and heterogeneity of preferences is easy to include. To account for differences in preferences between households, the parameters β, α, and γ can be made dependent on household and individual characteristics. In the first instance, it is assumed that only β_1, β_2, β_x, γ_1 and γ_2 depend on personal and household characteristics (see section 3.2 for a description of the characteristics to be included). Simple linear specifications are chosen to include the observed heterogeneity in β_1, β_2, β_x, γ_1 and γ_2.

Adding unobserved heterogeneity to these parameters, in the form of a normally distributed error term with zero mean and unknown variance, is quite simple, although exact maximisation would involve a likelihood function with multiple integrals. However, Van Soest (1995) outlines an easier method, replacing the expectation of the log likelihood by a simulated mean and optimising an approximate likelihood function instead of the exact likelihood function. It is straightforward to obtain a simulated mean by: drawing error terms from the distribution based on the current parameter estimates for the covariance matrix for each observation in the sample; calculating the log likelihood function based on these draws; and averaging the log likelihood function over a certain number of draws. Van Soest found that 10 draws seemed sufficient, so the estimation of unobserved heterogeneity in this paper is carried out with the same number of draws.

4.3. Expected Labour Supply

Once the complete model has been estimated, the results can be used to calculate the expected labour supply from the probabilistic outcomes for people with certain known characteristics and under known social security and taxation rules.

To obtain the expected labour supply of the husband, we first calculate the utility $U(x(h_1,h_2), h_1, h_2)$ for each possible combination of labour supply for both adults in the household. This is achieved by substituting the estimated parameter values into equation (10) after calculating the net income for the relevant combination. Once the utility values are known, a simple logit transformation provides the probability of each possible combination occurring according to the estimated model:

$$p(h_1,h_2) = \frac{\exp\left(U\left(x\left(h_1,h_2\right),h_1,h_2\right)\right)}{\sum\limits_{\substack{\text{over all}\\ h_1,h_2}} \exp\left(U\left(x\left(h_1,h_2\right),h_1,h_2\right)\right)} \quad (11)$$

These probabilities can then be used to calculate the expected value of preferred labour supply for the husband by simply aggregating the probabilities over all possible values of h_2 for each value of h_1. In this manner, the marginal probability of h_1 is obtained, which can then be used to calculate the expected value of h_1 in the usual way. The formula for this procedure looks as follows:

$$E(h_1) = \sum_{h_1}\left[\left(\sum_{h_2} p(h_1,h_2)\right)h_1\right] \quad (12)$$

The expected value for the wife's labour supply can be obtained in a similar way.

5. Results

Labour supply is estimated using imputed wage values for the non-workers as described in Section 2.2. The next subsection presents the results of the labour supply models for couples. In the second subsection, the estimated results are used to predict labour supply probabilities so that predicted and actual results can be compared.

5.1. Discussion of the Labour Supply Results

To show how the results of a model as discussed in section 4 are interpreted, we discuss the parameters of two-adult income units[18]. Table 2 gives the parameter estimates of the quadratic specification of the utility function for a model with six discrete labour supply points for men and eleven points for women. The location of the points is defined in a footnote to the table.

---

18 The results for single men, single women and sole parents are presented in Table A.1 in appendix A and are discussed more briefly, comparing the results with the results for couples.

The linear terms

The effects of different characteristics on the preference for leisure of both adults in the household are the first results to be discussed. We only discuss those parameters that are significant at the 5-percent level.

To begin with the parameterised preference for work for the male adult, a significant positive effect[19] is found for the linear term of age. This means that older men have a higher preference for work and thus a lower preference for leisure. However, on the other hand the quadratic term for age seems to have a significant negative effect on the preference for work, which combined with the linear effect of age means that the preference for work increases for men up to 37 years of age after which it decreases with age. Thus young men and older men have a lower preference for labour supply. A positive effect is further observed for households where the man has a higher level of education. The partner's education does not seem relevant. None of the variables related to the number and age of dependent children in the household influence the preference for work.

Table 2: Estimated Parameters of the Labour Supply Model [a]

	Estimated coefficient	**z-value [b]**
Quadratic terms		
income$\times$ *100,000*	-0.0213	-4.56
Labour supply husband $\times$ *100*	-0.6370	-63.15
Labour supply wife $\times$ *100*	-0.1912	-20.93
Crossproduct		
Inc. & lab. sup. husband $\times$ *10,000*	-0.3935	-14.54
Inc. & lab. sup. wife $\times$ *10,000*	-0.1717	-9.76
lab. sup. husband & wife $\times$ *100*	-0.0466	-8.51
Linear terms		
Income $\times$ 100		
constant	0.6991	35.62
Number of children	-0.0095	-3.72
Labour supply husband		
constant	0.3622	28.63
Children 0-2 yrs old	0.0001	0.03
Children 3-4 yrs old	-0.0025	-0.81
Children 5-9 yrs old	0.0004	0.16
Number of children	0.0009	0.95
Age/10	0.0633	12.38
Age squared/100	-0.0086	-14.44
Vocational education	0.0086	5.29
diploma	0.0065	2.61
degree	0.0114	4.54

19 This indicates a higher preference for work and thus a smaller taste for leisure.

Table 2: Continued

	Estimated coefficient	z-value [b]
Voc. education (partner)	0.0024	1.27
diploma (partner)	0.0005	0.18
degree (partner)	0.0023	0.87
Labour supply wife		
constant	-0.0042	-0.35
Children 0-2 yrs old	-0.0293	-13.32
Children 3-4 yrs old	-0.0198	-6.72
Children 5-9 yrs old	-0.0012	-0.54
Number of children	-0.0147	-19.24
Age/10	0.0652	11.61
Age squared/100	-0.0095	-13.44
Voc. education (partner)	-0.0001	-0.05
diploma (partner)	-0.0032	-1.44
degree (partner)	-0.0070	-3.25
Vocational education	0.0090	5.00
diploma	0.0169	7.29
degree	0.0287	13.22
Fixed cost husband/100	16.0933	30.24
Fixed cost wife/100	6.0702	25.97

a Six discrete points of labour supply are distinguished for each man: 0 hours for non-participants and people working less than 2.5 hours, 10 hours for people working from 2.5 to 15 hours, 20 hours for people working from 15 to 25 hours, 30 hours for people working from 25 to 35 hours, 40 hours for people working from 35 to 45 hours, and 50 hours for people working more than 45 hours. Eleven discrete points of labour supply are distinguished for each woman: 0 hours for non-participants and people working less than 2.5 hours, 5 hours for people working from 2.5 to 7.5 hours, 10 hours for people working from 7.5 to 12.5 hours, 15 hours for people working from 12.5 to 17.5 hours, 20 hours for people working from 17.5 to 22.5 hours, 25 hours for people working from 22.5 to 27.5 hours, 30 hours for people working from 27.5 to 32.5 hours, 35 hours for people working from 32.5 to 37.5 hours, 40 hours for people working from 37.5 to 42.5 hours, 45 hours for people working from 42.5 to 47.5 hours, and 50 hours for people working more than 47.5 hours.

b The z-value indicates the level of significance of the estimated coefficients. A value of 1.96 or more means that the parameter is significant at the 5% level at least. That is we are quite confident it is not equal to zero. The higher the z-value the more precise the estimated coefficient.

According to expectation, the preference for work of the female adult seems to be lower than that of her male partner, at least as far as this is reflected in the size of the constant term of β_2. A significant positive effect is observed for women with higher education levels. The effect of education seems more important for women than for men. This could be caused by the fact that almost all men are working or looking for work, whereas women's labour supply is more variable. Additionally, if the partner's education level is higher, then a woman's preference for work is reduced to some extent. However, the effects are smaller than those resulting from her own education and only when the partner has a degree or higher is the effect significant. From the linear and quadratic age parameters it can be derived that the maximum preference for work is around 34 years of age.

All variables related to children have a significant negative effect on a married woman's preference for work. The effect for children between five and nine years old is much smaller than for younger children and insignificant. As expected, and as is seen in many other studies

(Australian examples are Eyland, Mason and Lapsley, 1982; Ross, 1986 and Murray, 1996), having a newborn child or a child between three and four years of age has a large negative effect on the female preference for work. Children of primary school age, however, do not seem to affect the mother's preference for work any longer. Finally, women with more children have a lower preference for work.

To keep the model manageable the preference for income only depends on the number of children. Other characteristics were not significant and did not improve the model. One would expect a higher preference for income when the household size increases, however the reverse appears to be true. This could be a spurious relationship reflecting the often-observed correlation between low income and the number of children, which may be driven by similar household and personal characteristics.

Comparing the effects found for couples with those for singles and sole parents, similar variables are found to be important (see Table A.1). High education levels increase the preference for work for all groups. The effect on the preference for income is less clear; higher education levels increase the preference for some whilst decreasing it for others.

The effect of age on the preference for work for single men is similar to that for married men. The preference is at a maximum around 38 years of age. Furthermore, single women's maximum preference for work occurs around 36 years of age, which is close to the age at which this occurs for married women. For sole parents, this maximum lies around 41 years of age perhaps partly reflecting their children's age.

Comparing the effect of children for sole parents and married women it is obvious that the age of the youngest child is important for both groups. However, it is remarkable that sole parents seem to have the lowest preference for work when their youngest child is in the age group five to nine. On the other hand, their preference for income is also at its highest when they have children in this age group. Thus, the two effects work in opposite directions.

Finally, the model for sole parents contains one additional explanatory variable, gender, because this group consists of both men and women. The coefficient shows that sole mothers have a lower preference for work than sole fathers.

Overall, the characteristics included in the labour supply model had the expected effects on the preferences for labour supply. That is, the preference for employment goes up with age at first and declines again after an age of about 35 to 40 is reached. Well-educated individuals have higher labour supply preferences. Children decrease female preferences for labour supply, in particular when it concerns preschool children. The effect on male preferences is not significant, similar to what is found in other studies (Section 3 briefly discusses the results from other research).

Quadratic and crossproduct terms

Besides the linear terms, there are also quadratic terms involved in the quadratic utility function. Taking the first derivative with respect to labour supply of men, the following expression for the marginal utility of labour supply for men is obtained:

$$U_1 = \beta_1 + 2\alpha_{11}h_1 + \alpha_{12}h_2 + \alpha_{x1}(x - \gamma_1 - \gamma_2)$$

Similar expressions can be formulated for labour supply of women and for net income. From this formula and the results in Table 2, we conclude that couples seem to see each other's labour supply as substitutes. If one of the two persons works more, the marginal utility of work of the other person decreases (since α_{12}=-0.0466). This is contradictory to other Australian studies where it is found that if one in a couple had more leisure time then the other's marginal utility for leisure time increased (Kalb, 1999, 2000). The parameter was also positive for the data used here before the inclusion of the fixed costs of working parameter. The fixed cost of working parameter may need some re-fining in future studies, which could change the cross-product parameter again. This is discussed after the next paragraph.

The model presented here does not directly provide information on the effect of characteristics on labour supply like in a simple regression model. Instead, it provides the effect of characteristics on preferences for leisure of each person and on the preference for income. These preferences affect labour supply indirectly through the level of utility that can be obtained at each labour supply point. Therefore a positive cross product term for labour supply of the two adult household members indicates a preference for joint leisure time, but the labour supply outcomes are only partly driven by this cross product term. Other factors in the model (such as household income and wage levels) influence the final decision on labour supply as well. Thus, from a negative cross product term for the preference for labour supply of both members of the household, it does not automatically follow that if the husband increases labour supply that the spouse will then reduce her labour supply. Although the negative cross product term will make this more likely to happen, the effect from this factor is likely to be small compared to the effects of other factors. Thus although the negative coefficient is different from other studies, it is not a problem in the sense that the negative correlation would neutralize the positive influence of changes to incentives for couples.

There is also a significant effect of income on the marginal utility of labour supply or vice versa at the 5-percent level for both the husband and wife. Both effects are negative indicating that the marginal utility of labour supply goes down when income goes up and that the marginal utility of income goes down when the amount of labour supply goes up. A significant negative effect is also estimated for the models of the other groups.

Fixed cost of working

The fixed-cost-of-working parameters seem very large in the model for couples, particularly for men. The fixed costs parameters are not estimates of actual costs of working, because they also include non-pecuniary costs and they probably also pick up the lack of people working part-time. The latter may make it look like people do not want to work for an income under the full-time rate. For prediction purposes this is not a problem, but the model will need some further work to better understand the reasons for the large fixed-cost parameters. For example, the lack of people working part time may be a labour demand issue rather than a labour supply issue.

Comparing the estimated fixed costs of this model with the other models in the appendix, the explanation above is supported by the fact that we see the largest fixed costs for those who work part time the least. For example, sole parents have the lowest fixed cost and from some of the characteristic-specific components of the fixed costs it is clear that characteristics associated with a higher probability of part-time work reduce the amount of predicted fixed costs. This would explain why having a youngest child in between 5 to 9 years reduces fixed costs by the largest amount[20]. That is, with younger children the parent is more likely to be a non-participant whereas with older children the parent may be more likely to prefer full-time work.

The high fixed costs of working parameters in all models, except for sole parents, are combined with an increase in utility for part-time labour supply increases, thus making the low part-time hours the least attractive and compensating part of the fixed costs with the positive effect of labour supply at higher labour supply levels. Euwals and Van Soest (1999) find a similar effect for some individuals in their sample.

Unobserved heterogeneity

Finally, in all the models described here we allowed for unobserved heterogeneity in the linear labour supply preference parameters, the linear income preference parameter and in the fixed cost parameters by adding a normally distributed error term to these parameters. We found these additional parameters to be highly insignificant and their inclusion did have no effect on the estimated values of the other parameters or on the log likelihood value.

As an example, the results for sole parents are provided in Table A.2. In the first additional specification, the error terms in the different preference parameters are assumed to be independent so that only the variances of the error terms need to be estimated and in the second specification the full covariance matrix of the error terms is estimated.

---

20 Van Soest, Das and Gong (2002) find a similar result.

Quasi concavity

The quadratic utility function is not automatically regular. Therefore, one needs to check for quasi-concavity in each of the observed data points after estimating the model. For all groups, it is found that the two conditions, which are necessary for quasi concavity, are fulfilled at the observed hours in 100 per cent of the cases. From the above results, it can be concluded that the utility function is quasi-concave in the relevant regions of the model.

5.2. Goodness of Fit

The final analysis in this study compares the actually observed levels of labour supply to those predicted by the model (see Tables 3a to 3c). The probabilities of being in each of the categories of labour supply and the expected hours of labour supply are reported. Using a simulation procedure, drawing 1000 times from the estimated parameter distribution, empirical confidence intervals are constructed around the expected number of hours and the probabilities of being in each of the categories of labour supply. This procedure incorporates the uncertainty associated with the parameter estimates as they are reflected in the estimated standard deviations.

Table 3a: Actual and Expected Labour Supply for men (proportion in each category)

	single					married				
		Mean	Confidence interval				Mean	Confidence interval		
Hours per week	**Actual**		**5%**	**Median**	**95%**	**Actual**		**5%**	**Median**	**95%**
11 discrete labour supply points										
0-2.5	0.202	0.196	0.170	0.198	0.212					
2.5 – 7.5	0.005	0.000	0.000	0.000	0.001					
7.5 – 12.5	0.009	0.001	0.001	0.001	0.002					
12.5 – 17.5	0.008	0.005	0.003	0.005	0.007					
17.5 – 22.5	0.013	0.014	0.012	0.014	0.019					
22.5 – 27.5	0.015	0.037	0.033	0.036	0.044					
27.5 – 32.5	0.023	0.078	0.073	0.076	0.087					
32.5 – 37.5	0.075	0.132	0.126	0.131	0.140					
37.5 – 42.5	0.403	0.179	0.172	0.179	0.186					
42.5 – 47.5	0.088	0.193	0.185	0.193	0.200					
> 47.5	0.159	0.166	0.152	0.166	0.177					
6 discrete labour supply points										
0-2.5						0.134	0.133	0.128	0.133	0.138
2.5 – 15						0.011	0.000	0.000	0.000	0.000
15 – 25						0.015	0.010	0.010	0.010	0.011
25 – 35						0.037	0.120	0.116	0.120	0.124
35 – 45						0.496	0.390	0.384	0.390	0.396
> 45						0.307	0.346	0.338	0.347	0.354
Expected hours by age										
all	32.2	32.4	31.6	32.3	33.3	36.7	36.7	36.5	36.7	37.0
Age<30	32.4	32.1	31.3	32.1	33.0	37.0	36.4	35.7	36.4	37.0
Age 31-50	34.3	35.2	34.4	35.2	36.2	39.0	39.3	39.0	39.3	39.5
Age>50	24.7	24.2	22.6	24.1	26.3	31.1	30.9	30.4	30.9	31.3

Table 3b: Actual and expected labour supply for women (proportion in each category)

	single					married				
		Mean	Confidence interval				Mean	Confidence interval		
Hours per week	Actual		5%	Median	95%	Actual		5%	Median	95%
0-2.5	0.258	0.255	0.242	0.255	0.266	0.394	0.394	0.386	0.394	0.402
2.5 – 7.5	0.011	0.003	0.002	0.003	0.004	0.020	0.019	0.017	0.019	0.020
7.5 – 12.5	0.017	0.008	0.006	0.008	0.010	0.034	0.029	0.027	0.029	0.030
12.5 – 17.5	0.017	0.017	0.015	0.017	0.019	0.046	0.042	0.041	0.042	0.044
17.5 – 22.5	0.027	0.033	0.031	0.033	0.036	0.065	0.058	0.056	0.058	0.059
22.5 – 27.5	0.030	0.061	0.059	0.061	0.065	0.052	0.072	0.070	0.072	0.074
27.5 – 32.5	0.045	0.100	0.097	0.100	0.103	0.054	0.083	0.081	0.083	0.085
32.5 – 37.5	0.114	0.135	0.130	0.135	0.139	0.073	0.087	0.085	0.087	0.089
37.5 – 42.5	0.338	0.150	0.145	0.150	0.154	0.179	0.084	0.082	0.084	0.085
42.5 – 47.5	0.067	0.136	0.133	0.136	0.140	0.034	0.073	0.071	0.073	0.075
> 47.5	0.076	0.102	0.096	0.102	0.107	0.050	0.058	0.055	0.058	0.062
Expected hours by age										
all	27.4	27.5	27.0	27.5	28.0	19.1	19.1	18.8	19.1	19.4
Age<30	30.8	30.2	29.7	30.2	30.8	21.0	20.4	19.8	20.4	20.9
Age 31-50	32.0	34.0	33.4	34.0	34.6	19.7	19.9	19.5	19.9	20.2
Age>50	14.1	13.5	12.7	13.4	14.3	15.1	15.2	14.6	15.2	15.7

Table 3c: Actual and expected labour supply for sole parents (proportion in each category)

		Mean	Confidence interval		
Hours per week	Actual		5%	Median	95%
0-2.5	0.533	0.521	0.489	0.523	0.541
2.5 – 7.5	0.031	0.023	0.018	0.022	0.028
7.5 – 12.5	0.037	0.035	0.031	0.035	0.042
12.5 – 17.5	0.033	0.043	0.039	0.042	0.048
17.5 – 22.5	0.038	0.047	0.044	0.047	0.053
22.5 – 27.5	0.030	0.049	0.045	0.049	0.054
27.5 – 32.5	0.033	0.050	0.046	0.050	0.054
32.5 – 37.5	0.054	0.052	0.048	0.052	0.056
37.5 – 42.5	0.131	0.055	0.052	0.055	0.059
42.5 – 47.5	0.030	0.060	0.057	0.060	0.064
> 47.5	0.050	0.064	0.058	0.064	0.072
Expected hours					
all	14.5	14.8	14.1	14.7	15.5
Age<30	6.9	7.1	6.2	7.0	8.4
Age 31-50	16.8	17.2	16.4	17.1	18.0
Age>50	18.6	18.2	16.4	18.1	20.0

From the tables, it is clear that the lowest part-time hours categories are somewhat underpredicted and the category with the highest hours is somewhat overpredicted. It is also clear that the model cannot capture the peak around 40 hours per week in the observed hours. As a result this category is underpredicted, whereas the neighbouring categories are overpredicted.

These under and overpredictions of hours categories in the labour supply model are not transferred to the policy simulations by using the following approach. The impact of prediction errors in the labour supply model on the simulation results is reduced by basing the starting point of the simulations on the actual working hours in the data. That is, labour supply before the reform is fixed on observed labour supply. This prevents prediction errors in the model from impacting on the distribution of working hours in the base situation.

The error term included in the labour supply model to account for optimisation errors (see equation 7) is used to calibrate the model in such a way that observed labour supply is always the starting point. Basically, the procedure is that we draw from the possible values for the error term and only accept those draws for the calculation of the expected labour supply before and after the reform that places the individual at the observed labour supply in the pre-reform situation. The approach uses the unobserved characteristics (the value of the error term) as well as the observed characteristics (such as age or family composition which are used in the calculation of expected utility levels at each labour supply level). The two components jointly determine which labour supply point an individual prefers.

Fewer labour supply points are allowed for married men given the low number of married men working part-time hours (which can be caused by factors on both the supply and the demand side). However, given the probability approach used in the simulation of changes, small changes in labour supply can still be captured even in a ten-hour interval labour supply specification. A small change in labour supply means they may, for example, have a small probability of moving from 30 to 40 hours.

The results of the model can be summarized by calculating the expected hours of labour supply. The expected hours are given in the last row of table 3 and correspond well to the actual average hours of labour supply.

The expected effects of certain policy changes can be calculated by computing the expected numbers in each of the categories, accounting for the changed tax and benefit rules in the computer programs, and comparing these results to the expected numbers using the current tax and benefit rules. Examples of policy simulations using similar models to the ones described in this paper can be found in Creedy, Kalb and Kew (forthcoming). Two examples of simulations using the model for couples described in this paper can be found in Kalb and Kew (2002), which describes the effect of a change in taper rates, and in Kalb, Kew and Scutella (2002), which describes the effect of a change in the family payment taper rate.

From the range in the confidence intervals, it can be seen that most estimates are relatively precise. In addition to predicted values for the whole sample, the tables also present expected labour supply for three age categories. Expected labour supply by subgroup appears to follow the movements in actual hours quite closely. For the smaller subcategories (such as sole parents over 50 years old) the confidence intervals become wider, because individual deviations from the predicted values play a larger role, whereas in larger groups these are averaged out. However, as explained above, in policy simulations these deviations are accounted for in the simulation approach taken.

Comparing the labour supply in the three age groups for the different demographic groups, it is clear that labour supply is highest in the 31 to 50 age category for singles and married men. Labour supply is only slightly lower for the youngest age group, but individuals over 50 seem to reduce their labour supply considerably. Not unexpectedly, married women and sole parents behave differently. Married women have the highest labour supply when they are younger than 30 and there is much less decrease in labour supply going from the middle to the older aged group. A reverse pattern is observed for sole parents, who have the highest labour supply when they are over 50 and the lowest when they are under 30 years of age. This is most likely linked to the age of their children.

6. Conclusions

In this paper, four separate basic labour supply models for couples, single men, single women and sole parents are estimated for use in the Melbourne Institute Tax and Transfer Simulator. The preference parameters for labour supply and income and the parameters for fixed costs include observed heterogeneity in the form of the number and age of children in the income unit, age and education of the head and partner (if present), and the place of residence of the income unit. It was found that adding unobserved heterogeneity did not change the estimated values of the other parameters and the unobserved heterogeneity parameters were all very small and insignificant.

The results are similar for all demographic groups. The basic results seem sensible, with the preference for labour supply highest for people who are in their thirties with a high education level, although education levels seem somewhat more important for women than for men. The preference for labour supply is lower for women with children, in particular when the children are young, whereas no effect is found for married men. Finally, the predicted distribution over the different labour supply hours is similar to the actual distribution. To conclude, the four models in this paper seem a good starting point for further experimentation with alternative specifications and extensions in future research.

In using a labour supply model in microsimulation modelling, policy makers would be concerned about the validity of predictions from the model out of sample and after policy changes. A way of getting some information on the validity of results would be to carry out evaluations after new policies have been introduced and compare the outcomes of the evaluation with the predictions of the model. This is not easy, since finding a policy change and data at the right points in time and selecting appropriate comparison groups can be quite complicated. For example, Blundell and Hoynes (2000) attempt such a comparison and discuss the difficulties they encounter.

Appendix A: Additional labour supply models

Table A.1a: Estimated Parameters of one-adult Labour Supply Models

	Single men		Single women		Sole parents	
	Estimated coefficient	z-value[b]	Estimated coefficient	z-value[b]	Estimated coefficient	z-value[b]
Quadratic terms						
Income $\times 100,000$	-0.0132	-0.57	-0.1421	-2.03	-0.5858	-2.54
Labour supply $\times 100$	-0.4313	-12.79	-0.2711	-11.82	-0.0199	-0.51
Crossproduct						
Inc. & lab. sup. $\times 10,000$	-0.4969	-3.74	-1.5207	-8.65	-1.4297	-2.44
Linear terms						
Income $\times 100$						
constant	0.2214	2.85	0.7474	5.10	2.9512	2.18
Children 0-2 yrs old					0.4762	1.33
Children 3-4 yrs old					-0.1161	-0.40
Children 5-9 yrs old					0.8649	2.72
Number of children					0.1135	1.28
Age/10	0.1412	4.40	0.1013	1.42	-0.2572	-0.40
Age squared/100	-0.0154	-3.76	-0.0028	-0.31	0.0127	0.17
Vocational education	0.0209	1.82	-0.0166	-0.50	-0.0452	-0.30
diploma	-0.0005	-0.03	0.0289	0.60	-0.0165	-0.11
degree	0.0118	0.67	0.1162	2.40	c	c
female					0.0300	0.12
Labour supply						
constant	0.1502	4.55	0.0110	0.63	-0.1473	-4.08
Children 0-2 yrs old					-0.0335	-2.05
Children 3-4 yrs old					-0.0214	-1.57
Children 5-9 yrs old					-0.0534	-3.41
Number of children					-0.0020	-0.63
Age/10	0.0794	8.32	0.0914	12.88	0.0909	6.12
Age squared/100	-0.0103	-7.87	-0.0127	-13.63	-0.0111	-5.85
Vocational education	0.0170	4.67	0.0043	1.45	0.0169	3.89
diploma	0.0141	2.28	0.0207	5.33	0.0242	4.11
degree	0.0234	4.37	0.0308	8.06	c	c
female					-0.0486	-4.10
Fixed costs/100						
Constant	17.3972	6.74	5.2641	9.44	2.3593	5.95
Live in capital city	-0.3486	-1.65	-0.1248	-0.90	0.0563	1.10
Children 0-4 yrs old					-0.2301	-0.91
Children 5-9 yrs old					-0.6367	-2.62
Live in NSW	-0.3047	-1.24	0.0476	0.35	0.2290	3.38
Female					-0.4902	-1.88

a Eleven discrete points of labour supply are distinguished for each person: 0 hours for non-participants and people working less than 2.5 hours, 5 hours for people working from 2.5 to 7.5 hours, 10 hours for people working from 7.5 to 12.5 hours, 15 hours for people working from 12.5 to 17.5 hours, 20 hours for people working from 17.5 to 22.5 hours, 25 hours for people working from 22.5 to 27.5 hours, 30 hours for people working from 27.5 to 32.5 hours, 35 hours for people working from 32.5 to 37.5 hours, 40 hours for people working from 37.5 to 42.5 hours, 45 hours for people working from 42.5 to 47.5 hours, and 50 hours for people working more than 47.5 hours.

b The z-value indicates the level of significance of the estimated coefficients. A value of 1.96 or more indicates that the parameter is significant at the 5% level at least. That is we are quite confident it is not equal to zero. The higher the z-value, the more precise the estimated coefficient is.

c For sole parents people with a degree are categorized in a group with those having a diploma, because of the limited number of observations on sole parents with a higher level of education.

Table A.2: Estimated Parameters of the Labour Supply Model for Sole parents

	No unobserved heterogeneity		Variance of error terms only		Full covariance matrix of error terms	
	Estimated coefficient	z-value[b]	Estimated coefficient	z-value[b]	Estimated coefficient	z-value[b]
Income squared $\times$ *100,000*	-0.5858	-2.54	-0.5871	-2.56	-0.5888	-2.58
Labour supply squared $\times$ *100*	-0.0199	-0.51	-0.0198	-0.51	-0.0196	-0.51
Inc. & lab. sup. $\times$ *10,000*	-1.4297	-2.44	-1.4304	-2.45	-1.4334	-2.47
income$\times$ 100						
constant	2.9512	2.18	2.9597	2.18	2.9632	2.18
Children 0-2 yrs old	0.4762	1.33	0.4768	1.34	0.4803	1.36
Children 3-4 yrs old	-0.1161	-0.40	-0.1149	-0.39	-0.1128	-0.39
Children 5-9 yrs old	0.8649	2.72	0.8672	2.73	0.8632	2.73
Number of children	0.1135	1.28	0.1136	1.29	0.1142	1.30
Age/10	-0.2572	-0.40	-0.2604	-0.41	-0.2615	-0.41
Age squared/100	0.0127	0.17	0.0132	0.18	0.0134	0.18
Vocational education	-0.0452	-0.30	-0.0444	-0.29	-0.0432	-0.28
Diploma/degree	-0.0165	-0.11	-0.0166	-0.11	-0.0165	-0.11
female	0.0300	0.12	0.0275	0.11	0.0258	0.10
Labour supply						
constant	-0.1473	-4.08	-0.1474	-4.08	-0.1477	-4.11
Children 0-2 yrs old	-0.0335	-2.05	-0.0335	-2.05	-0.0336	-2.07
Children 3-4 yrs old	-0.0214	-1.57	-0.0214	-1.58	-0.0214	-1.59
Children 5-9 yrs old	-0.0534	-3.41	-0.0535	-3.41	-0.0532	-3.42
Number of children	-0.0020	-0.63	-0.0020	-0.63	-0.0020	-0.63
Age/10	0.0909	6.12	0.0909	6.12	0.0910	6.11
Age squared/100	-0.0111	-5.85	-0.0111	-5.84	-0.0111	-5.84
Vocational education	0.0169	3.89	0.0168	3.89	0.0168	3.87
Diploma/degree	0.0242	4.11	0.0242	4.11	0.0242	4.11
female	-0.0486	-4.10	-0.0486	-4.10	-0.0485	-4.11
Fixed costs /100						
Constant	2.3593	5.95	2.3578	5.98	2.3549	6.06
Live in capital city	0.0563	1.10	0.0562	1.10	0.0559	1.10
Children 0-4 yrs old	-0.2301	-0.91	-0.2306	-0.91	-0.2309	-0.92
Children 5-9 yrs old	-0.6367	-2.62	-0.6375	-2.63	-0.6338	-2.66
Live in NSW	0.2290	3.38	0.2287	3.38	0.2284	3.38
Female	-0.4902	-1.88	-0.4886	-1.88	-0.4870	-1.89
Unobserved heterogeneity						
Variance income			0.0005	0.61	0.0028	0.94
Variance labour supply			0.0000	0.00	0.0000	0.83
Variance fixed cost			0.0000	0.02	0.0010	0.86
Covariance inc. & ls					-0.0001	-0.93
Covariance inc & fc					-0.0016	-0.95
Covariance ls & fc					0.0001	0.87
Log likelihood	-2768.8		-2768.7		-2768.5	

a Eleven discrete points of labour supply are distinguished for each person: 0 for non-participants and people working less than 2.5 hours, 5 for people working from 2.5 to 7.5 hours, 10 for people working from 7.5 to 12.5 hours, 15 for people working from 12.5 to 17.5 hours, 20 for people working from 17.5 to 22.5 hours, 25 for people working from 22.5 to 27.5 hours, 30 for people working from 27.5 to 32.5 hours, 35 for people working from 32.5 to 37.5 hours, 40 for people working from 37.5 to 42.5 hours, 45 for people working from 42.5 to 47.5 hours, and 50 for people working more than 47.5 hours.

b The z-value indicates the level of significance of the estimated coefficients. A value of 1.96 or more indicates that the parameter is significant at the 5% level at least. That is we are quite confident it is not equal to zero. The higher the z-value, the more precise the estimated coefficient is.

References

Aaberge, R., U. Colombino and S. Strøm (1999). Labour supply in Italy: an empirical analysis of joint household decisions, with taxes and quantity constraints. *Journal of Applied Econometrics*, 14, 402-422.

Apps, P.F. and R. Rees (1996). Labour supply, household production and intra-family welfare distribution. *Journal of Public Economics*, 60, 199-219.

— (1997). Collective Labor Supply and Household Production. *Journal of Political Economy*, 105(1), 178-190.

— (2000). Household Production, Full Consumption and the Costs of Children. Discussion Paper no. 157, Institute for the Study of Labor (IZA), Bonn.

Australian Bureau of Statistics (1998), *Consumer Price Index*, catalogue number 6401.0, June quarter.

Bingley, P. and I. Walker (1997). The labour supply, unemployment and participation of lone mothers in in-work transfer programmes. *Economic Journal*, 107, 1375-1390.

Blundell, R. and H. Hoynes (2000). Has 'In-Work' Benefit Reform Helped the Labour Market. preliminary paper prepared for the CEP-IFS-NBER Premier League Meeting, December 8-9, 2000.

Bourguignon, F. and P.-A. Chiappori (1994). The collective approach to household behaviour. in *The measurement of household welfare* ed. by R. Blundell, I. Preston and I. Walker, Cambridge: Cambridge University Press.

Browning, M., F. Bourguignon, P.-A. Chiappori and V. Lechene (1994). Income and outcomes: a structural model of intrahousehold allocation. *Journal of Political Economy*, 102(6): 1067-1096.

Burtless, G. and J.A Hausman (1978). The effect of taxation on labor supply: evaluating the Gary Negative Income Tax experiment. *Journal of Political Economy*, 86(6), 1103-1130.

Callan, T. and A. van Soest (1996). Family Labour Supply and Taxes in Ireland. Discussion paper, mimeo, Tilburg University.

Creedy, J., A.S. Duncan, M. Harris, R. Scutella (2002). *Microsimulation Modelling of Taxation and The Labour Market: The Melbourne Institute Tax and Transfer Simulator*. Cheltenham: Edward Elgar.

Creedy, J., Kalb, G. and H. Kew (forthcoming). Flattening the Effective Marginal Tax Rate Structure in Australia: Policy Simulations Using the Melbourne Institute Tax and Transfer Simulator. *The Australian Economic Review*.

Deaton, A. and J. Muellbauer (1980). *Economics and Consumer Behavior*. New York: Cambridge University Press.

Duncan, A., C. Giles and J. MacCrae (1999). Household labour supply, childcare costs and in-work benefits: modelling the impact of the working families tax credit in the UK. *Paper, Institute for Fiscal Studies*, London.

Duncan, A. and M.N. Harris (2002). Simulating the Effect of Welfare Reforms among Sole Parents in Australia. *Economic Record*, 78, 249-263.

Duncan, A. and J. MacCrae (1999). Household labour supply, childcare costs and in-work benefits: modelling the impact of the working families tax credit in the UK. *Mimeo, Institute for Fiscal Studies*, London.

Eissa, N. and H.W. Hoynes (1999). The earned Income Tax Credit and the Labor Supply of Married Couples. Discussion Paper no. 1194-99, Institute for Research on Poverty, University of Wisconsin-Madison.

Euwals, R. (2001). Female labour supply, flexibility of working hours, and job mobility. *The Economic Journal*, 111(May), C120-C134.

Euwals, R. and A. van Soest (1999). Desired and actual labour supply of unmarried men and women in the Netherlands. *Labour Economics*, 6, 95-118.

Eyland, E.A., C.A. Mason and H.M. Lapsley (1982). Determinants of female employment. *Economic Record*, 58, 11-17.

Fraker, T. and R. Moffitt (1988). The effect of food stamps on labor supply: a bivariate selection model. *Journal of Public Economics*, 35, 25-56.

Gerfin, M. (1993). Simultaneous discrete choice model of labour supply and wages for married women in Switzerland. *Empirical Economics*, 18, 337-356.

Hagstrom, P.A. (1996). The food stamp participation and labor supply of married couples; an empirical analysis of joint decisions. *Journal of Human Resources*, 31(2), 383-403.

Hausman, J.A. (1979). The econometrics of labor supply on convex budget sets. *Economics Letters*, 3, 171-174.

— (1985). The econometrics of nonlinear budget sets. *Econometrica*, 53, 1255-82.

Heckman, J.J. (1979). Sampled selection bias as a specification error. *Econometrica*, 47(1), 153-161.

Hoynes, H.W. (1996). Welfare transfers in two-parent families: labor supply and welfare participation under AFDC-UP. *Econometrica*, 64(2), 295-332.

Kalb, G. (1999). Labour Supply and Welfare Participation in Australian Two-adult Households; Comparing 1986/1987 with 1994/1995. Working Paper No. BP-34, Centre of Policy Studies, Monash University, Australia.

Kalb, G. (2000). Labour Supply and Welfare Participation in Australian Two-Adult Households: Accounting for involuntary unemployment and the 'cost' of part-time work. Working Paper No. BP-35, Centre of Policy Studies, Monash University, Australia.

Kalb, G. and H. Kew (2002). The effect of a reduced allowance and pension taper rate; Policy simulations using the Melbourne Institute Tax and Transfer Simulator. Melbourne Institute Working Paper Series, WP No. 25/02, Melbourne Institute of Applied Economic and Social Research, The University of Melbourne.

Kalb, G., H. Kew and R. Scutella (2002). The effect of a reduced family payment taper rate; Policy simulations using the Melbourne Institute Tax and Transfer Simulator. Melbourne Institute Working Paper Series, WP No. 26/02, Melbourne Institute of Applied Economic and Social Research, The University of Melbourne.

Kalb, G. and Scutella. R. (2002), "Estimation of Wage Equations in Australia: Allowing for Censored Observations of Labour Supply", Melbourne Institute Working Paper Series, WP

No. 8/02, Melbourne Institute of Applied Economic and Social Research, The University of Melbourne.

Keane, M. and R. Moffitt (1998). A Structural Model of Multiple Welfare Program Participation and Labor Supply. *International Economic Review*, 39(3), 553-589.

Killingsworth, M.R. (1983). *Labor Supply*. New York: Cambridge University Press.

Laisney, F., M. Lechner, A. van Soest and G. Wagenhals (1992). A life cycle labour supply model with taxes estimated on German panel data: the case of parallel preferences. *Discussion Paper* no. 93-01, Zentrum für Europäische Wirtschaftsforschung GmbH, Labour Economics and Human Resources Series.

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

Moffitt, R. (1986). The econometrics of piecewise-linear budget constraints: a survey and exposition of the maximum likelihood method. *Journal of Business and Economic Statistics*, 4(3), 317-328.

— (1990). Taxation and Labor Supply in Industrial Countries. *Journal of Human Resources*, 25(3), 313-558.

Murray, J. (1996). Modelling the labour supply behaviour of sole parents in Australia. In M. McAleer, P.W. Miller and C. Ong, *Proceedings of The Econometric Society Australasian Meeting 1996*, 4: Microeconometrics. Perth: University of Western Australia, 507-546.

Ross, R.T. (1986). Analysis of the 1980 Sydney Survey of Work Patterns of Married Women: further results. *Economic Record*, 62, 325-337.

Van Soest, A. (1995). Structural models of family labor supply; a discrete choice approach. *Journal of Human Resources*, 30(1), 63-88.

Van Soest, A., M. Das and X. Gong (2002). A structural labour supply model with flexible preferences. *Journal of Econometrics*, 107, 345-374.

Van Soest, A., A. Kapteyn and P. Kooreman (1993). Coherency and regularity of demand systems with equality and inequality constraints. *Journal of Econometrics*, 57, 161-188.