The University of New South Wales
School of Economics

Modelling Decisions to Volunteer at a Household Level

MICHAEL ROSS ABBOTT

Bachelor of Commerce / Bachelor of Science

Honours in Business Economics

Supervisor: Professor Denzil Fiebig

Monday, 26th October 2009

Declaration

I hereby declare that this submission is my own work and any contributions or materials by other authors used in this thesis have been appropriately acknowledged. This thesis has not been previously submitted to any other university or institution as part of the requirements for another degree or award.

Michael Ross Abbott

26 October 2009

Acknowledgements

First and foremost, I would like to express my gratitude to my supervisor Professor Denzil Fiebig. Denzil's assistance and guidance during the honours year has been invaluable. I have also learnt a great deal from the opportunity to work as a research assistant with Denzil over several years.

In addition, I would like to acknowledge Dr Shiko Maruyama along with the other participants and faculty members at the 2009 National Honours Colloquium for their helpful comments. Associate Professor Denise Doiron and Dr Valentyn Panchenko also provided useful feedback.

As well, I would to thank Malcolm Chaikin for his generosity in providing financial support over the course of my degree through the Malcolm Chaikin Scholarships.

Finally, to the Honours Class of 2009, Kieran, Andy, Gwen, Dave, Andrew, Shish, Felix, Rahul, Deepika, Spiro, Hien, Tommy and Gordon, thank you for making the year memorable. Along with, of course, my family for their continuing love and support over many years.

Disclaimer

This paper uses unit record data from the Household, Income and Labour Dynamics in Australia (HILDA) Survey. The HILDA Project was initiated and is funded by the Australian Government Department of Families, Housing, Community Services and Indigenous Affairs (FaHCSIA) and is managed by the Melbourne Institute of Applied Economic and Social Research (Melbourne Institute). The findings and views reported in this paper, however, are those of the author and should not be attributed to either FaHCSIA or the Melbourne Institute.

Table of Contents

List of Tables

Abstract

Volunteering is widespread in Australia and within households it is common to observe couples making similar decisions on whether to volunteer or not. As a result, this thesis focuses on modelling interactions between heterosexual couples (married or de facto) in making volunteering decisions using a sample of 3255 couples from Wave 6 of the Household, Income and Labour Dynamics in Australia (HILDA) Survey. The key finding is that the positive correlation between male and female partners volunteering can be attributed to a large endogenous effect of volunteering. It is found that the female volunteering increases the male's probability of volunteering by around 0.15 while the male volunteering increases the female's probability of volunteering by around 0.20. In addition, it is found that volunteers are generally more educated, have children, are aged around 50 and attend religious services.

The models used in this thesis can be broadly classified as assuming cooperative and non-cooperative behaviour within couples and are based on equilibrium concepts from game theory (Stackelberg leader and pure strategy Nash equilibria). They are compared to a bivariate probit with no endogenous effects as a base and found to provide a significant improvement in fit. While it may be possible to choose a preferred model and therefore make some comment about the nature of decision making within couples, it is established that the substantive inferences are similar across the models. This can be interpreted as meaning they all provide a reasonable approximation to reality.

1 Introduction

In a broad sense, volunteering refers to individuals supplying labour without monetary reward. Volunteering is widespread in Australian society with the ABS 2006 Voluntary Work Survey (ABS, 2007) estimating that 34% of the Australian population aged 18 years and over participated in voluntary work, contributing some 713 million hours to the community. Indeed, it is possible to list many activities and organisations in Australian society in which volunteers play a crucial role such as running school canteens, coaching sports teams, delivering meals to the elderly and the State Emergency Service (SES). On this basis alone, it is clear that volunteering is an activity of substantial economic and social value to Australian society.

To further illustrate the value of volunteering, a number of attempts have been made to value in monetary terms the contribution of volunteers to Australian society. Bowman (2009) summarises several of the methods that have been used for this purpose. The value of volunteer work to the Australian economy in 1997 has been estimated as $24 billion (ABS, 2001) and $42 billion (Ironmonger, 2000). Volunteering Australia (2009) provides an updated figure from Ironmonger which values the contribution of volunteering as $70 billion in 2008. These numbers constitute a substantial percentage of Australian GDP.

Despite this, until recently volunteering and the non-profit sector in general have received relatively little attention from economists and researchers. This is beginning to change with the UN General Assembly passing a resolution in 2001 (the International Year of Volunteers) recognising the importance of volunteering and calling on member governments to encourage and support volunteering (United Nations, 2001). There is also current work being undertaken to develop guidelines for measuring the work of volunteers and for the International Labour

Organization to issue a *Manual on the Measurement of Volunteer Work* (John Hopkins University, 2008). In Australia, the Productivity Commission is currently undertaking a study of the contribution of the Not for Profit sector where volunteering plays a key role (Productivity Commission, 2009a).

This indicates that volunteering contributes substantially to Australian society and is currently an area of considerable research interest, providing a general motivation for the focus of this thesis on volunteering. More specifically, this thesis focuses on modelling volunteering decisions at the level of heterosexual couples and understanding the interactions between partners in making volunteering decisions. Freeman (1997) comments on the fact there is a positive association in volunteering among spouses, which is also evident in Australian data (in this case HILDA). This thesis will explore the possible explanations for this positive association by considering several models which capture alternative explanations of the observed correlation in volunteering decisions in Australian couples. From a policy perspective, it is useful to know if this observed correlation is attributable to an individual's partner volunteering directly impacting on their own probability of volunteering since this leads to a multiplier effect. That is, policies which successfully encourage a group of individuals to volunteer will have the added bonus of also encouraging their spouses to volunteer. Empirical evidence in this thesis suggests that this is the case.

From an econometric perspective, a bivariate probit model might seem like a natural choice for modelling a couple's volunteering decision which captures the fact the decision is made by two individuals. However, complications arise in a naive implementation of a bivariate probit model. In particular, the seemingly natural extension to accommodate each partner's decision to volunteer directly affecting their spouses' probability of volunteering by including each individual's volunteering decision as an explanatory variable for their partner's volunteering decision is problematic. The only allowable specifications are recursive having only the decision of one partner affecting their partner's probability of volunteering.

As a result, to develop a model which is reasonably complete in its ability to capture different sources of the observed correlation in a couple's volunteering decisions, it is necessary to make some assumptions regarding how couples interact. Since it is difficult a priori to determine whether such assumptions are valid or what their impact is on the results, this thesis will utilise a variety of models which draw on equilibrium concepts from game theory and can be broadly classified as cooperative or non-cooperative. In particular, the non-cooperative models will use the concept of a Nash equilibrium in pure strategies (Bjorn and Vuong, 1984) and a Stackelberg-leader equilibrium (Bjorn and Vuong, 1985). The equilibrium concept used in the cooperative model is developed from the Nash equilibrium in pure strategies by imposing Pareto optimality on the observed decisions (Kooreman, 1994). The key findings in this thesis are found to be robust across these models.

The contributions of this thesis to the relevant literature fall into two main areas. Firstly it aims to make a contribution to the rather sparse literature on volunteering in general and more specifically Australia. Secondly it aims to contribute to the literature on the use of behavioural models in empirical economics. There are very few applications of the models considered in this thesis in the literature and therefore this thesis aims to contribute further insight and discussion regarding their use.

In regard to the first aspect, to my knowledge, none of the literature to date on volunteering has considered the reasons behind the observed positive correlation in couples' volunteering decisions. In addition, there are very few Australian studies on volunteering. As a result it intends to make a contribution to the general literature on volunteering by providing an analysis of couples' volunteering decisions. In addition, it plans to contribute to the literature on volunteering in Australia by considering volunteering from an economic perspective. A better understanding of this important activity is crucial to developing policies which support and nurture volunteering in the community.

This thesis plans to contribute to the literature on the use of behavioural models by applying these models to a different decision made by couples. Previously these models have only been applied to areas such as paid labour supply, retirement choice and contraceptive choices. This will provide further insight into the use of these models as well as aiding in the interpretation of these models by detailing the computation of average partial effects for a number of variables.

Also, in order to estimate the Nash model (see Section 4.4.2 for details) by maximum likelihood estimation it is necessary to make some assumptions in terms of probability weights regarding how the observed outcome is determined when there is no Nash equilibrium in pure strategies or there are multiple Nash equilibria in pure strategies. This thesis investigates numerically the impact of varying these probability weights on the estimation results and finds that the substantive inferences are not affected. While there has been substantial theoretical work on alternative estimation methods which do not require such assumptions, little is known about the sensitivity of maximum likelihood estimation to these probability weights and as such this thesis will also make a contribution to this area.

2 Literature Review

2.1 Volunteering Behaviour

The models considered in the literature for volunteering behaviour are broadly based on the consumption and investment models of volunteering proposed by Menchik and Weisbrod (1987). The consumption model assumes that the giving of time is a normal utility bearing good while the investment model assumes that volunteering is an activity that raises one's future earning power. Menchik and Weisbrod (1987) also study volunteering empirically by estimating a tobit model for volunteer hours using data on individuals who are the sole wage earner in their household. It is not possible to impose restrictions in the model estimated by Menchik and Weisbrod (1987) which provide strong tests to discriminate between the consumption and investment models. The only approach suggested for discriminating between the models is to consider the effect of income and age on volunteering.

Under the consumption model, individuals divide their time between leisure, volunteering and market work, and their income between charitable donations and conventional consumption expenditures to maximise utility subject to a budget constraint. There are three possible explanations in the literature as to why people may wish to undertake volunteer work. These are altruistic motives, a public goods model and a private goods model.

Unger (1991) discusses the altruistic motive for volunteering and using a survey of a US Midwestern city concludes that there is evidence of altruism in volunteering behaviour. Duncan (1999) develops a public goods model in which an individual is only motivated by the public goods provided by charity. An individual's contributions to charity encompass both volunteering and monetary donations. Duncan (1999) finds that there is support for money and time donations being perfect substitutes which is predicted by the model but not for the crowd-out hypothesis, which predicts that increasing government spending on public goods will perfectly crowd-out charitable donations. Finally, individuals may derive private benefits from the act of giving due to factors such as a desire to avoid scorn of others or to receive social acclaim which is the key premise behind Menchik and Weisbrod's (1987) use of the consumption model. In addition, it may be the case that there is heterogeneity in why people volunteer or these different factors all have some impact on individuals. Also, since there are clear similarities between volunteering and donations, the literature on donations has some relevance. Andreoni (1990) notes in the context of donations, individuals may donate not only because of the value they place on the public good provided but also because of a "warm glow" effect. This is described as impure altruism.

An empirical application of the consumption model is provided by Freeman (1997). The paper notes that responses to surveys on volunteering may be highly sensitive to the wording of the question and context. The paper uses data from a supplement to the May 1989 Current Population Survey and the 1990 Gallup Survey conducted in the US to estimate a number of multiple regression models for the decision to volunteer and the number of hours volunteered. Freeman (1997) comments that volunteering does not appear to fit a standard opportunity cost of time explanation since volunteers are generally people with higher potential earnings or greater demands on their time, that is the employed, married persons, those with larger families, persons in the 35-54 peak earnings ages, the more highly educated, professionals and managers. This also appears to be the case in Australia with a larger proportion of the employed and those with diplomas and degrees volunteering their time (ABS, 2007). Freeman (1997) makes three further points which are there is a strong positive correlation between an individual volunteering and their spouse volunteering, the fact being asked to volunteer is an important determinant of volunteering behaviour (Bryant, et. al. (2003) provide further discussion on this point) and there is a strong positive correlation between donating to charity and donating time.

Under the investment model, individuals engage in volunteer work only to raise future potential earnings. The mechanism for this is the work experience and potentially valuable contacts gained through volunteering. Cugno and Ferrero (2004) modify this approach by considering volunteering as an investment activity in which individuals exert themselves for free in pursuit of a prize or reward, e.g. competition between environmental interest groups in the hope of being appointed to manage the resource in question. Katz and Rosenberg (2005) view volunteering as a way of signalling underlying traits to employers. More recently, the benefits of volunteering have been considered in terms of improved labour market mobility encompassing pecuniary and non-pecuniary benefits (Polidano, et. al., 2009).

Day and Devlin (1998) discuss volunteering from an investment perspective and using Canadian data attempt to estimate the return to volunteering. Their results suggest that the return to volunteering amounts to 6% to 7% of annual earnings, which they state is smaller in magnitude than most estimates of male-female earnings differentials. In addition they suggest that controlling for potential simultaneity of volunteering with income in their model is difficult. Day and Devlin (1998) attribute this simultaneity to two sources: the wage differential between volunteers and non-volunteers may motivate individuals to volunteer; and, individuals with higher incomes may be more likely to volunteer if volunteering is a normal good. Day and Devlin (1998) attempt to provide some insight into the issue by including dummy variables for different types of volunteering organisations in the regression. This is because if there is a positive correlation between income and volunteering solely because high income individuals are more likely to volunteer then the type of volunteer organisation should not matter.

Prouteau and Wolff (2006) also consider volunteering from an investment perspective focusing on volunteers who perform management tasks using French data from 1998 and 1999. Prouteau and Wolff (2006) claim that from an investment perspective it is important not only whether volunteers receive a wage premium but also whether this wage premium influences their decision to volunteer. Using a switching regression model with endogenous switching they find that volunteers receive a wage premium of between 5% and 6% in the public sector which does not have an effect on the decision to volunteer. In the private sector, the premium is negative but has a positive effect on the volunteering decision. As such their results do not support the investment model and the paper also finds that there exists an intrinsic benefit in volunteering in terms of relational goods, defined as "outputs of a communicative and affective nature produced through social interactions", which suggests that a consumption motive is more appropriate.

Hackl, et. al. (2007) also analyse volunteering from an investment perspective and estimate an average wage premium for volunteering of up to 18.5% after accounting for the potential endogeneity of the volunteering decision in the wage equation using instrumental variables. This estimate of the wage premium seems high and may be attributable to the authors' choices of instrumental variables, namely whether or not an individual has been engaged in a club during childhood or adolescence and whether or not they have a volunteering partner. Given the use of cross-sectional data where a representative individual from the household is interviewed, using whether or not an individual has a volunteering partner as an instrumental variable makes the strong assumption that the causality only runs one way. That is, an individual's partner volunteering affects their own volunteering but not the other way around.

This literature surrounding the consumption and investment models of volunteering has been extended in a number of directions. Several studies have focused on the role of policy in affecting volunteering behaviour while discussing considerations that are important for this thesis from a modelling perspective. Firstly, depending on the relationship between donations of time and money, altering the taxation treatment of donations may have an effect. In a study of the effect of the tax price of money on charitable giving and volunteering, Brown and Lankford (1992) find evidence that volunteering and monetary donations are gross complements. The estimated elasticities are -1.7 for money giving, -2.1 for women's time and -1.1 for men's time. They also discuss whether it is appropriate to include the wage rate or hours available to volunteer as explanatory variables in a volunteering equation. The point raised is that if hours of work are fully flexible the wage rate represents the opportunity cost of time, however if hours of work are constrained, available hours becomes the theoretically relevant variable. Andreoni, et. al. (1996) also finds evidence that donations of money and time are gross complements.

Day and Devlin (1996) consider the possibility that government expenditure may crowd out volunteering since in some cases provision of services by volunteers may be a substitute for provision of services by government. A weekend "working-bee" at the local primary school to install new play equipment could be replaced by the government employing workers to install the equipment. On the other hand, increased government funding may enhance the ability of volunteer organisations to absorb new volunteers. Potentially, the current additional Federal Government funding for schools will provide a good natural experiment to investigate this in Australia. Day and Devlin (1996) jointly estimate two equations for the decision to participate in volunteer work and the number of hours of volunteer work undertaken using the 1987 Survey of Volunteer Activity conducted by Statistics Canada. Day and Devlin (1996) account for the possible endogeneity of income (since volunteering may increase income under the investment model) by replacing income with an estimate of the individual's spouse's income and variables that have been shown to be important determinants of income. The paper finds that in general increases in government expenditure are associated with increases in the probability of volunteering, with the direction of the effect varying by sector.

In addition, one study has relaxed the assumption that volunteer labour is a homogenous activity. Segal and Weisbrod (2002) consider differences in volunteering behaviour across sectors, in their case health, education and religion. They find differences in the marginal effects of personal demographics, household composition and tax status across sectors.

Day and Devlin (1997) use Canadian data from 1987 and find that as much as one-third of the male-female earnings gap may be attributable to differential returns to volunteer work. Day and Devlin (1997) find evidence that on average male volunteers earn around 11% more than their non-volunteering counterparts while female volunteers and non-volunteers earn around the same. This may be attributable to differences in the type of volunteer activity undertaken.

A number of studies using older data have specifically considered the volunteer labour supply of women. This is partly explained by Schram and Dunsing (1981) who note that volunteers were usually married women at this time. Mueller (1975) presents an early study of the determinants of volunteer work by women using a simple regression analysis. Schram and Dunsing (1981) use multiple regression to study whether married women participate in volunteer work and the degree to which they participate. Carlin (2001) uses 1975 to 1976 US time diary data for married women and studies the decision to volunteer using a probit model and the number of hours volunteered using a tobit model. Carlin (2001) finds that the wage elasticity for hours volunteered is positive while rises in the net wage have a small negative effect on participation. In addition, having more children increases the probability of volunteering but reduces the number of hours volunteered. There is no clear conclusion regarding whether hours available or the wage rate should be included to represent the opportunity cost of time and while the study lacks appropriate data there is a tentative finding that donations of time and money appear to be substitutes overall.

2.2 Econometric approach

The papers to date on volunteering have used a variety of econometric models as discussed in the previous section, however none of the papers have focused on the correlation between male and female partners volunteering noted by Freeman (1997). This may be partly due to a lack of appropriate data to undertake such a study. Since appropriate data is available for Australia through HILDA, this thesis will focus on this correlation and as such this section of the literature review will focus on methods to capture this correlation. Polidano, et. al. (2009) is the only other paper to use the HILDA dataset however they do not consider the possible interaction effects between partners.

In regard to the focus on the decision to volunteer rather than the number of hours volunteered, a number of the papers discussed have modelled the volunteering decision (some further examples are Vaillaincourt (1994) using a probit model and Gomez and Gunderson (2003) using a logit model which is interpreted using the concept of a household production function). Day and Devlin (1996) claim that since many individuals only volunteer a small number of hours it may well be the case that the decision to volunteer is more important in determining the total number of volunteer hours than the number of volunteer hours an individual chooses to supply. Polidano, et. al. (2009) comment on the possibility of substantial recall error when asking individuals how many hours they volunteer. This is part of the motivation for their use of an Ordered Generalised Extreme Value model for various levels of volunteering. The basis for this is that there is less error in an individual recalling whether they volunteered or not and which category they fall into (i.e. high, medium or low number of hours volunteered). This clearly indicates that the initial decision to volunteer is important in its own right, which motivates the focus on the decision to volunteer in this thesis. Further issues regarding the decision of how many hours to volunteer when taking into account the correlation between male and female partners volunteering are left for future work.

Manski (1993) notes that we may view the observed correlation as arising from three main sources, which motivates the following latent variable framework (Equation (2-1)):

1. Exogenous (contextual) effects - Volunteering is correlated with observable factors which are in turn correlated with whether or not couples form. For instance, volunteering may be correlated with education and generally couples have similar education levels which would give rise to the observed correlation.
2. Correlated effects - Volunteering is correlated with unobservable factors which are correlated between couples. This can be taken into account by allowing correlation in the error terms (ε_1 and ε_2).
3. Endogenous effects - An individual's partner's volunteering decision directly affects their own volunteering decision which is allowed for by including the dummy variables (y_1 and y_2).

$$\begin{cases} y_1^* = x'_1\beta_1 + \gamma_1 y_2 + \varepsilon_1 \\ y_2^* = x'_2\beta_2 + \gamma_2 y_1 + \varepsilon_2 \\ y_i = 1 \; if \; y_i^* > 0, 0 \; otherwise \end{cases} \tag{2-1}$$

In the framework given in Equation (2-1), y_1^* represents the male's propensity to volunteer, y_2^* represents the female's propensity to volunteer, y_1 is a binary outcome variable for the male's decision to volunteer, y_2 is a binary outcome variable for the female's decision to volunteer and x_1 and x_2 contain observable factors we wish to control for.

To estimate this model we may assume a bivariate normal distribution for the errors which results in the simultaneous probit model. However, a well known difficulty with this model is that it is incoherent without a suitable restriction being placed on the parameters (that is it produces total probabilities in excess of one). Kooreman (1994) comments that the root of the coherency problem is that the relationship between the errors and the outcome variables is not one to one.

The relevant coherency condition for the model in this case is $\gamma_1\gamma_2 = 0$ (Heckman, 1978). This makes the resulting system triangular restricting the simultaneity in the model. This is problematic since any resulting model specification would assume that the male makes their decision independently of the female and then the female makes their decision conditional on the male's decision or vice-versa. This does not allow us to fully capture the endogenous effects and as such seems to be an overly restrictive class of models. It should be noted that the bivariate probit model with $\gamma_1 = \gamma_2 = 0$ accommodates exogenous and correlated effects but not endogenous effects.

With this in mind, some papers have considered other ways to link the latent variables (y_1^* and y_2^*) with the observed outcomes (y_1 and y_2) which do not result in a need to impose restrictive coherency conditions. Bjorn and Vuong (1984, 1985) develop two models based on equilibrium concepts from game theory in which both endogenous variables can be included without coherency restrictions. Bjorn and Vuong (1984) assume that the observed outcome is the Nash equilibrium in pure strategies of a two player game with payoffs given by the latent variables (referred to as the Nash model in Section 2.2). Bjorn and Vuong (1985) assume that the observed outcome is the Stackelberg equilibrium of a Stackelberg leader game played between the partners with payoffs given by the latent variables (referred to as the Stackelberg model in Section 2.2).

To my knowledge this approach has not been applied in the context of volunteering decisions, however it has been applied to several other decisions where there is the possibility of interdependence between partners' decisions. Bjorn and Vuong (1984, 1985) both apply the models to a study of husband/wife labour force participation using the Panel Study on Income Dynamics. Bresnahan and Reiss (1991) provide further discussion of these models and consider their applicability to family labour supply. Kooreman (1994) discusses these models as well as one assuming Pareto optimality of observed outcomes and applies them to the joint labour force decisions of husbands and wives using a sample of Dutch households.

Chao (2002) applies the Nash and Stackelberg-leader models along with a consensus model (in this case a bivariate probit) to contraceptive choice by treating use of the pill as a female decision and use of the condom as a male decision. The data used comes from the 1995 National Survey of Family Growth. Jia (2005) formulates a mixed model of household retirement choice by using a multinomial logit model for cooperative households and a Stackelberg leader model with the male as leader for non-cooperative households. This model is applied to data based on administrative registers from Statistics Norway. Further evidence regarding the presence of non-cooperative behaviour within households is available in Dercon and Krishnan (2000) and Udry (1996) who use data from Ethiopia and Burkina Faso respectively.

In the Stackelberg leader model, the assumption that the observed outcomes are generated by a Stackelberg leader game is sufficient to define how the observed outcome variables are linked to the latent variables. In the Nash model this is not the case due to non-existence of equilibrium and multiple equilibria. More recent literature has focused on different solutions to this problem, which Tamer (2003) notes is an example of an incomplete econometric structure. Berry and Reiss (2006) discuss possible solutions to this problem:

1. Introduce an equilibrium selection mechanism which specifies which equilibrium is picked as part of the econometric model. This has been implemented by introducing additional parameters into the model which specify the probability each equilibrium is picked in the case of multiple equilibria or non-existence of equilibrium. This approach has been adopted by Bjorn and Vuong (1984), Kooreman (1994) and Chao (2002) and in each of these papers the probabilities were determined arbitrarily prior to estimation. Tamer (2003) extends this approach to also estimate the additional parameters from the exogenous variables in the case when both γ_1 and γ_2 have the same sign. Bajari, et. al. (2007) develop an estimator that can be used to estimate both the utilities and equilibrium selection parameters for static, discrete games.
2. Aggregating outcomes – e.g. Bresnahan and Reiss (1991) note that in the case of a two firm entry game, the model uniquely predicts the number of firms that will enter but not their identities. Such an approach has its costs when considering male and female volunteering decisions since ignoring issues with the heterogeneity of the male and female, we would only be able to determine the number volunteering rather than their identities which is clearly of interest.
3. Since this is an example of an incomplete econometric model, Manski and Tamer (2002) suggest using a bounds approach to estimation. In this case bounds on the parameters are derived from the necessary conditions for pure strategy Nash equilibria. This approach involves non-standard estimation techniques and has been discussed in detail by Andrews et. al. (2004).
4. Relaxing the assumption that the individuals move simultaneously. This can be done by allowing sequential entry with pre-determined orders or to allow only the most efficient player to move first.

In this study the first approach will be followed. The major benefit of imposing such an equilibrium selection rule is that standard maximum likelihood methods can be used meaning the models can be estimated in *Stata.* However, more recent literature focuses on approaches which avoid imposing arbitrary equilibrium selection assumptions since the effects of imposing an incorrect equilibrium selection rule are somewhat unclear. To provide some guidance regarding the likely effects of incorrectly imposing an equilibrium selection rule I intend to use a sensitivity analysis to better understand the effects on the parameter estimates. Depending on the results of this sensitivity analysis, the probability distribution used in the final estimation may need to be justified in some way rather than simply chosen arbitrarily.

In addition to these non-cooperative approaches to modelling household decisions, there is also a substantial literature on modelling household decisions cooperatively (see Vermeulen (2002)) that provides a useful alternative way to view interactions within the household. Due to the fact many of the models in this literature are not appropriate for modelling discrete choices of each of the partners there is a somewhat limited selection of cooperative models which could be applied. Kooreman (1994) provides a possible cooperative model which adjusts the Nash equilibrium concept used by Bjorn and Vuong (1984) to ensure that the outcome is always Pareto optimal (this is clearly necessary for a cooperative model), which is described as the Pareto Nash model. While the full details of the equilibrium concept are explained in Section 4.5, to provide some insight we consider the case where the payoffs defined by the latent variables result in the couple facing a prisoners' dilemma game. The Pareto Nash equilibrium concept would choose the "cooperative" outcome where both players are not playing their best response to the other's strategy rather than the Nash equilibrium in pure strategies.

As a result of considering the bivariate probit model along with cooperative and non-cooperative models for household decision making, we have several competing models of household choice. While the bivariate probit model is nested within the cooperative and non-cooperative models, the cooperative and non-cooperative models are non-nested. In this context, it is important to consider methods of non-nested model selection to determine the most appropriate model and as a result possibly provide some indication of the relevance of different models of household decision making. Pollak and Wales (1991) provide an approach to this known as the Likelihood Dominance Criterion which will be utilised in this thesis.

An alternative approach offered by Jia (2005) takes into account the possibility that different households interact in different ways by allowing the true model to be a mixture of some of the competing models considered. However there are important differences between the models used in Jia's (2005) approach and the data used, and the approach I intend to take and the available data. In particular, Jia (2005) used a much larger data set with substantially fewer explanatory variables. As such, implementing such an approach in this thesis may require some changes to the modelling strategy.

2.3 Alternative modelling approaches

In addition to the cooperative and non-cooperative approaches to modelling couple decision making discussed in Section 2.2, there is a third approach which assumes that couples are a unitary decision making unit. Such an approach was used by Jia (2005), which adopts a simplistic cooperative model that is equivalent to the unitary model of the household and estimates it using a conditional logit regression. Adopting this approach is problematic for several reasons. Firstly, the available data is not appropriate for a conditional logit regression since the independent variables do not vary over choices, only individuals. Secondly, a conditional logit regression assumes independent and identically distributed error terms. This means we are unable to capture correlated effects which are a potentially significant source of the observed correlation in volunteering decisions.

This leads to two possible approaches to implementing a unitary model for volunteering decisions. Initially, both approaches involve generating an outcome variable for whether neither volunteers, only the female volunteers, only the male volunteers or both volunteer. Then, using multinomial logit or multinomial probit with this outcome variable implies there is a single decision maker within the household and as such is in line with a unitary model. Unfortunately, using multinomial probit in this situation is difficult due to the fact there are no clearly appropriate exclusion restrictions on the regressors. As discussed by Keane (1992), identification is fragile without appropriate exclusion restrictions on the regressors.

While a multinomial logit model could be estimated, it suffers from the same problem as a conditional logit regression. That is, it assumes independent and identically distributed error terms implying that it is unable to capture correlated effects which are a potentially important source of the observed correlation in volunteering decisions. In addition, as a result of assuming independent and identically distributed error terms it assumes independence of irrelevant alternatives (IIA). This seems unlikely to hold in this case since for instance if a couple where both partners are volunteering changed status, it appears substantially more likely that they would change to a situation where one partner volunteered rather than neither volunteering. While some initial exploratory work was carried out using a multinomial logit model, due to these significant limitations it was not seen as an appropriate model for volunteering decisions.

Finally, if we simply wished to obtain an estimate of the impact of one partner's volunteering decision on the others rather than modelling the decision making structure, a possible approach to this may be using instrumental variables. This aims to correct for the endogeneity of the partner's volunteering decision. Unfortunately, it appears unlikely that an appropriate instrumental variable could be found within HILDA since very limited information is available on an individual's parents (unless they still live in the same household), which is the most likely source of an appropriate instrumental variable.

3 Data

3.1 The HILDA Survey

3.1.1 Summary of characteristics of the HILDA Survey

The data used in this study was obtained from Wave 6 of the Household, Income and Labour Dynamics in Australia (HILDA) Survey. A summary of the relevant information on HILDA provided by Watson (2008) follows. The HILDA Survey began with Wave 1 in 2001 which contained a large national probability sample of Australian households occupying private dwellings. In total, 19 914 individuals from 7682 households were interviewed in Wave 1. Each subsequent wave attempts to interview all members of the households providing at least one interview in Wave 1 and as such HILDA is a panel data set. These waves are conducted at yearly intervals with most of the fieldwork being completed in spring of each year. The data for Wave 6 was collected in 2006.

In addition, the sample includes any new household members resulting from changes in the composition of the original households. Individuals who join a household after Wave 1 are only interviewed in subsequent waves if they remain in the household or become Continuing Sample Members (CSMs). CSMs are defined to include all members of wave 1 households, any children subsequently born to or adopted by CSMs and all new entrants to a household who have a child with a CSM.

In Wave 1, the HILDA survey used four instruments which were the Household Form, the Household Questionnaire, the Person Questionnaire and the Self-Completion Questionnaire. In subsequent waves, the Person Questionnaire (PQ) was replaced with the Continuing Person Questionnaire for people who have been interviewed in a previous wave and the New Person

Questionnaire for people who have never been interviewed before. A description of each of these instruments is provided below:

- Household Form (HF) – records basic information about the composition of the household immediately after making contact
- Household Questionnaire (HQ) – this collects general information about the household mainly covering childcare arrangements, housing, household spending (until wave 5) and in waves 2 and 6, household wealth. While the intention is for this part of the survey to be administered to only one member of the household, in practice often more than one person is present at the interview or different parts of the survey are administered to different people.
- Continuing Person Questionnaire (CPQ) – this collects a wide variety of information on individual characteristics such as employment, income and health. It is administered to every member of the household aged 15 years and over who has previously completed a person questionnaire with parental consent being sought before interviewing persons aged under 18 years.
- New Person Questionnaire (NPQ) – this differs from the CPQ in that additional biographical history questions are included which only need to be asked once. These questions were included in the Person Questionnaire for those interviewed in Wave 1.
- Self-Completion Questionnaire (SCQ) – this contains mainly attitudinal questions which respondents may feel slightly uncomfortable answering in a face-to-face interview and as such respondents complete the questionnaire individually and return it later to the interviewer. This questionnaire is given to all individuals who completed a person questionnaire (i.e. the CPQ, NPQ or PQ).

3.1.2 Key features of the dataset relevant for the study

The key characteristic of HILDA which is relevant for the analysis in this study is that HILDA uses the household as the sampling unit, attempts to interview all members of the household and records the relationships within the household. As such for all those individuals in the dataset who have a partner, it is possible to identify their partner within the dataset and obtain their partner's survey responses (assuming that both of them also complete the person questionnaires). This differentiates the dataset from other datasets containing information on volunteering such as the Australian Bureau of Statistics' Voluntary Work Survey, 2000, (only the 2000 release is available as a Confidentialised Unit Record File) which only contains data on a representative individual from the household (ABS, 2001). Nevertheless, this other dataset is still useful for comparing rates of volunteering in HILDA to those estimated by the ABS.

While one of the key features of HILDA is that it is a panel data set, this study will not use the panel structure of the data and instead focuses on a single cross-section for several reasons. Firstly, several of the models I intend to consider have not been implemented for panel data previously and doing so is likely to be non-trivial. While an alternative may be to consider changes in volunteering over time (for instance changes when individuals find a partner) which would not require altering the models, the data in HILDA is unlikely to be sufficient for this purpose. In the case of changes when an individual finds a partner, data is not available in HILDA on their partner prior to the partner entering the household. More generally, for the majority of individuals surveyed there is usually no change in volunteering status between waves and as such there is likely to be little variation in the dependent variable when considering changes in volunteering status. Finally, several variables which are relevant for the analysis are not available in each wave.

While data on volunteering behaviour is available in all waves of HILDA, wave 6 of HILDA will be used for the analysis for several reasons. Firstly, Wave 6 of HILDA provides an additional question on volunteering behaviour which is not provided in previous waves. Given the apparent sensitivity of responses to questions on volunteering to the way the question is asked, this will be used for comparison with the other data on volunteering behaviour.

In addition, Waves 5 through 7 provide data on who within the household is responsible for a variety of decisions. Similar data is used in the context of monetary donations. Andreoni, et. al. (2003) considers the relationship between who within the household is primarily responsible for giving decisions and donations by the household. While the data in HILDA is not identical, it may still be useful for further work which considers differences between couples where both individuals make the same volunteering decision (i.e. both don't volunteer/both volunteer) and make the opposite volunteering decision (i.e. one volunteers while the other doesn't).

Finally, within the special topics section, Wave 6 of HILDA contains questions on wealth which a priori would seem to be an important determinant of volunteering choices. For these reasons and the fact Wave 6 is one of the most recent available, the analysis will be undertaken using Wave 6 of HILDA.

3.2 Sample selection

Wave 6 of HILDA contains data on 12 905 responding individuals from 6538 fully responding households and 601 partially responding households. This sample was obtained by approaching a total of 9351 households. To obtain the sample of 3255 couples used for estimating the models, the following steps were carried out:

1. Remove individuals who do not have a partner or data on their partner is missing (5447 individual level observations, leaving a total of 3729 couples)
2. Remove same-sex couples (41 observations consisting of 14 male couples and 27 female couples)
3. Remove couples who are part of multi-family households (76 observations)
4. Remove couples where either partner did not complete the Self-Completion Questionnaire (SCQ) since the question the dependent variable is generated from is in the SCQ (305 observations)
5. Remove couples where either partner gave an implausible value for the number of hours volunteered (19 observations)
6. Remove couples where either partner did not know or gave an implausible value for hours worked (2 observations)
7. Remove couples where either partner was required by Centrelink or a Job Network provider to do volunteer work (9 observations)
8. Remove couples where either partner stated that they are not looking for work because they do voluntary/unpaid work to avoid possible endogeneity problems (5 observations)
9. Remove couples where either partner stated that they are separated or never married and not de facto in the person questionnaire (8 observations)
10. Remove couples where either partner stated they are married, but living with spouse less than half the time (9 observations)

3.3 Treatment of missing data

As detailed in Section 3.2, only two observations were removed from the dataset due to missing data in the explanatory variables. In the other cases the missing data was handled as follows:

- For the income and wealth variables the imputed values were used.
- If the missing data occurred in a continuous variable, the continuous variable was set to zero and a dummy variable was generated which was 1 if the variable was missing and 0 otherwise. Information on these variables is available in Section 9.1, with the frequencies of missing observations being available in Table 3-6 and Table 3-7.
- If the missing data occurred in a discrete variable with more than two values which was used to generate a set of dummy variables, a dummy variable was generated which was 1 if the variable was missing and 0 otherwise. Information on these variables is available in Section 9.1.
- If the missing data occurred in a variable which was used to generate a dummy variable, the dummy variable was set to zero when the data was missing. Table 3-1 summarises the number of observations affected.

Table 3-1 – Missing data in dummy variables

Dummy variable	HILDA Variable used to generate	No. of missing observations and source
arrived_less_than_10y	fanyoa	1 – Don't know
part_arrived_less_than_10y	part_fanyoa	2 – Refused/Not stated
aboriginal_torres	fanatsi	1 – Refused/Not stated
part_irregular_hours	part_fjbmsch	1 – Refused/Not stated
retire_in_5y	frtiage1	1 – Refused/Not stated 69 – Don't know
part_retire_in_5y	part_frtiage1	81 – Don't know
worried_about_job	fjomwf	15 – Refused/Not stated
part_worried_about_job	part_fjomwf	2 – Multiple response on SCQ 5 – Refused/Not stated
child_under_10	ftcyng	5 – Don't know
pressed_for_time	flsrush	20 – Refused/Not stated
part_pressed_for_time	part_flsrush	28 – Refused/Not stated
religion	flsnwser	2 – Multiple response on SCQ 44 – Refused/Not stated
part_religion	part_flsnwser	30 – Refused/Not stated

3.4 Dependent variable

3.4.1 Summary statistics and Definition

The dependent variable is derived from the time-use question (B25) in the Self Completion Questionnaire. This question asks individuals how much time they spend on a variety of activities in a typical week with part h) being "Volunteer or charity work (for example, canteen work at the local school, unpaid work for a community club or organisation)". To obtain a dummy variable for whether each partner volunteers, this was recoded to 1 if the individual undertakes one or more hours of volunteer or charity work and 0 otherwise. Table 3-2 shows a cross-tabulation of the dependent variable for the 3255 couples in the sample which illustrates the correlation between each couple's volunteering decisions noted by Freeman (1997). We can see that there are substantially more couples where both volunteer (i.e. (1,1) outcomes) than would be expected under independence. In particular, the relative frequency of (1,1) is $\frac{301}{3255}=0.092$ compared with an expected relative frequency under independence of $\frac{628}{3255}\times\frac{793}{3255}=0.047$.

Table 3-2 - Volunteering outcomes for males/females

Male \ Female	0	1	Total
0	2135	492	2627
1	327	301	628
Total	2462	793	3255

The second question on volunteering behaviour in Wave 6 asks "In general, how often do you do the following things: i) Volunteer your spare time to work on boards or organising committees of clubs, community groups or other non-profit organizations?" with the responses being never, rarely, occasionally, sometimes, often and very often. Given the sensitivity of responses to questions on volunteering to the way the question is asked noted by Freeman (1997), it is worthwhile comparing responses to this question to the dependent variable derived previously.

Table 3-3 provides cross-tabulations for this variable and the dependent variable for male and female partners and it is clear that there is a strong correlation between the two variables. However, there are two key discrepancies which should be considered. Firstly, a number of individuals with a value of 0 for the dependent variable have values of often or very often for the other measure of volunteering. This is possibly due to the activity these individuals participate in not being held on a weekly basis. Secondly, a number of individuals with a value of 1 for the dependent variable have values of never or rarely for the other measure of volunteering. This could be attributed to these individuals volunteering time without being on the board or organising committee. As such, it appears that the two measures are relatively consistent after considering differences in what they are measuring.

Table 3-3 - Comparison of volunteering variables

Male partners				Female partners			
	0	**1**	**Total**		**0**	**1**	**Total**
Multiple response	2	1	3	Multiple response	1	0	1
Refused/Not stated	42	4	46	Refused/Not stated	27	6	33
Never	1386	34	1420	Never	1337	58	1395
Rarely	680	76	756	Rarely	620	101	721
Occasionally	212	58	270	Occasionally	230	91	321
Sometimes	175	125	300	Sometimes	166	153	319
Often	100	205	305	Often	66	231	297
Very Often	30	125	155	Very Often	15	153	168
Total	2627	628	3255	Total	2462	793	3255

The volunteering rate indicated by the dependent variable can also be compared with that estimated by the ABS in the 2006 Voluntary Work Survey (ABS, 2007). The ABS estimates that for husbands, wives or partners throughout Australia, 34.9% of males and 39.6% of females volunteered over the previous twelve months. This is substantially higher than the rate of volunteering indicated by the dependent variable (19.3% for males and 24.4% for females), which is expected since the dependent variable adopts a more narrow definition of volunteering. In addition, the Voluntary Work Survey focuses on volunteering activities and includes a substantial preamble to define and explain the concept of volunteering, while volunteering is only addressed briefly in HILDA. Freeman (1997) argues that this focus on charitable activity may lead people to remember more fully their volunteering or label actions they might otherwise have seen in a different light as volunteering, thus leading to higher measured rates of volunteering.

3.4.2 Changes in status across waves

To better understand the nature of volunteering decisions it is useful to consider changes in a couple's volunteering status over time. Since the question used to generate the dummy variable for volunteering status is present in Waves 1 through 6 of the HILDA Survey, it is possible to use the panel nature of the HILDA dataset to ascertain the volunteering status for many of the couples in previous periods. Table 3-4 and Table 3-5 show the changes in volunteering status from Waves 5 to 6 and 1 to 6 respectively. The Not in sample category refers to those observations which were in the sample of 3255 couples from Wave 6 which was used for estimation, but data for either partner's volunteering status was not available in the previous wave. There are a variety of reasons for why this may be the case including the couple forming between the waves, refusal to answer the volunteering question in the previous wave or failure to complete the Self-Completion Questionnaire in the previous wave.

Table 3-4 – Changes in volunteering status from Wave 5 to Wave 6

Wave 5 \ Wave 6	**Neither volunteers**	**Male volunteers**	**Female volunteers**	**Both volunteer**	**Total**
Not in sample	331	37	57	18	443
Neither volunteers	1,570	105	137	35	1,847
Male volunteers	82	131	22	42	277
Female volunteers	133	17	237	51	438
Both volunteer	19	37	39	155	250
Total	2,135	327	492	301	3,255

Table 3-5 – Changes in volunteering status from Wave 1 to Wave 6

Wave 1 \ Wave 6	Neither volunteers	Male volunteers	Female volunteers	Both volunteer	Total
Not in sample	784	85	116	50	1,035
Neither volunteers	988	81	134	55	1,258
Male volunteers	113	84	29	33	259
Female volunteers	169	24	153	55	401
Both volunteer	81	53	60	108	302
Total	2,135	327	492	301	3,255

Table 3-4 indicates that there is a reasonable amount of persistence in volunteering decisions from Wave 5 to Wave 6 with 2093 of the 2812 (74.4%) couples for which data was available in both waves having the same volunteering status in both waves. In addition, of the couples that do change volunteering status, in most cases the volunteering status of only one partner changes (626 couples) rather than both partners changing status (93 couples). It also appears that there may be a small amount of correlation between volunteering and the likelihood of not being in the sample in the previous wave with 15.5% of those couples where neither volunteered in Wave 6 not being in the sample in Wave 5 and 6.0% of those couples where both volunteered in Wave 6 not being in the sample in Wave 5.

Table 3-5 shows that there is less persistence in volunteering status between Waves 1 and 6, which is to be expected given the longer period of time between the waves. Nevertheless, 1333 of the 2220 (60.0%) couples for which data was available in both waves had the same volunteering status in both waves and of those that did change volunteering status, in most cases the volunteering status of only one partner changes.

This indicates a potential issue with the modelling approach used in this thesis since it does not capture the effect of persistence in volunteering decisions over time. It is clear though that the approach used here is a first step towards such a model where a couple's previous volunteering decisions affect their current volunteering decision. The modelling structure that has been used is not necessarily inconsistent with the observed persistence since this may simply be a result of many of the factors affecting volunteering decisions remaining constant across periods. This would imply that most couples maintain their volunteering status without any direct impact of previous volunteering decisions. As such, more complex behavioural structures involving repeated interactions between the partners rather than a one shot simultaneous Nash or Stackelberg leader game may not be a more accurate representation of the true decision making structure. Instead, the primary advantage may be an ability to exploit the panel structure of the dataset to eliminate couple specific effects.

3.5 Key explanatory variables

Given the nature of the study, there are a wide variety of plausible explanatory variables available in HILDA which could be included as controls, many of which have been identified in previous studies. There is only one major exception in terms of variables which have previously been identified in the literature as important but are not available. This is that no data is available on whether an individual was asked to volunteer which Freeman (1997) identifies as an important determinant of volunteering behaviour.

Section 9.1 in the appendix details the variables that will be used in the study and their definitions. Since many variables relate to an individual rather than a couple, prefixes are used to distinguish between the male and female's characteristics (in some cases the letters I and H are used to identify whether the variable is measured at the individual (I) or household (H) level). The prefix "part_" is placed in front of the variable name to indicate it refers to the female partner's characteristics. If the variable does not have a prefix it refers to the male partner's characteristics or a household characteristic. The base category for each set of dummy variables is provided in the table.

While the rationale for using many of these variables such as age, demographic characteristics, education level, health, hours worked and number of children is clear, the motivation for including a small number of them may not be clear. The variables arrived_less_than_10y, yr_curr_addr, owns_home and likely_to_move are intended to aid in capturing an individual's level of integration with their local community which is likely to have an impact on whether or not they volunteer. The variables fixedterm_casual, irregular_hours, more_than_one_job, office_worker, satisfied_balance_work and worried_about_job are intended to capture characteristics of an individual's employment which may be relevant in determining whether or not they volunteer (e.g. if their hours are irregular they may have difficulty committing to a volunteering activity). Some of these are motivated by Polidano, et. al. (2009) who find a positive relationship between volunteering and being in labour market transition. Finally, pressed_for_time is included since it may have an impact on an individual's ability to commit to a volunteering activity and retire_in_5y is included on the basis that being close to retirement may mean an individual takes on additional voluntary activities as they transition into retirement.

3.6 Summary statistics

Means for the dummy explanatory variables are provided in Table 3-6 (these are all measured at the individual level), while the mean, standard deviation, minimum and maximum are reported for the continuous explanatory variables in Table 3-7. The results in Table 3-6 indicate that a substantial proportion of the sample is born overseas (around 20%), however very few are recent arrivals (around 2-3% arrived in the 10 years prior to the survey). In addition, the majority of couples live in major urban areas and around 25% of individuals have a postgraduate or bachelors level qualification with slightly more females having one of the two qualifications than males. Finally, a large proportion of the sample owns their own home (over 80%) and has self-assessed health status of good, very good or excellent.

Table 3-6 – Summary statistics for dummy explanatory variables

Variable	Male Mean	Female Mean	Variable	Male Mean	Female Mean
overseasborn_english	0.1263	0.1054	irregular_hours	0.1551	0.1244
overseasborn_other	0.1084	0.1177	fixedterm_casual	0.1143	0.1677
arrived_less_than_10y	0.0206	0.0283	satisfied_balance_work	0.6003	0.5020
aboriginal_torres	0.0123	0.0123	retire_in_5y	0.0578	0.0535
state_vic	0.2498	0.2498	office_worker	0.4737	0.5475
state_qld	0.2123	0.2123	unemployed	0.0172	0.0181
state_sa	0.0965	0.0965	retired	0.2089	0.2452
state_wa	0.1011	0.1011	worried_about_job	0.1579	0.0989
state_tas	0.0326	0.0326	child_under_10	0.2876	0.2866
state_nt	0.0061	0.0061	long_term_hlth_cond	0.1828	0.1662
state_act	0.0197	0.0197	health_excellent	0.0952	0.1041
other_urban	0.2335	0.2335	health_verygood	0.3318	0.3733
bounded_locality	0.0316	0.0316	health_good	0.3662	0.3389
rural_balance	0.1515	0.1515	health_fair	0.1413	0.1186
postgrad	0.1051	0.0992	health_poor	0.0326	0.0289
bachelor	0.1339	0.1521	owns_home	0.8575	0.8154
post_school	0.4181	0.2544	pressed_for_time	0.3539	0.4452
yr12	0.0962	0.1398	likely_to_move	0.1195	0.1183
hours_vary	0.0243	0.0200	religion	0.1330	0.1579
more_than_one_job	0.0513	0.0562			

The results in Table 3-7 indicate that the males in the sample are on average slightly older than the females and work substantially more hours than females (the average for males is close to double that for females). In addition, males have higher individual disposable incomes than females in the sample.

Table 3-8 presents the mean and standard deviation for selected continuous explanatory variables and the mean for selected dummy explanatory variables across groups of observations. Numbers in brackets in the first row refer to whether the male volunteers and female volunteers respectively. i.e. (1,0) implies the male volunteers while the female does not. Table 3-8 indicates that there is substantial variation in some of the explanatory variables across the four categories.

Table 3-7 – Summary statistics for continuous variables

Variable	Mean	Std Dev	Min	Max
age	49.15	15.55	16	93
part_age	46.68	15.20	17	93
hours_worked	32.01	22.84	0	100
part_hours_worked	18.59	18.89	0	112
ind_disp_income	44.18	37.64	-86.9	375.639
part_ind_disp_income	26.47	23.76	-61	375.639
num_child	0.9447	1.1728	0	8
part_num_child	1.0277	1.2127	0	8
yr_curr_addr	10.55	11.82	0	80.35
part_yr_curr_addr	10.25	11.40	0	70.56
yr_curr_addr_missing	0.0061			
part_yr_curr_addr_missing	0.0065			
hh_gross_inc	92.84	76.15	-86.9	842.344
hh_home_equity	302.51	417.94	-450	4426.46
hh_net_worth	806.33	1369.06	-918.7	12798.92

These results suggest a number of relationships in the data. In particular, volunteers are generally older and more highly educated with a greater proportion of volunteers holding postgraduate degrees. As would be expected, hours worked appears to have a negative relationship with volunteering with volunteers generally working fewer hours than non-volunteers.

In addition, volunteering couples (those where both volunteer) have a greater proportion of the males being office workers than non-volunteering couples (the opposite is true for the females). Volunteers are more likely to be retired, more likely to attend religious services and less likely to have young children (defined as under 10). Finally, volunteers generally have higher net worth.

Table 3-8 – Selected summary statistics across groups of observations

Sample	(1,1)	(1,0)	(0,1)	(0,0)
Observations	301	327	492	2135
Variable	**Mean (Std Dev)**	**Mean (Std Dev)**	**Mean (Std Dev)**	**Mean (Std Dev)**
age	53.41 (13.53)	50.77 (14.83)	50.10 (13.17)	48.08 (16.28)
postgrad	0.1761	0.1682	0.1220	0.0815
bachelor	0.1561	0.1437	0.1545	0.1246
post_school	0.4120	0.4557	0.4370	0.4089
yr12	0.0764	0.0703	0.0711	0.1087
hours_worked	29.87 (24.58)	28.76 (23.16)	35.14 (22.47)	32.09 (22.54)
irregular_hours	0.1894	0.1529	0.1626	0.1489
office_worker	0.5548	0.5076	0.5061	0.4496
retired	0.2625	0.2477	0.1748	0.2033
child_under_10	0.2292	0.2813	0.3374	0.2852
num_child	1.1993 (1.4469)	1.1070 (1.2838)	1.2012 (1.2352)	0.8248 (1.076)
long_term_hlth_cond	0.2359	0.1927	0.1707	0.1766
likely_to_move	0.1063	0.0795	0.1037	0.1311
religion	0.3422	0.2202	0.1626	0.0834
part_age	50.87 (13.31)	48.28 (14.58)	47.67 (12.84)	45.62 (15.90)
part_postgrad	0.1229	0.1284	0.1199	0.0867
part_bachelor	0.1595	0.1193	0.1606	0.1541
part_post_school	0.2625	0.2385	0.2581	0.2548
part_yr12	0.0930	0.1498	0.1484	0.1429
part_hours_worked	13.85 (16.99)	18.72 (18.35)	16.06 (17.51)	19.83 (19.37)
part_irregular_hours	0.1395	0.1407	0.1585	0.1119
part_office_worker	0.4950	0.5780	0.5528	0.5489
part_retired	0.3355	0.2599	0.2175	0.2365
part_child_under_10	0.2392	0.2936	0.3374	0.2806
part_num_child	1.2525 (1.4547)	1.1927 (1.3394)	1.3008 (1.2584)	0.9077 (1.1242)
part_owns_home	0.8970	0.8654	0.8699	0.7836
part_likely_to_move	0.0864	0.0979	0.0996	0.1302
part_religion	0.3854	0.2141	0.2337	0.0998
hh_home_equity	382.97 (505.70)	353.52 (390.92)	396.15 (559.43)	261.78 (360.64)
hh_net_worth	1117.98 (1763.39)	952.97 (1447.11)	1055.53 (1627.41)	682.50 (1202.07)

4 Modelling approach

4.1 Modelling Volunteering Behaviour

This thesis will focus on the consumption motive for volunteering rather than the investment motive due to the fact estimated wage premiums for volunteers are small and as a result evidence for the investment model is limited (see Day and Devlin (1998) and Prouteau and Wolff (2006)). Under the consumption model, utility is assumed to depend directly or indirectly on hours volunteered (as discussed in Section 2.1 there are three possible motivations for this). As a result, following Freeman (1997) individuals are assumed to maximize utility dependent on goods (G), leisure (L) and charity (C) subject to a budget constraint, where volunteer time (T_v) and donations (D) are the two inputs used to produce charity. More formally, individuals solve the following problem:

$$\max\ U(G, L, C)$$

$$\text{subject to}\ \begin{array}{c} C = C(T_v, D) \\ G + D = WT_w + Y\ \text{(income constraint)} \\ T_w + T_v + L = 1\ \text{(time constraint)} \end{array}$$

where W is wages, T_w is time worked, Y is non-wage income and D is charitable donations. In this problem the exogenous variables are W and Y and as such, the solution to this problem will yield a derived demand for volunteer time as a function of W and Y. This solution will depend on the functional forms chosen for U and C. For simplicity, we follow Freeman (1997) and write the solution in linear form as:

$$T_v = a + bW + cY + v$$

In addition, the functional form of the utility function may differ across individuals depending on characteristics such as culture, education and family characteristics. This provides a justification for adding these characteristics to the previous equation. Although substantially more structure must be imposed in the previous problem to obtain the regression models that will be used in this thesis, it is intended to provide some justification for the form of the regressions being used.

Also, while this model predicts that higher wage workers should volunteer less since the productivity of volunteer time is the same for all workers, and that there should be substitution of donations for volunteering as wages rise, these predictions are a result of the charitable production function. Freeman (1997) points out that by using $C = C(WT_v, D)$ instead, the increased productivity in volunteering for higher wage workers can offset the increased opportunity cost of time as measured through the wage rate.

As a result, the consumption model does not provide any strong predictions for the sign of the coefficient *b* as it depends on positive income effects (individuals have greater demand for charity as income increases) and negative substitution effects (the opportunity cost of volunteering increases with the wage rate, which may be offset by increased productivity in volunteering). Coefficient c measures the income effect of charitable activity and as such may be negative or positive depending on the nature of charitable activity as a good. This highlights the difficulty in testing the consumption model empirically since the model does not provide testable predictions for parameter values (see Prouteau and Wolff (2006)).

In addition, it may be the case that seemingly endogenous variables in the maximization problem should be included in the linear equation for volunteer time. As Brown and Lankford (1992) discuss, whether W measures the worker's opportunity cost of time depends on whether hours of work are fully flexible. If in the previous maximization problem, workers are forced to solve it sequentially by committing to a paid job with fixed hours before allocating time to volunteering, the wage rate will not measure the opportunity cost of time. Instead, hours worked becomes the theoretically relevant control variable and should be included in the model instead. This issue is related to the issue of endogeneity discussed in Section 4.2 and will be discussed in Section 6.1.

4.2 Possible endogeneity of hours worked and income

There is a possibility that some of the explanatory variables being used are not exogenous to the volunteering decision. The variables which are most likely to cause concern in this regard are income and hours worked. While this endogeneity will not be accounted for in the models used in this thesis, the purpose of this section is to highlight the likely severity of this issue along with drawing attention to it as a possible issue with the estimation results.

Under the investment model of volunteering, income will clearly be endogenous since volunteering increases income. Given the limited evidence for the investment model, this in itself is not a major cause for concern. In addition few papers in the literature have attempted to control for the endogeneity of income (Day and Devlin (1996) is one exception where determinants of income were included in the regression rather than income itself).

Whether hours worked is endogenous essentially depends on how free an individual is to set their hours worked. As Brown and Lankford (1992) discuss there are a variety of theoretical models which imply that hours are constrained. In the context of volunteering, a worker may have to commit to a remunerative job with fixed hours before committing time to volunteering which implies that hours worked is exogenous to the volunteering decision. If on the other hand hours are unconstrained then the time allocated to working may be decided jointly with time allocated to volunteering, implying that hours worked is endogenous. Anecdotal evidence suggests the nature of most employment contracts in Australia is such that the prior situation is more likely to be correct. In addition, there is some evidence in HILDA to suggest that volunteering has little impact on hours worked since very few individuals indicated they are not looking for work because they do voluntary/unpaid work. As discussed in Section 3.2, these observations were removed from the sample.

4.3 Bivariate probit

For comparison with the non-cooperative choice models which are a key component of this thesis, a bivariate probit model will be estimated. This model does not attempt to capture any endogenous effects of volunteering behavior and as such takes the form in Equation (4-1) where y_1 and y_2 are dummy variables for whether the male and female volunteer and x_1 and x_2 are observable factors we wish to control for. This is motivated by Equation (2-1).

$$\begin{cases} y_1^* = x'_1\beta_1 + \varepsilon_1 \\ y_2^* = x'_2\beta_2 + \varepsilon_2 \\ y_i = 1 \; if \; y_i^* > 0, 0 \; otherwise \end{cases} \tag{4-1}$$

4.4 Non-cooperative choice models

This section details how assuming the observed outcomes are generated by a Nash or Stackelberg leader game played between the partners can be used to link the latent variables to the observed outcomes. It also defines exactly what is meant by these concepts.

4.4.1 Stackelberg leader game – Definition and Estimation

The Stackelberg leader game assumes that the observed outcome is the Stackelberg equilibrium of a Stackelberg leader game played between the partners with payoffs given by the latent variables. This section follows the approach of Bjorn and Vuong (1985) in developing the Stackelberg leader model to illustrate how it is implemented. We will develop the model for the case where the male is the Stackelberg leader and the female is the follower and as such index the players by i=m,f (corresponding to the male and female respectively) rather than i=1,2. The case where the female is the Stackelberg leader and the male is the follower is developed in a similar manner and as such we can simply swap the subscripts to obtain the relevant likelihood function for this case.

In making a decision whether to take action 0 or 1, player m takes into account player f's payoffs (since this is a complete information game, these are known to player m). Therefore, player m chooses action j=0,1 such that player f's action conditional on j gives player m the greatest possible payoff. The possible reaction functions for player f along with the conditions for them to hold are given in Table 4-1 (note that $U^f\left(y_m, y_f\right)$ denotes the female's utility when the male takes action y_m and the female takes action y_f).

Table 4-1 - Player f's reaction functions

F_1	Always take action 1	$U^f(y_m,1) - U^f(y_m,0) \geq 0 \; for \; y_m = 0 \; or \; 1$
F_2	Always take the same action as player m	$U^f(y_m,1) - U^f(y_m,0) \geq 0 \; for \; y_m = 1$ & $U^f(y_m,1) - U^f(y_m,0) < 0 \; for \; y_m = 0$
F_3	Always take action 0	$U^f(y_m,1) - U^f(y_m,0) < 0 \; for \; y_m = 0 \; or \; 1$
F_4	Always take the opposite action to player m	$U^f(y_m,1) - U^f(y_m,0) < 0 \; for \; y_m = 1$ & $U^f(y_m,1) - U^f(y_m,0) \geq 0 \; for \; y_m = 0$

The utility comparison player m makes in determining their action depends on the reaction function of player f. The conditions arising from these comparisons (or outcomes of these comparisons) are detailed in Table 4-2. Note that we let $\overline{M_i}$ denote the negation of M_i and $U^m\left(y_m, y_f\right)$ refers to the male's utility when the male takes action y_m and the female takes action y_f.

Table 4-2 - Player m's utility comparisons

Reaction function	Utility Condition for player m	Condition Label
F_1	$U^m(1,1) \geq U^m(0,1)$	M_1
F_2	$U^m(1,1) \geq U^m(0,0)$	M_2
F_3	$U^m(1,0) \geq U^m(0,0)$	M_3
F_4	$U^m(1,0) \geq U^m(0,1)$	M_4

Given the reaction function for player f and the outcome of player m's utility comparison, the Stackelberg equilibria of the game can be found. These are given in Table 1 of Bjorn & Vuong (1985) and are reproduced in Table 4-3.

Table 4-3 – Stackelberg equilibria

Pair	Outcome	Pair	Outcome
F_1 and M_1	(1,1)	F_3 and M_3	(1,0)
F_1 and $\overline{M_1}$	(0,1)	F_3 and $\overline{M_3}$	(0,0)
F_2 and M_2	(1,1)	F_4 and M_4	(1,0)
F_2 and $\overline{M_2}$	(0,0)	F_4 and $\overline{M_4}$	(0,1)

To estimate this model, we decompose the utilities into deterministic and random components, as shown in Equation (4-2).

$$\begin{gathered} U^m(1, y_f) = x_m'\beta_1^m + \alpha_1^m y_f + \varepsilon_1^m \\ U^m(0, y_f) = x_m'\beta_0^m + \alpha_0^m y_f + \varepsilon_0^m \\ U^f(y_m, 1) = x_f'\beta_1^f + \alpha_1^f y_m + \varepsilon_1^f \\ U^f(y_m, 0) = x_f'\beta_0^f + \alpha_0^f y_m + \varepsilon_0^f \end{gathered} \tag{4-2}$$

Since only differences in utilities are relevant in determining the outcome, we define $\varepsilon^i = \varepsilon_1^i - \varepsilon_0^i$ for i=m,f and assume that the pair (ε^m, ε^f) is normally distributed with zero means, unit variances and correlation ρ. As a result, each reaction function for player f will occur only if some conditions on the errors are satisfied and each utility condition for player m will hold only if certain conditions on the errors are satisfied. Therefore, the probability of each pair of reaction functions and utility conditions can be found in terms of the parameters in the model. From Table 4-3 we can find the probability of each pair of observed outcomes in terms of the probabilities of pairs of a reaction function and utility condition. Equation (4-3) provides these results.

$$\begin{gathered} \Pr(0,0) = \Pr(F_2 \text{ } and \text{ } \overline{M_2}) + \Pr(F_3 \text{ } and \text{ } \overline{M_3}) \\ \Pr(1,0) = \Pr(F_3 \text{ } and \text{ } M_3) + \Pr(F_4 \text{ } and \text{ } M_4) \\ \Pr(0,1) = \Pr(F_1 \text{ } and \text{ } \overline{M_1}) + \Pr(F_4 \text{ } and \text{ } \overline{M_4}) \\ \Pr(1,1) = \Pr(F_1 \text{ } and \text{ } M_1) + \Pr(F_2 \text{ } and \text{ } M_2) \end{gathered} \tag{4-3}$$

Using Table 4-1 and Equation (4-2) we can derive the conditions on the error terms for each of player f's reaction functions to hold which are detailed in Table 4-4.

Table 4-4 – Conditions on error term for player f's reaction functions

F_1	$\varepsilon^f > -x_f'(\beta_1^f - \beta_0^f) - \min(0, \alpha_1^f - \alpha_0^f)$
F_2	$-x_f'(\beta_1^f - \beta_0^f) - (\alpha_1^f - \alpha_0^f) < \varepsilon^f < -x_f'(\beta_1^f - \beta_0^f)$ and $(\alpha_1^f - \alpha_0^f) \geq 0$
F_3	$\varepsilon^f < -x_f'(\beta_1^f - \beta_0^f) - \max(0, \alpha_1^f - \alpha_0^f)$
F_4	$-x_f'(\beta_1^f - \beta_0^f) < \varepsilon^f < -x_f'(\beta_1^f - \beta_0^f) - (\alpha_1^f - \alpha_0^f)$ and $(\alpha_1^f - \alpha_0^f) < 0$

Using Table 4-2 and Equation (4-2) we can derive the conditions on the error term for each of player m's utility conditions to hold which are detailed in Table 4-5.

Table 4-5 – Conditions on error term corresponding to player m's utility conditions

M_1	$\varepsilon^m \geq -x_m'(\beta_1^m - \beta_0^m) - (\alpha_1^m - \alpha_0^m)$
M_2	$\varepsilon^m \geq -x_m'(\beta_1^m - \beta_0^m) - \alpha_1^m$
M_3	$\varepsilon^m \geq -x_m'(\beta_1^m - \beta_0^m)$
M_4	$\varepsilon^m \geq -x_m'(\beta_1^m - \beta_0^m) + \alpha_0^m$

Using Equation (4-3), Table 4-4 and Table 4-5, we can derive probability statements for each outcome in terms of the parameters. Using these, we can use maximum likelihood estimation to estimate the parameters of the model. We note that only the differences in some parameters can be identified since only their differences appear in Table 4-4 and Table 4-5 and therefore we define $\beta^m = (\beta_1^m - \beta_0^m)$, $\beta^f = (\beta_1^f - \beta_0^f)$, $\alpha^f = (\alpha_1^f - \alpha_0^f)$. The probability of each pair of observed outcomes in terms of the model parameters is given in Table 4-6.

.

Table 4-6 – Outcome probabilities in terms of model parameters

Pr (0,0) $= \Phi(-x_m'\beta^m, -x_f'\beta^f, \rho)$
$-\Omega(-x_m'\beta^m, -x_f'\beta^f, -x_m'\beta^m - \alpha_1^m, -x_f'\beta^f - \alpha^f, \rho)\ if\ \ \alpha^f \geq 0$
$= \Phi(-x_m'\beta^m, -x_f'\beta^f, \rho)\ otherwise$

Pr (1,0) $= \Phi(x_m'\beta^m, -x_f'\beta^f - \alpha^f, -\rho)\ if\ \ \alpha^f \geq 0$
$= \Phi(x_m'\beta^m, -x_f'\beta^f - \alpha^f, -\rho)$
$+\Omega(-x_m'\beta^m, -x_f'\beta^f - \alpha^f, -x_m'\beta^m + \alpha_0^m, -x_f'\beta^f, \rho)\ otherwise$

Pr (0,1) $= \Phi(-x_m'\beta^m - \alpha_1^m + \alpha_0^m, x_f'\beta^f, -\rho)\ if\ \ \alpha^f \geq 0$
$= \Phi(-x_m'\beta^m - \alpha_1^m + \alpha_0^m, x_f'\beta^f, -\rho)$
$+\Omega(-x_m'\beta^m + \alpha_0^m, -x_f'\beta^f - \alpha^f, -x_m'\beta^m - \alpha_1^m + \alpha_0^m, -x_f'\beta^f, \rho)\ otherwise$

Pr (1,1) $= \Phi(x_m'\beta^m + \alpha_1^m - \alpha_0^m, x_f'\beta^f + \alpha^f, \rho)$
$-\Omega(-x_m'\beta^m - \alpha_1^m, -x_f'\beta^f, -x_m'\beta^m - \alpha_1^m + \alpha_0^m, -x_f'\beta^f - \alpha^f, \rho)\ if\ \ \alpha^f \geq 0$
$= \Phi(x_m'\beta^m + \alpha_1^m - \alpha_0^m, x_f'\beta^f + \alpha^f, \rho)\ otherwise$

Where $\Omega(a, b, c, d, \rho)$ is the integral of the bivariate standard normal density over the region $a \geq \varepsilon^m \geq c,\ b \geq \varepsilon^f \geq d$ (Note: If c>a or d>b we swap the limits and take the negative of the integral)

and $\Phi(a, b, \rho)$ is the bivariate standard normal cumulative density function with correlation parameter ρ

4.4.2 Nash game – Definition and Estimation

The Nash game assumes that the observed outcome is the Nash equilibrium in pure strategies of a two player game with payoffs given by the latent variables. This section shows how this is implemented, following the approach of Bjorn and Vuong (1984). We begin by interpreting the latent variables as the utility differences specified in Equation (4-4). Denote the utility of agent i=m,f (corresponding to the male and female respectively) by $U^i(y_m, y_f)$.

$$\begin{aligned} y_m^* &= U^m(1, y_f) - U^m(0, y_f) \\ y_f^* &= U^f(y_m, 1) - U^f(y_m, 0) \end{aligned} \qquad (4\text{-}4)$$

Bjorn and Vuong (1984) show how to use the Nash equilibria of a game between the two agents to relate the latent variables to the observed discrete variables. In developing the model, we follow their approach of considering the reaction functions of the relevant players.

To do so we decompose $U^i(y_m, y_f)$ into a deterministic and a random component as shown in Equation (4-5). Table 4-7 gives the four possible reaction functions for player f along with the relevant conditions on the utilities. The possible reaction functions for player m are identical and we label these M_1 through M_4 corresponding to F_1 through F_4.

$$\begin{aligned} U^m(1, y_f) &= x_m'\beta_1^m + \alpha_1^m y_f + \varepsilon_1^m \\ U^m(0, y_f) &= x_m'\beta_0^m + \alpha_0^m y_f + \varepsilon_0^m \\ U^f(y_m, 1) &= x_f'\beta_1^f + \alpha_1^f y_m + \varepsilon_1^f \\ U^f(y_m, 0) &= x_f'\beta_0^f + \alpha_0^f y_m + \varepsilon_0^f \end{aligned} \tag{4-5}$$

Table 4-7 – Player f's reaction functions

Label	Description	Condition
F_1	Always take action 1	$U^f(y_m, 1) - U^f(y_m, 0) \geq 0 \; for \; y_m = 0 \; or \; 1$
F_2	Always take the same action as player m	$U^f(y_m, 1) - U^f(y_m, 0) \geq 0 \; for \; y_m = 1$ & $U^f(y_m, 1) - U^f(y_m, 0) < 0 \; for \; y_m = 0$
F_3	Always take action 0	$U^f(y_m, 1) - U^f(y_m, 0) < 0 \; for \; y_m = 0 \; or \; 1$
F_4	Always take the opposite action to player m	$U^f(y_m, 1) - U^f(y_m, 0) < 0 \; for \; y_m = 1$ & $U^f(y_m, 1) - U^f(y_m, 0) \geq 0 \; for \; y_m = 0$

In the Nash game, only the differences in the utilities given in Equation (4-5) are used to determine the outcomes. The relevant differences are given in Equation (4-4). As a result, we can only identify the following differences in the parameters in Equation (4-5) $\alpha^i = \alpha_1^i - \alpha_0^i, \beta^i = \beta_1^i - \beta_0^i$ for i=m,f.

In addition, we must impose an assumption on the distribution of the differences in the error terms in order to estimate the model. We assume that $\varepsilon^m = \varepsilon_1^m - \varepsilon_0^m$ and $\varepsilon^f = \varepsilon_1^f - \varepsilon_0^f$ are jointly normally distributed with zero means, unit variances and correlation ρ. Therefore, the expressions for y_m^* and y_f^* are given in Equation (4-6).

$$\begin{aligned} y_m^* &= x_m'\beta^m + \alpha^m y_f + \varepsilon^m \\ y_f^* &= x_f'\beta^f + \alpha^f y_m + \varepsilon^f \end{aligned} \tag{4-6}$$

The conditions for each reaction function will hold only if certain conditions on the error are satisfied. These conditions are detailed in Table 4-8 for player f's reaction functions and they are similar for player m. As a result of the distributional assumption placed on the error terms, these can be used to derive probability statements for each pair of reaction functions and therefore for the observed outcomes.

Table 4-8 – Conditions on error term for player f's reaction functions

F_1	$\varepsilon^f > -x_f'(\beta_1^f - \beta_0^f) - \min(0, \alpha_1^f - \alpha_0^f)$
F_2	$-x_f'(\beta_1^f - \beta_0^f) - (\alpha_1^f - \alpha_0^f) < \varepsilon^f < -x_f'(\beta_1^f - \beta_0^f)$ and $(\alpha_1^f - \alpha_0^f) \geq 0$
F_3	$\varepsilon^f < -x_f'(\beta_1^f - \beta_0^f) - \max(0, \alpha_1^f - \alpha_0^f)$
F_4	$-x_f'(\beta_1^f - \beta_0^f) < \varepsilon^f < -x_f'(\beta_1^f - \beta_0^f) - (\alpha_1^f - \alpha_0^f)$ and $(\alpha_1^f - \alpha_0^f) < 0$

We now use the assumption that the observed outcomes are Nash Equilibrium outcomes of a game played by the two agents to relate the latent variables to the observed outcomes. This is insufficient to define how the observed outcomes are generated without additional assumptions due to the fact a Nash Equilibrium in pure strategies may not exist or multiple Nash Equilibria in pure strategies may arise. Table 1 of Bjorn and Vuong (1984) gives the possible Nash Equilibria in pure strategies of the game given the two player's reaction functions and these are reproduced in Table 4-9.

The difficulty associated with non-existence and multiple Nash equilibria becomes evident in Table 4-9 since each pair of reaction functions does not necessarily imply a unique Nash equilibrium. As a result, we must introduce some additional assumptions to relate the probability of each pair of outcomes to the probability of pairs of reaction functions occurring. At present, the following equilibrium selection rule has been used to deal with this problem:

- In the case of multiple Nash equilibria, we will distribute the probability of occurrence equally over the two outcomes indicated in Table 4-9.

- In the case of non-existence of a Nash equilibrium we will distribute the probability of occurrence equally over all four outcomes.

Table 4-9 – Nash Equilibria in pure strategies

Player m/Player f	**F_1**	**F_2**	**F_3**	**F_4**
M_1	(1,1)	(1,1)	(1,0)	(1,0)
M_2	(1,1)	(1,1) & (0,0)	(0,0)	None
M_3	(0,1)	(0,0)	(0,0)	(0,1)
M_4	(0,1)	None	(1,0)	(1,0) & (0,1)

Using this, the probability of each observed outcome can be linked to the probability of each pair of reaction functions occurring as shown in Equation (4-7).

$$
\begin{aligned}
\Pr(0,0) \quad &= \Pr(F_2 \; and \; M_3) + \Pr(F_3 \; and \; M_2) + \Pr(F_3 \; and \; M_3) \\
&+0.5\Pr(F_2 \; and \; M_2) + 0.25\Pr(F_4 \; and \; M_2) + 0.25\Pr(F_2 \; and \; M_4) \\
\Pr(1,0) \quad &= \Pr(F_3 \; and \; M_1) + \Pr(F_3 \; and \; M_4) + \Pr(F_4 \; and \; M_1) \\
&+0.5\Pr(F_4 \; and \; M_4) + 0.25\Pr(F_4 \; and \; M_2) + 0.25\Pr(F_2 \; and \; M_4) \\
\Pr(0,1) \quad &= \Pr(F_1 \; and \; M_3) + \Pr(F_1 \; and \; M_4) + \Pr(F_4 \; and \; M_3) \\
&+0.5\Pr(F_4 \; and \; M_4) + 0.25\Pr(F_4 \; and \; M_2) + 0.25\Pr(F_2 \; and \; M_4) \\
\Pr(1,1) \quad &= \Pr(F_1 \; and \; M_1) + \Pr(F_1 \; and \; M_2) + \Pr(F_2 \; and \; M_1) \\
&+0.5\Pr(F_2 \; and \; M_2) + 0.25\Pr(F_4 \; and \; M_2) + 0.25\Pr(F_2 \; and \; M_4)
\end{aligned} \tag{4-7}
$$

Using Equation (4-7) and Table 4-8 we can derive probability statements for each outcome in terms of the model parameters which are given in Table 4-10. Maximum likelihood techniques can then be used to estimate the parameters in the model.

Table 4-10 – Outcome probabilities in terms of model parameters

Pr (0,0) $= \Phi(-x_m'\beta^m, -x_f'\beta^f, \rho)$
$-0.5\Omega(-x_m'\beta^m, -x_f'\beta^f, -x_m'\beta^m - \alpha^m, -x_f'\beta^f - \alpha^f, \rho)\ if\ \alpha^m \geq 0, \alpha^f \geq 0$
$= \Phi(-x_m'\beta^m, -x_f'\beta^f, \rho)$
$+0.25\Omega(-x_m'\beta^m, -x_f'\beta^f - \alpha^f, -x_m'\beta^m - \alpha^m, -x_f'\beta^f, \rho)\ if\ \alpha^m \geq 0, \alpha^f < 0$
$= \Phi(-x_m'\beta^m, -x_f'\beta^f, \rho)$
$+0.25\Omega(-x_m'\beta^m - \alpha^m, -x_f'\beta^f, -x_m'\beta^m, -x_f'\beta^f - \alpha^f, \rho)\ if\ \alpha^m < 0, \alpha^f \geq 0$
$= \Phi(-x_m'\beta^m, -x_f'\beta^f, \rho)\ if\ \alpha^m < 0, \alpha^f < 0$

Pr (1,0) $= \Phi(x_m'\beta^m, -x_f'\beta^f - \alpha^f, -\rho)\ if\ \alpha^m \geq 0, \alpha^f \geq 0$
$= \Phi(x_m'\beta^m, -x_f'\beta^f - \alpha^f, -\rho)$
$+0.25\Omega(-x_m'\beta^m, -x_f'\beta^f - \alpha^f, -x_m'\beta^m - \alpha^m, -x_f'\beta^f, \rho)\ if\ \alpha^m \geq 0, \alpha^f < 0$
$= \Phi(x_m'\beta^m, -x_f'\beta^f - \alpha^f, -\rho)$
$+0.25\Omega(-x_m'\beta^m - \alpha^m, -x_f'\beta^f, -x_m'\beta^m, -x_f'\beta^f - \alpha^f, \rho)\ if\ \alpha^m < 0, \alpha^f \geq 0$
$= \Phi(x_m'\beta^m, -x_f'\beta^f - \alpha^f, -\rho)$
$-0.5\Omega(-x_m'\beta^m - \alpha^m, -x_f'\beta^f - \alpha^f, -x_m'\beta^m, -x_f'\beta^f, \rho)\ if\ \alpha^m < 0, \alpha^f < 0$

Pr (0,1) $= \Phi(-x_m'\beta^m - \alpha^m, x_f'\beta^f, -\rho)\ if\ \alpha^m \geq 0, \alpha^f \geq 0$
$= \Phi(-x_m'\beta^m - \alpha^m, x_f'\beta^f, -\rho)$
$+0.25\Omega(-x_m'\beta^m, -x_f'\beta^f - \alpha^f, -x_m'\beta^m - \alpha^m, -x_f'\beta^f, \rho)\ if\ \alpha^m \geq 0, \alpha^f < 0$
$= \Phi(-x_m'\beta^m - \alpha^m, x_f'\beta^f, -\rho)$
$+0.25\Omega(-x_m'\beta^m - \alpha^m, -x_f'\beta^f, -x_m'\beta^m, -x_f'\beta^f - \alpha^f, \rho)\ if\ \alpha^m < 0, \alpha^f \geq 0$
$= \Phi(-x_m'\beta^m - \alpha^m, x_f'\beta^f, -\rho)$
$-0.5\Omega(-x_m'\beta^m - \alpha^m, -x_f'\beta^f - \alpha^f, -x_m'\beta^m, -x_f'\beta^f, \rho)\ if\ \alpha^m < 0, \alpha^f < 0$

Pr (1,1) $= \Phi(x_m'\beta^m + \alpha^m, x_f'\beta^f + \alpha^f, \rho)$
$-0.5\Omega(-x_m'\beta^m, -x_f'\beta^f, -x_m'\beta^m - \alpha^m, -x_f'\beta^f - \alpha^f, \rho)\ if\ \alpha^m \geq 0, \alpha^f \geq 0$
$= \Phi(x_m'\beta^m + \alpha^m, x_f'\beta^f + \alpha^f, \rho)$
$+0.25\Omega(-x_m'\beta^m, -x_f'\beta^f - \alpha^f, -x_m'\beta^m - \alpha^m, -x_f'\beta^f, \rho)\ if\ \alpha^m \geq 0, \alpha^f < 0$
$= \Phi(x_m'\beta^m + \alpha^m, x_f'\beta^f + \alpha^f, \rho)$
$+0.25\Omega(-x_m'\beta^m - \alpha^m, -x_f'\beta^f, -x_m'\beta^m, -x_f'\beta^f - \alpha^f, \rho)\ if\ \alpha^m < 0, \alpha^f \geq 0$
$= \Phi(x_m'\beta^m + \alpha^m, x_f'\beta^f + \alpha^f, \rho)\ if\ \alpha^m < 0, \alpha^f < 0$

Where $\Omega(a, b, c, d, \rho)$ is the integral of the bivariate standard normal density over the region $a \geq \varepsilon^m \geq c, b \geq \varepsilon^f \geq d$ (Note: If c>a or d>b we swap the limits and take the negative of the integral)

and $\Phi(a, b, \rho)$ is the bivariate standard normal cumulative density function with correlation parameter ρ

4.5 A cooperative choice model – Mixed Pareto Optimality/Nash

The equilibrium concept used in this model is a modification of that used in the Nash game. To define how the observed outcome is generated from the latent variables we follow the approach of Kooreman (1994) by dividing the possible situations into three cases as follows. In each case the same general approach of attempting to find the Nash equilibrium first and comparing it to the Pareto optimal allocations is followed. The payoffs in the game are given by the utilities specified in Equation (4-2).

1. Unique Nash equilibrium
   a. If the Nash equilibrium is also Pareto optimal then this is assumed to be the observed outcome.
   b. If the Nash equilibrium is not Pareto optimal there exists exactly one outcome at which both the male and the female are better off when compared to the Nash equilibrium. The players are assumed to choose this Pareto efficient allocation.
2. Two Nash equilibria (at least one of these will be Pareto optimal)
   a. If only one is Pareto optimal, this is assumed to be the observed outcome.
   b. If both Nash equilibria are Pareto optimal, the players are assumed to choose one of them with equal probability.
3. No Nash equilibrium – in this case the game may have two, three or four Pareto optimal allocations. The players are assumed to randomly choose one of these Pareto optimal allocations with equal probabilities.

Using this equilibrium concept, given the rankings of the four possible utilities each player may receive it is possible to find the observed outcome. After imposing an assumption for the distribution of the differences in the error terms in Equation (4-2) it is possible to find the probability of a particular utility ranking occurring as a function of the model parameters.

This makes it possible to find the probability of each of the four observed outcomes as a function of the model parameters.

Since there are 4! or 24 possible rankings of the utilities for each partner meaning there are 24^2 possible combinations of utility rankings for both partners, this approach becomes cumbersome without imposing some assumptions to reduce the number of possible utility rankings each partner may have. Table AV of Kooreman (1994), which is reproduced in Table 4-11 (note that U_{ij}^m denotes the male's utility when the male takes action *i* and the female takes action *j* while U_{ij}^f refers to the female's utility when the male takes action *i* and the female takes action *j*), gives the outcomes of the game for each of the 36 possible combinations of utility rankings in the case when $\alpha_1^m > 0, \alpha_0^m > 0$, $\alpha_1^f > 0$ and $\alpha_0^f > 0$.

In using the model in this thesis, we will assume that $\alpha_1^m > 0, \alpha_0^m > 0$, $\alpha_1^f > 0$ and $\alpha_0^f > 0$ in order to reduce the number of possible cases. This is equivalent to assuming that an individual's partner volunteering has a positive effect on their own utility regardless of whether they volunteer or not. It is also distinct from assuming the direction or existence of an endogenous effect in volunteering (i.e. that an individual's partner volunteering directly affects their own probability of volunteering) since this depends on the differences $\alpha_1^m - \alpha_0^m$ and $\alpha_1^f - \alpha_0^f$. In the *Stata* code the constraint is imposed by squaring the structural parameters before they enter the log-likelihood function.

While it may not be clear that such an assumption is valid and further investigation may be required, at this stage it is necessary to reduce the number of possible cases and it is also possible to provide some justification for it. In particular, if there is some spill-over benefit of volunteering between partners then this would imply that the structural parameters are all positive. A spill-over benefit may occur because the individual's partner volunteering makes the household appear charitable which in turn makes the individual appear more charitable gaining them social approval. This fits with the characterisation of volunteering as a conscience good or activity by Freeman (1997).

Table 4-11 – Utility rankings and observed outcomes for Pareto Nash game

	$U_{00}^f < U_{10}^f < U_{01}^f < U_{11}^f$	$U_{00}^f < U_{01}^f < U_{10}^f < U_{11}^f$	$U_{01}^f < U_{00}^f < U_{10}^f < U_{11}^f$	$U_{00}^f < U_{01}^f < U_{11}^f < U_{10}^f$	$U_{01}^f < U_{00}^f < U_{11}^f < U_{10}^f$	$U_{01}^f < U_{11}^f < U_{00}^f < U_{10}^f$
$U_{00}^m < U_{01}^m < U_{10}^m < U_{11}^m$	(1,1)	(1,1)	(1,1)	(1,0)	(1,0)	(1,0)
$U_{00}^m < U_{10}^m < U_{01}^m < U_{11}^m$	(1,1)	(1,1)	(1,1)	(1,0)	(1,0)	(1,0)
$U_{10}^m < U_{00}^m < U_{01}^m < U_{11}^m$	(1,1)	(1,1)	(1,1)	(1,1) or (1,0)	(1,1)	(0,0)
$U_{00}^m < U_{10}^m < U_{11}^m < U_{01}^m$	(0,1)	(0,1)	(1,1) or (0,1)	(1,0) or (0,1)	(1,0)	(1,0)
$U_{10}^m < U_{00}^m < U_{11}^m < U_{01}^m$	(0,1)	(0,1)	(1,1)	(0,1)	(1,1)	(0,0)
$U_{10}^m < U_{11}^m < U_{00}^m < U_{01}^m$	(0,1)	(0,1)	(0,0)	(0,1)	(0,0)	(0,0)

To illustrate how the log-likelihood function in the appendix is derived, it is detailed here for the case when both partners don't volunteer, i.e. (0,0). In the other cases it is more straightforward to calculate the probabilities for four cases depending on whether each of $\alpha_0^f - \alpha_1^f$ and $\alpha_0^m - \alpha_1^m$ are positive or negative. Using Table 4-11, we have:

$$\begin{aligned}
P(0,0) &= P\left(U_{10}^m < U_{11}^m < U_{00}^m < U_{01}^m \,\&\, U_{01}^f < U_{00}^f < U_{10}^f < U_{11}^f\right) \\
&\quad + P\left(U_{10}^m < U_{11}^m < U_{00}^m < U_{01}^m \,\&\, U_{01}^f < U_{00}^f < U_{11}^f < U_{10}^f\right) \\
&\quad + P\left(U_{10}^m < U_{11}^m < U_{00}^m < U_{01}^m \,\&\, U_{01}^f < U_{11}^f < U_{00}^f < U_{10}^f\right) \\
&\quad + P\left(U_{10}^m < U_{00}^m < U_{11}^m < U_{01}^m \,\&\, U_{01}^f < U_{11}^f < U_{00}^f < U_{10}^f\right) \\
&\quad + P\left(U_{10}^m < U_{00}^m < U_{01}^m < U_{11}^m \,\&\, U_{01}^f < U_{11}^f < U_{00}^f < U_{10}^f\right) \\
&= P\left(\varepsilon^m < -x_m'\beta^m - \alpha_1^m \,\&\, -x_f'\beta^f + min\left(0, \alpha_0^f - \alpha_1^f\right) < \varepsilon^f < -x_f'\beta^f\right) \\
&\quad + \mathrm{P}\left(\varepsilon^m < -x_m'\beta^m - \alpha_1^m \,\&\, -x_f'\beta^f - \alpha_1^f < \varepsilon^f < -x_f'\beta^f + \min\left(0, \alpha_0^f - \alpha_1^f\right)\right) \\
&\quad + P\left(\varepsilon^m < -x_m'\beta^m - \alpha_1^m \,\&\, \varepsilon^f < -x_f'\beta^f - \alpha_1^f\right) \\
&\quad + \mathrm{P}\left(-x_m'\beta^m - \alpha_1^m < \varepsilon^m < -x_m'\beta^m + \min(0, \alpha_0^m - \alpha_1^m) \,\&\, \varepsilon^f < -x_f'\beta^f - \alpha_1^f\right) \\
&\quad + \mathrm{P}\left(-x_m'\beta^m + \min(0, \alpha_0^m - \alpha_1^m) < \varepsilon^m < -x_m'\beta^m \,\&\, \varepsilon^f < -x_f'\beta^f - \alpha_1^f\right)
\end{aligned}$$

$$
\begin{aligned}
&= P\left(\varepsilon^m < -x_m'\beta^m - \alpha_1^m \;\&\; \varepsilon^f < -x_f'\beta^f\right) \\
&\quad + \mathrm{P}\left(-x_m'\beta^m - \alpha_1^m < \varepsilon^m < -x_m'\beta^m + \min(0, \alpha_0^m - \alpha_1^m) \;\&\; \varepsilon^f < -x_f'\beta^f - \alpha_1^f\right) \\
&\quad + \mathrm{P}\left(-x_m'\beta^m + \min(0, \alpha_0^m - \alpha_1^m) < \varepsilon^m < -x_m'\beta^m \;\&\; \varepsilon^f < -x_f'\beta^f - \alpha_1^f\right) \\
&= P\left(\varepsilon^m < -x_m'\beta^m \;\&\; \varepsilon^f < -x_f'\beta^f\right) \\
&\quad - P\left(-x_m'\beta^m - \alpha_1^m < \varepsilon^m < -x_m'\beta^m \;\&\; -x_f'\beta^f - \alpha_1^f < \varepsilon^f < -x_f'\beta^f\right) \\
&= \Phi\left(-x_m'\beta^m, -x_f'\beta^f, \rho\right) - \Omega(-x_m'\beta^m, -x_f'\beta^f, -x_m'\beta^m - \alpha_1^m, -x_f'\beta^f - \alpha_1^f, \rho)
\end{aligned}
$$

Where $\Omega(a, b, c, d, \rho)$ is the integral of the bivariate standard normal density over the region $a \geq \varepsilon^m \geq c,\; b \geq \varepsilon^f \geq d$ (Note: If c>a or d>b we swap the limits and take the negative of the integral)

and $\Phi(a, b, \rho)$ is the bivariate standard normal cumulative density function with correlation parameter ρ

4.6 Methods for checking accuracy of *Stata* code used to estimate models

Estimating many of the models in this thesis required *Stata* code to be written since they are not built-in to *Stata* and there is no user-written code available to estimate the models. Using the *ml* command in *Stata* it is possible to estimate the models by writing a program for each model which evaluates the log-likelihood function given the indexes (i.e. $x_m'\beta^m$ and $x_f'\beta^f$), the structural parameters and the correlation coefficient. The code for the relevant program for each model is provided in the appendix.

Given the use of non-standard code it is important to verify the accuracy of the code as errors in evaluating the log-likelihood function may result in serious errors in the estimation results (e.g. Cameron and Quiggin (1994) and the following correction Haab (1998)). In addition to carefully checking the derivation of the log-likelihood function and *Stata* code for errors, the following checks can be and were used to verify the accuracy of the code:

- When the structural parameters are constrained to zero in the game theoretic models they reduce to the bivariate probit with no endogenous effects. As such, the estimation results with the structural parameters constrained to zero were compared to those from the bivariate probit and were found to be identical.

- After estimating the models, the predicted probabilities for each observation of each of the four outcomes can be obtained from the log-likelihood function by taking the exponential of the log-likelihood contribution with only y_m and y_f changed to the appropriate values. The fact that these predicted probabilities should sum to one for each observation was used as an additional check.
- Initially, an attempt was made to replicate the results of Bjorn and Vuong (1984, 1985) and Chao (2002) however due to issues in obtaining the data used in these papers and reducing it to the same subsample, only qualitatively similar results could be obtained.

5 Estimation Results

5.1 Bivariate probit parameter estimates

A bivariate probit model with no endogenous effects was estimated to provide a base model comparator for the game theoretic models. Table 5-1 details these estimation results and provides a listing of the explanatory variables included in the male and female volunteering equations. The explanatory variables are grouped into categories depending on their nature which are demographic, location/residence, education, employment, health and wealth/income. The estimates for the parameters on the explanatory variables in the game theoretic models are similar to those found in Table 5-1.

Table 5-1 – Bivariate probit estimation results

	Variable	I/H[1]	Male eqn. Coefficient (Std Err)[2]	Female eqn. Coefficient (Std Err)[2]
Demographic	overseasborn_english	I	-0.1741 (0.0857) **	-0.1098 (0.0863)
	overseasborn_other	I	-0.3884 (0.0967) ***	-0.2937 (0.0890) ***
	arrived_less_than_10y	I	-0.2478 (0.2466)	-0.2558 (0.1964)
	aboriginal_torres	I	-1.0533 (0.4452) **	0.2222 (0.2393)
	age	I	0.0396 (0.0136) ***	0.0919 (0.0136) ***
	age_sq	I	-0.00037 (0.00013) ***	-0.0009 (0.0001) ***
	child_under_10	I	-0.2123 (0.0797) ***	-0.1054 (0.0795)
	num_child	I	0.1764 (0.0296) ***	0.0934 (0.0287) ***
	religion	I	0.6433 (0.0715) ***	0.6641 (0.0662) ***
Location/Residence	state_vic	I	0.0360 (0.0735)	0.1036 (0.0705)
	state_qld	I	-0.1068 (0.0806)	-0.0123 (0.0756)
	state_sa	I	0.1066 (0.0984)	0.0843 (0.0961)
	state_wa	I	-0.1727 (0.1042) *	-0.0785 (0.0968)
	state_tas	I	0.0255 (0.1594)	-0.2815 (0.1636) *
	state_nt	I	0.3693 (0.3373)	0.5742 (0.3100) *
	state_act	I	-0.1046 (0.2038)	0.3295 (0.1827) *
	other_urban	I	0.1653 (0.0707) **	0.1559 (0.0664) **
	bounded_locality	I	0.1256 (0.1546)	0.0228 (0.1514)
	rural_balance	I	0.2052 (0.0789) ***	0.2308 (0.0743) ***
	yr_curr_addr	I	0.0006 (0.0027)	0.0052 (0.0028) *
	yr_curr_addr_missing	I	-0.1951 (0.4225)	0.4038 (0.3083)
	likely_to_move	I	-0.0828 (0.0916)	0.0505 (0.0868)

[1] If the variable is for the individual (I), it has the part_ prefix when included in the female equation. If it is for the household (H), exactly the same variable is included in both equations.
[2] *,**,*** denote significance at the 10%, 5% and 1% levels respectively. Values are reported to 4 decimal places and 5 decimal places when zero to 3 decimal places.

Educ.	postgrad	I	0.5311 (0.1053) ***	0.2812 (0.0963) ***
	bachelor	I	0.2981 (0.1003) ***	0.2644 (0.0851) ***
	post_school	I	0.2817 (0.0733) ***	0.1515 (0.0675) **
	yr12	I	0.1338 (0.1133)	0.1099 (0.0845)
Employment	hours_worked	I	-0.0086 (0.0022) ***	-0.0146 (0.0024) ***
	hours_vary	I	-0.4451 (0.1975) **	-0.2479 (0.1856)
	more_than_one_job	I	0.2833 (0.1137) **	0.3980 (0.1053) ***
	irregular_hours	I	0.1403 (0.0764) *	0.1396 (0.0801) *
	fixedterm_casual	I	0.0651 (0.0890)	-0.0055 (0.0749)
	satisfied_balance_work	I	0.1841 (0.0796) **	0.0126 (0.0804)
	retire_in_5y	I	-0.1856 (0.1271)	-0.1702 (0.1227)
	office_worker	I	0.2218 (0.0732) ***	0.1038 (0.0984)
	unemployed	I	0.1357 (0.2410)	0.1989 (0.2021)
	retired	I	0.3661 (0.1451) **	0.0370 (0.1054)
	worried_about_job	I	-0.0678 (0.0786)	0.0848 (0.0886)
Health	long_term_hlth_cond	I	0.1088 (0.0812)	-0.1277 (0.0805)
	health_excellent	I	0.4346 (0.1770) **	0.2547 (0.1591)
	health_verygood	I	0.2674 (0.1626) *	0.2883 (0.1441) **
	health_good	I	0.1876 (0.1617)	0.2164 (0.1441)
	health_fair	I	0.0704 (0.1744)	-0.0177 (0.1608)
	health_poor	I	-0.3048 (0.2350)	-0.1352 (0.2207)
	pressed_for_time	I	0.0351 (0.0611)	0.1282 (0.0562) **
Wealth/Inc	hh_gross_inc	H	-0.0012 (0.0007) *	-0.00028 (0.00047)
	hh_home_equity	H	-0.00001 (0.00008)	0.00014 (0.00007) *
	hh_net_worth	H	0.00005 (0.00002) **	0.00005 (0.00002) **
	ind_disp_income	I	0.0010 (0.0013)	-0.0011 (0.0015)
	owns_home	I	0.1298 (0.1015)	0.0086 (0.0813)
	Constant	N/A	-2.6718 (0.3841) ***	-3.4587 (0.3607) ***

5.2 Selected parameter estimates from bivariate probit and game theoretic models

Table 5-2 includes some key parameter estimates from the bivariate probit model with no endogenous effects, the two possible recursive formulations for the bivariate probit with endogenous effects (the specification is obtained from Equation (2-1) by setting either $\gamma_1 = 0$ or $\gamma_2 = 0$ depending on whether we wish to include the male's decision in the female's equation or vice-versa) and the game theoretic models. The set of explanatory variables included in each of these models is identical. While the bivariate probit with endogenous effects has not been discussed previously, it is included here for comparison. For the game theoretic models, the parameters are defined in Section 4.4 - Non-cooperative choice models, in particular see Equation (4-2). The parameter ρ is the correlation between the error terms. In some cases it is not possible to individually identify α_1^m and α_0^m or α_1^f and α_0^f but only their differences. In these cases, the two cells are merged and the difference is reported.

Table 5-2 - Key parameter estimates

Model	$\boldsymbol{\rho}$	Log-likelihood	$\boldsymbol{\alpha_1^m}$	$\boldsymbol{\alpha_0^m}$	$\boldsymbol{\alpha_1^f}$	$\boldsymbol{\alpha_0^f}$
Bivariate probit (BVP)	0.3592 (0.0331)	-2942.00				
BVP with male decision inc.	0.0636 (0.1381)	-2939.28			0.53685 * (0.23751)	
BVP with female decision inc.	-0.0792 (0.1473)	-2936.97	0.7665 ** (0.2437)			
Nash game	-0.9298 (0.0520)	-2934.85	1.4537 ** (0.1364)		1.2357 ** (0.1351)	
Stackelberg male leader	-0.4250 (0.1809)	-2932.90	1.1221 ** (0.3815)	0.3591 (0.3134)	0.5043 * (0.2040)	
Stackelberg female leader	-0.4101 (0.1720)	-2932.29	0.8055 ** (0.2297)		0.7839 ** (0.2812)	0.4813 (0.3048)
Pareto Nash game	-0.4101 (0.1720)	-2932.29	0.8055 ** (0.2297)	0.00000	0.7840 ** (0.2813)	0.4813 (0.3048)
Number of observations used in all models: 3255						

Note: Standard errors in round brackets, *, ** denote significance at 5%, 1% level

In addition to the results in Table 5-2, some selected estimates of the parameters on the explanatory variables in the game theoretic models are reported in Table 5-3. While the full set of explanatory variables detailed in Table 5-1 was included in all the models, the estimated coefficients were generally similar to those reported in Table 5-1 for the bivariate probit model. As a result, due to space constraints only some key parameter estimates are reported in Table 5-3.

Table 5-3 – Selected coefficient estimates from game theoretic models

	Stackelberg male leader	Stackelberg female leader	Nash game	Pareto Nash game
Male equation				
overseasborn_other	-0.3234 *** (0.1029)	-0.3102 *** (0.0995)	-0.1946 ** (0.0835)	-0.3103 *** (0.0995)
age	0.0246 * (0.0145)	0.0231 (0.0141)	0.0150 (0.0120)	0.0231 (0.0141)
age_sq	-0.0002 (0.0001)	-0.0002 (0.0001)	-0.0001 (0.0001)	-0.0002 (0.0001)

postgrad	0.5329 *** (0.1079)	0.5201 *** (0.1045)	0.4297 *** (0.0867)	0.5201 *** (0.1045)
bachelor	0.3276 *** (0.1006)	0.3204 *** (0.0976)	0.2827 *** (0.0803)	0.3204 *** (0.0976)
yr12	0.1458 (0.1124)	0.1449 (0.1094)	0.1599 * (0.0878)	0.1449 (0.1094)
hours_worked	-0.0073 *** (0.0023)	-0.0071 *** (0.0023)	-0.0051 *** (0.0018)	-0.0071 *** (0.0023)
unemployed	0.1425 (0.2389)	0.1440 (0.2334)	0.0999 (0.1935)	0.1440 (0.2334)
retired	0.3533 ** (0.1470)	0.3452 ** (0.1431)	0.2640 ** (0.1182)	0.3452 ** (0.1431)
child_under_10	-0.2187 *** (0.0813)	-0.2124 *** (0.0795)	-0.2113 *** (0.0743)	-0.2124 *** (0.0795)
num_child	0.1341 *** (0.0339)	0.1296 *** (0.0329)	0.0932 *** (0.0284)	0.1296 *** (0.0329)
religion	0.4960 *** (0.0936)	0.4776 *** (0.0897)	0.3701 *** (0.0671)	0.4776 *** (0.0897)
_cons	-2.4949 *** (0.4034)	-2.4608 *** (0.3946)	-2.1612 *** (0.3400)	-2.4608 *** (0.3946)
Female equation				
part_overseasborn_other	-0.3079 *** (0.0883)	-0.3173 *** (0.0904)	-0.2532 *** (0.0756)	-0.3173 *** (0.0904)
part_age	0.0894 *** (0.0138)	0.0912 *** (0.0142)	0.0734 *** (0.0128)	0.0912 *** (0.0142)
part_age_sq	-0.0009 *** (0.0001)	-0.0009 *** (0.0001)	-0.0007 *** (0.0001)	-0.0009 *** (0.0001)
part_postgrad	0.2199 ** (0.0974)	0.2305 ** (0.1001)	0.1600 ** (0.0811)	0.2305 ** (0.1001)
part_bachelor	0.1898 ** (0.0861)	0.1924 ** (0.0878)	0.1263 * (0.0698)	0.1924 ** (0.0878)
part_yr12	0.1010 (0.0828)	0.1081 (0.0850)	0.0650 (0.0700)	0.1081 (0.0850)
part_hours_worked	-0.0140 *** (0.0024)	-0.0142 *** (0.0025)	-0.0111 *** (0.0021)	-0.0142 *** (0.0025)
part_unemployed	0.1688 (0.1924)	0.1635 (0.1963)	0.1609 (0.1557)	0.1635 (0.1964)
part_retired	0.0206 (0.1027)	0.0191 (0.1052)	0.0277 (0.0858)	0.0191 (0.1052)
part_child_under_10	-0.0467 (0.0811)	-0.048 (0.0832)	-0.0028 (0.0739)	-0.0480 (0.0832)
part_num_child	0.0719 ** (0.0299)	0.0771 ** (0.0311)	0.0582 ** (0.0274)	0.0771 ** (0.0311)
part_religion	0.5731 *** (0.0821)	0.5921 *** (0.0863)	0.4537 *** (0.0626)	0.5921 *** (0.0863)
_cons	-3.5438 *** (0.3609)	-3.6014 *** (0.3717)	-3.1924 *** (0.3315)	-3.6014 *** (0.3717)

Standard errors in round brackets, *, **, *** denote significance at the 10%, 5%, 1% levels

5.3 Average partial effects of selected explanatory variables

To assist in interpreting the parameter estimates, for each of the game theoretic models a selection of average partial effects on the male and female probabilities of volunteering were calculated. These have been reported in Table 5-4. Due to the nature of the models, a change in the male or female characteristics will have an impact on both the male and female probabilities of volunteering. In general these probabilities were calculated as follows:

1. Calculate the predicted probabilities of the four possible observed outcomes for the base group. If we wish to calculate the effect of a change in a male characteristic, the base group consists of all couples in the sample where the male has the initial value for the variable we wish to calculate the effect of, i.e. if we wish to calculate the effect of the male retiring, the base group would consist of all couples where the male is not retired. If we wish to calculate the effect of a change in a female characteristic, the base group consists of all couples in the sample where the female has the initial value for the relevant variable.
2. Depending on whether we wish to calculate the effect for a change in a male or female characteristic, adjust the male or female index by the change in the index caused by the change in the variable. i.e. if we wish to calculate the effect of the male's education increasing from year 12 to a bachelors degree we would adjust the male index by the difference between the estimated coefficients on the bachelors degree dummy variable and the year 12 dummy variable.
3. Recalculate the predicted probabilities of the four possible observed outcomes for the base group.
4. Calculate the change in the male probability of volunteering (defined as $P(Male\ volunteering) = P(1,1) + P(1,0)$) and the female probability of volunteering (defined as $P(Female\ volunteering) = P(1,1) + P(0,1)$) for each observation in the base group.

5. Average these changes in the male and female probabilities of volunteering over the base group to obtain the average partial effects.

It is clear that the number of observations used to calculate the average partial effects will vary with the size of the appropriate base group. In addition, since in these models the effect of an explanatory variable on the predicted probabilities will depend on the values of all the other explanatory variables, it is important to choose appropriate values of these other explanatory variables. By using the method proposed here, the values of these other explanatory variables will be determined by the sample rather than chosen arbitrarily, thus ensuring that appropriate values of the other explanatory variables are used.

The standard errors for the average partial effects are reported in brackets in Table 5-4, which were calculated using the bootstrap command in *Stata.* This process involves the following steps:

1. Resampling with replacement the 3255 couple level observations to obtain a new sample of 3255 couple level observations. Since the resampling is carried out at the couple level, each individual still has the same partner in the new sample.
2. Estimating the relevant game theoretic model on the new sample, storing the relevant parameter estimates and index values.
3. Use these values to calculate the average partial effects with the method described previously and store these average partial effects.

While this process was repeated 500 times for each model, the actual number of bootstrap replications used to calculate each standard error is reported in Table 5-4. The discrepancy arises because the estimation procedure for these models appears to be somewhat fragile and as a result while the estimation procedure was repeated 500 times, it sometimes failed to converge (similar issues were encountered in Sections 7.1 and 7.2). In addition, the maximum number of iterations was restricted to 500 rather than the default 16 000 in *Stata* since otherwise a small number of replications consumed a large proportion of the processing time. As a result, in a small number of cases the optimisation algorithm terminated and the results were reported before the usual convergence criteria were satisfied (details of these cases are available on request). These results were still used to compute the statistics in Table 5-4.

Table 5-4 – Average partial effects of selected explanatory variables

	Stackelberg male		**Stackelberg female**		**Nash**		**Pareto Nash**	
	Male	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**
Male variables								
Yr 12 to Post school	0.034 (0.020)	0.007 (0.006)	0.033 (0.020)	0.007 (0.006)	0.021 (0.020)	0.008 (0.006)	0.033 (0.020)	0.007 (0.006)
Bachelor to Post grad	0.058 (0.030)	0.011 (0.009)	0.058 (0.031)	0.010 (0.009)	0.050 (0.031)	0.018 (0.009)	0.058 (0.030)	0.010 (0.008)
Age from 20 to 30	0.017 (0.011)	0.002 (0.002)	0.016 (0.012)	0.002 (0.003)	0.012 (0.011)	0.002 (0.002)	0.016 (0.011)	0.002 (0.002)
Hrs. worked from 30 to 40	-0.019 (0.007)	-0.004 (0.003)	-0.018 (0.007)	-0.004 (0.002)	-0.015 (0.007)	-0.006 (0.003)	-0.018 (0.007)	-0.004 (0.002)
Retired	0.095 (0.046)	0.019 (0.015)	0.095 (0.047)	0.018 (0.017)	0.085 (0.046)	0.032 (0.017)	0.095 (0.046)	0.018 (0.015)
Religion	0.138 (0.029)	0.027 (0.018)	0.136 (0.031)	0.025 (0.022)	0.123 (0.030)	0.046 (0.019)	0.136 (0.030)	0.025 (0.017)
Female variables								
Yr 12 to Post school	0.001 (0.006)	0.004 (0.024)	0.001 (0.005)	0.003 (0.024)	0.002 (0.005)	0.006 (0.023)	0.001 (0.006)	0.003 (0.024)
Bachelor to Post grad	0.002 (0.007)	0.009 (0.029)	0.003 (0.007)	0.011 (0.029)	0.004 (0.008)	0.011 (0.031)	0.003 (0.008)	0.011 (0.030)
Age from 20 to 30	0.014 (0.007)	0.073 (0.012)	0.015 (0.007)	0.073 (0.012)	0.018 (0.006)	0.067 (0.011)	0.015 (0.006)	0.073 (0.011)
Hrs. worked from 30 to 40	-0.009 (0.004)	-0.037 (0.006)	-0.009 (0.005)	-0.037 (0.006)	-0.013 (0.004)	-0.035 (0.006)	-0.009 (0.004)	-0.037 (0.006)
Retired	0.001 (0.007)	0.006 (0.030)	0.001 (0.007)	0.005 (0.030)	0.003 (0.008)	0.009 (0.029)	0.001 (0.007)	0.005 (0.029)
Religion	0.043 (0.020)	0.185 (0.031)	0.044 (0.023)	0.187 (0.032)	0.063 (0.021)	0.169 (0.031)	0.044 (0.021)	0.187 (0.030)
Bootstrap reps.	403		461		307		400	

For some of the models the estimated average partial effects are missing for a large number of replications. If they are not missing at random, it is possible for this to have a substantial impact on the estimated standard errors. While it is difficult to confirm whether or not this is the case without attempting to use alternative maximization algorithms to reduce the number of missing values, some reassurance is provided by comparing the estimated standard errors across models. This is because the estimated standard errors are very similar across models despite the number of successful bootstrap replications varying considerably. The standard errors are expected to be similar across models due to the comparable coefficient estimates and standard errors (see Section 5.2) and average partial effects.

5.4 Conditional probabilities of volunteering

To better understand the implications of the results for interactions between male and female volunteering decisions a number of conditional probabilities of volunteering have been calculated. These probabilities have been calculated as $P(y^m = 1 | y^f = 1) = \frac{P(1,1)}{P(1,1)+P(0,1)}$, $P(y^m = 1 | y^f = 0) = \frac{P(1,0)}{P(1,0)+P(0,0)}$, $P(y^f = 1 | y^m = 1) = \frac{P(1,1)}{P(1,1)+P(1,0)}$ and $P(y^f = 1 | y^m = 0) = \frac{P(0,1)}{P(0,1)+P(0,0)}$. The values in Table 5-5 were obtained by finding the predicted probabilities for each observation of the four observed outcomes, calculating the conditional probabilities for each observation and reporting the summary statistics for these conditional probabilities in the table.

Table 5-5 – Conditional probabilities of volunteering

Model & Probability	Mean	Std. Dev.	Min	Max
$P(y^m = 1 \mid y^f = 1)$				
Stackelberg male leader	0.3196	0.1465	0.0006	0.8714
Stackelberg female leader	0.3239	0.1436	0.0007	0.8583
Nash	0.3014	0.1746	0.0000	0.9140
Pareto Nash	0.3239	0.1436	0.0007	0.8583
$P(y^m = 1 \mid y^f = 0)$				
Stackelberg male leader	0.1499	0.1161	0.0004	0.8596
Stackelberg female leader	0.1510	0.1205	0.0004	0.8973
Nash	0.1483	0.1067	0.0013	0.8211
Pareto Nash	0.1510	0.1205	0.0004	0.8973
$P(y^f = 1 \mid y^m = 1)$				
Stackelberg male leader	0.4097	0.1698	0.0095	0.9509
Stackelberg female leader	0.4124	0.1622	0.0090	0.9429
Nash	0.3834	0.2044	0.0000	0.9746
Pareto Nash	0.4124	0.1622	0.0090	0.9429
$P(y^f = 1 \mid y^m = 0)$				
Stackelberg male leader	0.2047	0.1422	0.0059	0.9608
Stackelberg female leader	0.2053	0.1453	0.0054	0.9698
Nash	0.2034	0.1324	0.0106	0.9682
Pareto Nash	0.2053	0.1453	0.0054	0.9698

6 Discussion

6.1 Estimated impacts of explanatory variables

The coefficient estimates and significance levels are generally comparable across all the models that have been considered, with the main exception being the magnitude of the coefficient estimates is often smaller in the Nash model. In addition, the coefficient estimates are largely in line with our expectations. As such, while the following discussion focuses on the coefficient estimates reported in Table 5-1 for the bivariate probit model with no endogenous effects, the comments are generally the same for the coefficient estimates for the game theoretic models reported in Table 5-3. The average partial effects reported in Table 5-4 are also referenced to provide some information regarding the economic significance of the estimates.

The estimated impact of being overseas born is negative for both the male and the female and is statistically significant at the 1% level in the case of being overseas born in a non-English speaking country. This can be attributed to the fact an individual born overseas may feel less integrated with their local community and as such is less likely to volunteer. In addition, age enters the model in a quadratic form and as such has a positive effect until a turning point at 53.6 years for males and 52.2 years for females and thereafter has a negative effect. Both age and age squared are individually statistically significant in both equations at the 1% level and are jointly statistically significant at the 5% level in the male equation and the 1% level in the female equation. The average partial effect for a change in the female's age on the female's probability is also statistically significant at the 1% level. This matches findings by Freeman (1997) that volunteering peaks in the 35-54 age range. When considering the average partial effects of age in Table 5-4, it can be seen that changes in age have a much larger effect for females than males and as such while age has an economically significant effect for females (when taking into account the relatively low rates of volunteering in the sample), it does not for males. In fact, the effect of a change in the female's age on the male's probability of volunteering as reported in Table 5-4 (this effect is due to the impact of the female's volunteering decision on the males) is around the same as the effect of a similar change in the male's age on the male's probability of volunteering.

Number of children has a positive and statistically significant effect (1% level) on an individual's probability of volunteering, which is similar to findings by Carlin (2001). Having a young child has a negative effect on volunteering which may be attributable to difficulties in finding appropriate care for children while volunteering. Attendance at religious services has a large positive and statistically significant effect on an individual's probability of volunteering which is in line with results from Gomez and Gunderson (2003). When considering the average partial effects reported in Table 5-4, the effect is economically significant with an increase in the male probability of volunteering of around 0.14 and a larger increase of around 0.19 for the female (both of these effects are statistically significant at the 1% level). Gomez and Gunderson (2003) suggest that this is due to traits within individuals which motivate them to engage in socially oriented activities and the fact many religions encourage such selfless behaviour.

The fact none of the state dummy variables are significant at the 5% level indicates that there is little evidence for state effects in volunteering behaviour. There is some evidence that the size of the area an individual lives in has an effect on volunteering with living in an area classified as rural balance having a positive and statistically significant (1% level) effect on an individual's probability of volunteering, relative to the base category which is Major Urban. This is likely to be attributable to stronger communities in these areas. Years at current address and whether an individual is likely to move appear to have little impact on their probability of volunteering.

In addition, the coefficients on the education dummy variables indicate that the more highly educated are more likely to volunteer since all of these coefficients are positive and many are statistically significant at the 1% level (note that the base group is those who did not complete Year 12). This is in line with many previous studies on volunteering, for example Freeman (1997) finds that the more highly educated are more likely to volunteer. Table 5-4 indicates that while economically the effect of education is significant for the male, it is not for the female (in most of the models the average partial effect of Bachelor to Post grad for the male is statistically significant at the 5% level).

In regard to the employment variables, hours worked has a negative and statistically significant effect (1% level) on both the male and female's probability of volunteering. Table 5-4 indicates that the effect is larger for females than males, with the effect being economically quite small for males and around twice the size for females (the estimated average partial effect is statistically significant at the 1% level in all the models excluding the Nash model where it is significant at the 5% level). While this is expected when considering the budget constraint in Section 4.1 since working additional hours reduces the amount of time available for volunteering, in a previous study Freeman (1997) finds that the data does not support this simple substitution story. Interestingly, whether an individual has more than one job has a positive and statistically significant (5% - male equation; 1% - female equation) effect on their probability of volunteering. This may be attributable to factors such as those who work multiple jobs are better able to manage time commitments and therefore fit in volunteering activities. Of the other employment variables, only satisfied_balance_work, office_worker and retired are positive and statistically significant (only in the male equation at the 5% level, 1% level and 5% level respectively). The effect of being retired is not economically significant for the female but is for the male with Table 5-4 indicating an increase in the male's probability of volunteering of around 0.1 (the average partial effect for the male is statistically significant at the 5% level in all the models except the Nash model where it is significant at the 10% level). It is expected that satisfied_balance_work and retired have a positive effect since those who feel they are better able to balance work and non-work commitments would be more likely to volunteer and retired people are more likely to volunteer due to greater availability of time. In regard to office_worker it is difficult to form any strong expectations of the direction of the effect.

In relation to the health variables, the majority of these are not statistically significant, however the self assessed health status dummy variables indicate a positive relationship between self assessed health status and volunteering. That is, individuals in better health are more likely to volunteer.

Finally, the wealth and income variables do not appear to play a substantial role in determining volunteering behaviour. Only the household net worth variable is statistically significant (5% level) and has a positive sign as would be expected. The effect of the other income/wealth variables appears small when compared to the impacts of some of the other explanatory variables. Brown and Lankford (1992) provide some discussion as to whether hours worked or income is the relevant control variable and given the nature of many employment contracts in the Australian labour market it is not surprising that income has a limited effect on volunteering behaviour.

These results provide some support for there being exogenous or contextual effects (Manski, 1993) in volunteering behaviour. This is because, for example, individuals of similar age and education levels are more likely to become couples and individuals with higher education levels and ages around 50 are more likely to volunteer.

6.2 Estimated endogenous effects of volunteering

Each of the models considered provides support for there being endogenous effects in volunteering behaviour. Table 5-2 indicates that all the models which are able to capture at least some endogenous effects provide a substantial improvement in the log-likelihood over the bivariate probit with no endogenous effects. Since the bivariate probit with no endogenous effects is nested in each of the other models the significance of this difference can be formally tested using a likelihood ratio test. We find that the improvement is statistically significant at the 1% level in all models except the bivariate probit with the male decision included. In this model it is significant at the 5% level. In addition, the structural parameters reported in Table 5-2 are often statistically significant and reasonably large in comparison to other coefficient estimates reported in Table 5-3.

Further evidence of the presence of endogenous effects in volunteering behaviour is provided by comparing the correlation parameter ρ across models. In the bivariate probit model with no endogenous effects the correlation parameter is positive and statistically significant, however in all the game theoretic models it is negative and statistically significant. In addition, when endogenous effects are added to the bivariate probit model in a recursive formulation (that is the male's volunteering decision is included in the female's volunteering equation or vice-versa) the correlation parameter decreases substantially and becomes insignificant. This indicates that the bivariate probit with no endogenous effects is incorrectly specified and as a result the endogenous effects are being picked up by the correlation parameter.

To provide some guidance as to the impact of each partner's volunteering decision on the others, the conditional probabilities of volunteering reported in Table 5-5 can be used. It should be noted that the differences in these probabilities (i.e. $P(y^m = 1|y^f = 1) - P(y^m = 1|y^f = 0)$ and $P(y^f = 1|y^m = 1) - P(y^f = 1|y^m = 0)$) are affected not only by the magnitudes of the structural parameters but also by the magnitude of the correlation parameter. However, since the correlation parameter is negative in all the game theoretic models (and therefore acts to reduce the size of these differences), the impact of the endogenous effect is at least as large as these differences.

Table 5-5 indicates that the conditional probabilities of volunteering are similar across the four game theoretic models. In addition, it appears that the impact of the partner's volunteering decision on their spouse's probability of volunteering is substantial with a male whose partner volunteers being around 15 percentage points more likely to volunteer and a female whose partner volunteers being around 20 percentage points more likely to volunteer. While volunteering has been modelled as a discrete decision, this suggests that volunteering (when modelled in terms of hours volunteered) is a strategic complement and therefore the individual's reaction functions are upward sloping. This indicates that policies which encourage individuals to volunteer are likely to have substantial multiplier effects since an individual choosing to volunteer has a substantial positive impact on their partner's probability of volunteering.

6.3 Model selection

Each of the models considered in this thesis implies a different structure for couple decision making and as such if it is possible to differentiate between them and choose a preferred model this will provide some insights into the nature of decision making within couples. As discussed in Section 2.2, a possible approach to this is to use the likelihood dominance criterion (LDC) (Pollak and Wales, 1991) or alternatively various information criterions could be considered such as AIC and BIC.

Pollak and Wales (1991) develop the LDC by introducing the 'dominance ordering'. The dominance ordering is constructed by considering nesting the two non-nested hypotheses in a composite which cannot be estimated. Since the composite cannot be estimated we must focus on choosing between the two hypotheses rather than comparing them with the composite (i.e. we ignore the possibility of rejecting or accepting both against the composite). Pollak and Wales (1991) show that it is not necessary to estimate the composite to determine which hypothesis is preferred (interpreted as the likelihood ratio test accepting the hypothesis against the composite while the other is rejected against the composite) since the ranking of the two hypotheses depends only on the parametric size of the composite and the two log-likelihood values for the individual models. The LDC generalises the dominance ordering by considering a range of admissible composite parametric sizes. Where the dominance ordering does not agree for all admissible parametric sizes, the LDC is indecisive. If the dominance ordering agrees for all admissible parametric sizes, the LDC gives the same ordering of the hypotheses.

To apply the LDC we note that the Nash model has 103 parameters, both of the Stackelberg leader models have 104 parameters and the Pareto Nash model has 105 parameters. The log-likelihood values for each model are reported in Table 5-2 while the relevant critical values (at the 1% level) for the LDC are reported in Table 6-1. Since the difference in the log-likelihood between the Nash model and the two Stackelberg leader models exceeds the critical value of 1.2877 in both cases, we conclude that the LDC prefers the Stackelberg leader models to the Nash model.

Table 6-1 – Critical values for likelihood dominance criterion

$n_2 \setminus n_1$	103	104
104	0.5805, 1.2877	
105	1.1607, 2.3550	0.5802, 1.2877

While we could also apply the LDC to choosing between the Pareto Nash model and the two Stackelberg leader models, it is more worthwhile to explore the fact there is very little difference in the log-likelihood values for the Stackelberg female leader model and the Pareto Nash model. This similarity is not restricted to the log-likelihood values as can be seen by comparing the estimates of the other parameters in Table 5-2 and Table 5-3. The reason for this is that given the restrictions placed on the structural parameters in order to estimate the Pareto Nash model and the differences between the structural parameters, the Pareto Nash model generally predicts the same outcomes as the Stackelberg leader models. It is possible for the Pareto Nash model to also "mimic" the Stackelberg male leader model, however it has tended towards the Stackelberg female leader model due to the lower log-likelihood value for this model.

As such, on the basis of this it may be possible to claim that the Stackelberg female leader model is the preferred model, however it is not clear that this is an appropriate conclusion. This is because as will be illustrated later in Table 7-2, the log-likelihood value for the Nash model is somewhat sensitive to the weights that are used in the estimation procedure. If we were to take the maximum log-likelihood value across the simulations as the log-likelihood value for the Nash model, then the difference between it and the Stackelberg female leader model is very close to being in the region where the likelihood dominance criterion is indecisive.

In addition, it is possible that the fact the Pareto Nash model mimics the Stackelberg female leader model is a result of the restrictions imposed in order to estimate the model. Further research is required in order to estimate the model without these restrictions and determine whether this is the case. Also, the similarity in fit between the two Stackelberg leader models could be interpreted as meaning that some households have the female acting as leader while others have the male acting as leader which provides some motivation for considering a mixture model rather than concluding the female leader model is preferred.

Finally, it is difficult to differentiate between the models in terms of the substantive inferences. This is due to the similarities between the models discussed in Sections 6.1 and 6.2. In terms of the key research question which relates to the correlation in volunteering decisions, it does not appear to make a great deal of difference which model is used. As a result given the data available on volunteering decisions, it seems difficult to determine what the underlying decision making structure is.

6.4 Implications for volunteering behaviour

As mentioned in Section 6.2, it appears that there are substantial multiplier effects in volunteering behaviour. This is because the observed correlation in volunteering decisions does not seem to be solely due to exogenous or observable characteristics, but can also be attributed to the presence of endogenous effects. There are a variety of possible rationalisations for the presence of these endogenous effects. While it is not possible to distinguish between them on the basis of the available data, we may consider the explanations and corresponding situations given below:

- There is likely to be a strong tendency for couples to be involved in the same organisation and it may be possible that one has encouraged the other to join. For instance a father may begin coaching a team at a local sports club and as a result finds out they are in need of help in the canteen so asks his wife to assist. This fits with

Freeman's (1997) findings that being asked to volunteer is an important determinant of volunteering behaviour.

- An individual may be more familiar with the benefits of volunteering if their partner volunteers. For instance, consider a household where the wife decides to join a group of volunteers that look after a local park. After visiting the park and hearing from his wife about the scope of the volunteers' work, the husband then decides to join a similar volunteer organisation.
- Partners may wish to share the social aspect of volunteering. For instance suppose a male retires and as a result decides to join a volunteer organisation to meet new people other than old work colleagues. He may then encourage his partner to join the same organisation so that she also meets the same people.

7 Sensitivity checks and opportunities for further research

7.1 Simulation exercise for testing sensitivity of Nash model

7.1.1 Setup

As discussed in Section 2.2, issues arise when using the Nash equilibrium in pure strategies as an equilibrium concept due to the fact an equilibrium may not exist or there may be multiple equilibria. While there are a variety of methods available to deal with this problem, an equilibrium selection rule was imposed to produce the estimation results in Section 5, with the weights used detailed in Section 4.4.2. Since there is no strong justification for assigning the weights used in Section 4.4.2, this section aims to assess the impact of varying the weights over their possible ranges on the estimation results. The sets of weights which must be chosen in order to estimate the model by maximum likelihood estimation are detailed in Table 7-1.

Table 7-1 – Weights used to estimate Nash model

Weight	Description
a_1	$P(0,0)$ in the case where both (0,0) and (1,1) are possible equilibriums
a_2	$P(1,1)$ in the case where both (0,0) and (1,1) are possible equilibriums
b_1	$P(1,0)$ in the case where both (1,0) and (0,1) are possible equilibriums
b_2	$P(0,1)$ in the case where both (1,0) and (0,1) are possible equilibriums
c_1	$P(0,0)$ in the case where an equilibrium does not exist and $\alpha^m \geq 0$ and $\alpha^f < 0$
c_2	$P(1,0)$ in the case where an equilibrium does not exist and $\alpha^m \geq 0$ and $\alpha^f < 0$
c_3	$P(0,1)$ in the case where an equilibrium does not exist and $\alpha^m \geq 0$ and $\alpha^f < 0$
c_4	$P(1,1)$ in the case where an equilibrium does not exist and $\alpha^m \geq 0$ and $\alpha^f < 0$
d_1	$P(0,0)$ in the case where an equilibrium does not exist and $\alpha^m < 0$ and $\alpha^f \geq 0$
d_2	$P(1,0)$ in the case where an equilibrium does not exist and $\alpha^m < 0$ and $\alpha^f \geq 0$
d_3	$P(0,1)$ in the case where an equilibrium does not exist and $\alpha^m < 0$ and $\alpha^f \geq 0$
d_4	$P(1,1)$ in the case where an equilibrium does not exist and $\alpha^m < 0$ and $\alpha^f \geq 0$

Each run of the simulation proceeds as follows:

1. Draw 12 Uniform(0,1) random variables labelled u_i, i=1,...,12.

2. Calculate the weights from these random variables as follows:

   $a_i = \frac{u_i}{u_1+u_2}, i = 1,2,\ b_i = \frac{u_{i+2}}{u_3+u_4}, i = 1,2,\ c_i = \frac{u_{i+4}}{u_5+u_6+u_7+u_8}, i = 1, \dots, 4$ and $d_i = \frac{u_{i+8}}{u_9+u_{10}+u_{11}+u_{12}}, i = 1, \dots, 4.$

3. Using these weights, estimate the Nash model with the same specification detailed in Section 5 (i.e. same explanatory variables and data set).
4. Save the estimated parameters and log-likelihood value along with the weights used into a dataset.

7.1.2 Results

The process described in Section 7.1.1 was repeated 500 times and the results saved, resulting in a total of 476 simulations being completed successfully. In the other 24 cases, the maximum likelihood algorithm used in *Stata* failed to converge due to *Stata* being unable to calculate the numerical derivatives. While it may be possible to estimate the Nash model in these cases by coding the expressions for the derivatives in *Stata* and using an alternative maximization algorithm, this approach was not pursued. Some key results from these 476 simulations are summarised in Table 7-2 and Table 7-3, with the conditional probabilities having the same definition as that used in Section 5.4. In addition to these results, the estimated coefficients on the explanatory variables for each run were also stored. These have not been reported since there was generally little variation in the estimated coefficients across runs.

Table 7-2 – Summary statistics for key parameters in simulation

Variable	Original estimate	Mean	Std. Dev.	Min	Max
$P(y^f = 1 \| y^m = 0)$	0.2034	0.2026	0.0006	0.2010	0.2035
$P(y^f = 1 \| y^m = 1)$	0.3834	0.3929	0.0073	0.3802	0.4016
$P(y^m = 1 \| y^f = 0)$	0.1483	0.1471	0.0008	0.1452	0.1483
$P(y^m = 1 \| y^f = 1)$	0.3014	0.3083	0.0053	0.2974	0.3138
Log-likelihood	-2934.85	-2935.01	0.9504	-2936.89	-2933.61
ρ	-0.9298	-0.5197	0.3195	-0.9823	-0.0824
α^f	1.2357	0.7487	0.3868	0.1441	1.3680
α^m	1.4537	1.0143	0.3406	0.6596	1.6302

* Based on 476 observations, Orig. estimate refers to estimate with equal probability weights

Table 7-3 – Correlation coefficients for key parameters in simulation

	ρ	α^f	α^m
ρ	1	-0.9976	-0.9874
α^f	-0.9976	1	0.9888
α^m	-0.9874	0.9888	1

7.1.3 Discussion of results

From Table 7-2 it is clear that varying the weights used has little impact on some of the substantive inferences from the model, in particular the conditional probabilities of volunteering for the male and female, since there is very little variation in the estimates across runs. In addition, the small amount of variation in the estimated coefficients on the explanatory variables across runs implies that there is little change in the inferences associated with the explanatory variables. Also, as would be expected the original estimates are between the minimum and maximum values attained in the simulation.

As a result it appears that the main impact of varying the weights is on the separation of the endogenous and correlated effects in the model as measured by ρ, α^f and α^m. Table 7-2 indicates that there is substantial variation in each of these three parameters across the simulations, however in all the simulations α^f and α^m are both positive while ρ remains negative. This indicates that in all the simulations there is evidence of a positive endogenous effect of volunteering. An individual's partner volunteering directly increases their own likelihood of volunteering, however there is some uncertainty regarding the magnitude of this endogenous effect. The fact ρ remains negative across all the simulations indicates that the observed positive correlation in volunteering cannot be attributed to correlation in the unobservables.

While it seems striking that despite the substantial variation in the estimates of ρ, α^f and α^m across the simulations there is little variation in the substantive inferences from the model, this can be explained by considering the relationship between changes in ρ, α^f and α^m across simulations as detailed in Table 7-3. The large negative correlations between ρ/α^f and ρ/α^m across simulations indicate that decreases in ρ are associated with increases in α^f and α^m. This implies that as the estimated size of the positive endogenous effect increases, the magnitude of the negative correlation between the unobservables increases to offset this change. As such, the simulation study indicates that the only major impact of setting the weights is on the separation of the endogenous and correlated effects.

7.1.4 Simulation associated with Pareto Nash model

While the simulation exercise has only been undertaken for the Nash model, it is the case that multiple equilibria may also arise in the Pareto Nash model. As discussed in Section 4.5, this issue was worked around by imposing an equilibrium selection rule in the case of multiple equilibria with equal weights. As a result a similar sensitivity analysis could be undertaken for the Pareto Nash model, however a priori we would expect this is likely to indicate that there are no changes in any of the substantive inferences from the model. This is because for the estimated values of the structural parameters, multiple equilibria do not arise and as such for changing the weights to have any impact on the model it must cause the sign of either

$\alpha_0^f - \alpha_1^f$ or $\alpha_0^m - \alpha_1^m$ to change. This seems unlikely given that the estimate of this difference in all the models and in the sensitivity analysis has been negative.

To illustrate why this is the case, consider Table 4-11 from Section 4.5. For multiple equilibria to occur it must be the case that either $U_{00}^f < U_{01}^f < U_{11}^f < U_{10}^f$ or $U_{00}^m < U_{10}^m < U_{11}^m < U_{01}^m$. Since using Equation (4-2) we have $U_{11}^f - U_{01}^f = \alpha_1^f$, $U_{10}^f - U_{00}^f = \alpha_0^f$, $U_{11}^m - U_{10}^m = \alpha_1^m$ and $U_{01}^m - U_{00}^m = \alpha_0^m$, it follows that $U_{00}^f < U_{01}^f < U_{11}^f < U_{10}^f \Rightarrow \alpha_1^f < \alpha_0^f$ and $U_{00}^m < U_{10}^m < U_{11}^m < U_{01}^m \Rightarrow \alpha_1^m < \alpha_0^m$. However, the estimated values of the structural parameters in the Pareto Nash model imply that $\alpha_0^f < \alpha_1^f$ and $\alpha_0^m < \alpha_1^m$, which illustrates the claim that the sign of this difference must change.

7.2 Sub sampling the dataset

Some exploratory work regarding estimating the models for different subsets of the data was carried out. This was motivated by the possibility that different models may be more appropriate for different subsets of the data and there may be differences in the effects of the explanatory variables for different groups. For instance, it may be the case that couples where both are retired behave differently to couples where both are working.

While there are a reasonably large number of couples in the sample, the uneven split across the four outcomes (see Table 3-2) means that for many subsets of the data there are very few observations in some of the categories despite still having a reasonable number of observations in total. As a result, even when using a substantially smaller number of explanatory variables than that detailed in Section 5.1, many of the models failed to converge. In the cases when the models did converge the estimated effects of the explanatory variables were broadly similar along with there being a similar relationship in fit (as measured by the log-likelihood) between the models. The only substantial differences that arose were in the separation of the endogenous and correlated effects, i.e. the estimates for the structural parameters and the correlation coefficient.

7.3 Opportunities for further research

While a number of threats to the analysis have been identified and discussed throughout the thesis, there are a number of key areas for further research arising from this thesis. Firstly, as discussed in Section 3.4.2, there is a reasonable amount of persistence in volunteering behaviour over time. As such, it may be worthwhile exploring this aspect of volunteering behaviour and considering extensions of the models used in this thesis which capture this.

In addition, more modern estimation methods for the Nash model (introduced in Section 4.4.2) do not require the use of equilibrium selection rules. While the sensitivity analysis detailed in Section 7.1 assesses the impact of varying the equilibrium selection rule used, it may be useful to evaluate the impact of removing it altogether on the results.

Finally, as discussed in Section 4.5 it was necessary to impose reasonably strong assumptions to estimate the Pareto Nash model. It may be worthwhile considering ways of estimating the model without making these assumptions or assessing the likely impact of these assumptions on the estimation results.

8 Conclusion

Volunteering is rapidly becoming an area of substantial research interest. This is hardly surprising given recent comments by the Productivity Commission (2009b), 'If you count the contribution of 4.6 million volunteers, with an imputed value of $15 billion, this would make it a similar contribution to the retail industry.' Given the retail industry is one of the largest employers in Australia, this highlights the value of volunteering to the Australian community.

This thesis adds to the limited empirical evidence on volunteering in Australia by explicitly modelling interactions between couples in making volunteering decisions. This allows the reasons for the observed correlation in couple's volunteering decisions to be uncovered. It was found that there is strong evidence for the presence of a substantial endogenous effect of volunteering. That is, an individual's partner choosing to volunteer has a direct positive impact on their own likelihood of volunteering. This positive impact is estimated to be a 0.15 increase in probability for males whose partners volunteer and a 0.20 increase for females whose partners volunteer.

This is particularly useful from a policy perspective since it indicates that a strong multiplier effect is present. This arises because successful policies which target specific groups of people and encourage them to volunteer will have the added benefit of encouraging their partners to volunteer. This may substantially enhance the benefit of any Government policies to increase volunteering.

In addition, this thesis finds that volunteering is positively associated with a number of observable factors. Volunteers are generally more educated, have children, are aged around 50 and attend religious services. These findings are similar to those found in overseas studies discussed in this thesis.

A number of methods have been used in this thesis to assess the sensitivity of these results. In particular, in order to separate the different reasons for the observed correlation in couple's volunteering decisions it is necessary to make assumptions about the nature of interactions between couples. To assess the impact of these assumptions a variety of models have been used in this thesis which assume couples interact in different way. The key results are robust across these models providing some evidence that they are either not affected by these assumptions or all the models are providing a reasonable approximation to reality. In addition, methods were used to assess the impact of the equilibrium selection rule used in the Nash model. Once again, the key results were found to be relatively stable when varying the equilibrium selection rule.

However, there are some possible threats to the conclusions found in this thesis, which provide opportunities for further research. Firstly, as has been discussed, it was necessary to impose strong assumptions to estimate the Pareto Nash model which may have a substantial impact on the estimation results. Also, the models which have been used are not able to account for the persistence in volunteering behaviour which has been found nor do they allow investigation of the number of hours individuals volunteer. Moreover, while it seems unlikely that income and hours worked are endogenous, it has not been formally tested in the models that have been used.

In addition, it may be possible to gain substantially more information with a dataset that has a greater focus on volunteering. In particular, there is some evidence to suggest that with greater prompting and focus on volunteering in the survey, individuals are more likely to recall their volunteering activities. As such, there is some possibility of measurement error in the HILDA dataset due to the fact volunteering only forms a very small part of the survey.

There is also some possibility of omitted variable bias since a small number of variables identified in previous studies as important are not available in the HILDA dataset. For example, Freeman (1997) finds that being asked to volunteer is an important determinant of volunteering behaviour and this is not available in HILDA.

Despite these possible shortcomings, this study clearly provides a logical first step in improving understanding of volunteering behaviour in Australia. It has demonstrated that a number of relationships between observable characteristics and volunteering behaviour found in overseas studies also hold in Australian data. Its key contributions lie in the finding of a large endogenous effect of volunteering and in providing an additional application of behavioural models in empirical economics. This thesis demonstrates that these models can be particularly useful in separating various justifications for there being observed correlation in decisions made by couples. In addition, it has provided some exploration of the estimation methods used for these models.

Overall, volunteering is clearly a fruitful area for further research. Little is understood about this important activity despite its clear value to the Australian community. For instance, there is almost no research available on the demand for volunteers since engaging volunteers is not necessarily zero cost for non-profit organisations. As such, the intention of this thesis is to provide some initial understanding of volunteering from an empirical perspective and add to the knowledge base for future research. While it has uncovered some of the motivations for individuals to volunteer, it is clear that much work is required from a theoretical perspective to comprehend volunteering behaviour.

9 Appendix

9.1 Variable Definitions

Variable	Definition
aboriginal_torres	1 if individual is of Aboriginal or Torres Strait Islander origin, 0 otherwise
arrived_less_than_10y	1 if individual first came to Australia to live less than 10 years ago (defined as (fanyoa-1996)>0), 0 otherwise
other_urban	1 if individual's Section of State is other urban, 0 otherwise
rural_balance	1 if individual's Section of State is rural balance, 0 otherwise
bounded_locality	1 if individual's Section of State is bounded locality, 0 otherwise
Base category for above dummy variables is individual's Section of State is Major Urban	
bachelor	1 if individual's highest level of education is a bachelors degree, 0 otherwise
post_school	1 if individual's highest level of education is an Advanced Diploma, Diploma or Certificate, 0 otherwise
postgrad	1 if individual's highest level of education is a masters, doctorate, graduate diploma or graduate certificate, 0 otherwise
yr12	1 if individual's highest level of education is Year 12, 0 otherwise
Base category for above dummy variables is individual's highest level of education is below Year 12	
child_under_10	1 if age of individual's youngest own child is less than 10, 0 otherwise
health_excellent	1 if individual's Self-assessed health is excellent, 0 otherwise
health_fair	1 if individual's Self-assessed health is fair, 0 otherwise
health_good	1 if individual's Self-assessed health is good, 0 otherwise
health_poor	1 if individual's Self-assessed health is poor, 0 otherwise
health_verygood	1 if individual's Self-assessed health is very good, 0 otherwise
Base category for self-assessed health dummy variables is individual refused/not stated	
long_term_hlth_cond	1 if individual has a long-term health condition which means they can't work or limits the type or amount of work they can do, 0 otherwise
age	Inidivdual's age last birthday at date of interview
age_sq	Individual's age last birthday at date of interview squared
hh_gross_inc	Household financial year gross income (excluding windfall) in thousands (missing values are imputed and variable is top-coded)
yr_curr_addr	Years at current address
yr_curr_addr_missing	1 if years at current address is missing (in this case yr_curr_addr=0), 0 otherwise
hh_home_equity	Household Home Equity in thousands (missing values are imputed and variable is top-coded)
hh_net_worth	Household Net Worth in thousands (missing values are imputed and variable is top-coded)
fixedterm_casual	1 if individual is employed on a fixed-term contract or casual basis, 0 otherwise
hours_worked	Hours individual usually works per week in all jobs
owns_home	1 if individual owns or is buying own home or any other residential property, 0 otherwise

num_child	Number of own resident children
ind_disp_income	Individual’s financial year disposable income in thousands (missing values are imputed and the variable is top-coded)
hours_vary	1 if individual’s hours vary (in this case hours_worked=0), 0 otherwise
irregular_hours	1 if individual works a rotating shift, split shift, is on call, has an irregular schedule or some other irregular schedule, 0 otherwise
likely_to_move	1 if individual is likely or very likely to move in the next 12 months, 0 otherwise
more_than_one_job	1 if individual currently has more than one job, 0 otherwise
office_worker	1 if individual is currently employed as an office worker (derived from fjbmocc1), 0 otherwise
overseasborn_english	1 if individual is born outside Australia in Main English speaking country, 0 otherwise
overseasborn_other	1 if individual is born outside Australia in a country which is not Main English speaking, 0 otherwise
Base category for above two dummy variables is Australian born	
pressed_for_time	1 if individual often/almost always feels pushed for time (flsrush equals 1 or 2), 0 otherwise
religion	1 if individual often or very often makes time to attend services at a place of worship, 0 otherwise
retire_in_5y	1 if individual plans to retire within the next 5 years (defined as frtiage1-age<5 & frtiage1-age>=0), 0 otherwise
retired	1 if individual has completely retired from paid work or was never in paid work but considers themselves retired, 0 otherwise
satisfied_balance_work	1 if individual is satisfied with the flexibility to balance work and non-work commitments (fjbmsflx>5), 0 otherwise
state_act	1 if individual lives in Australian Capital Territory, 0 otherwise
state_nt	1 if individual lives in Northern Territory, 0 otherwise
state_qld	1 if individual lives in Queensland, 0 otherwise
state_sa	1 if individual lives in South Australia, 0 otherwise
state_tas	1 if individual lives in Tasmania, 0 otherwise
state_vic	1 if individual lives in Victoria, 0 otherwise
state_wa	1 if individual lives in Western Australia, 0 otherwise
Base category for above state dummy variables is individual lives in New South Wales	
unemployed	1 if individual is unemployed and looking for full time or part time work, 0 otherwise
volunteer	1 if individual undertakes 1 or more hours of volunteer or charity work, 0 otherwise
worried_about_job	1 in individual worries about the future of their job (defined as fjomwf>4), 0 otherwise

9.2 Stata code for Log-likelihood function for Stackelberg leader model

```
program stackelberg_lf
      version 10.0
      args lnf xb1 xb2 alpha1 alpha10 alpha2 rho
      quietly replace `lnf' = ln( binormal( -`xb1', -`xb2', `rho') - ///
            ( binormal( -`xb1', -`xb2', `rho' ) - ///
            binormal( -`xb1'-`alpha1', -`xb2', `rho' ) - ///
```

```
        binormal( -`xb1', -`xb2'-`alpha2', `rho' ) + ///
        binormal( -`xb1'-`alpha1', -`xb2'-`alpha2', `rho' ) )) ///
        if $ML_y1==0 & $ML_y2==0 & `alpha2'>=0
    quietly replace `lnf' = ln( binormal( -`xb1', -`xb2', `rho') ) ///
        if $ML_y1==0 & $ML_y2==0 & `alpha2'<0

    quietly replace `lnf' = ln( binormal( -`xb1' - `alpha1' + ///
        `alpha10', `xb2', -`rho') ) ///
        if $ML_y1==0 & $ML_y2==1 & `alpha2'>=0
    quietly replace `lnf' = ln( binormal( -`xb1' - `alpha1' + ///
        `alpha10', `xb2', -`rho') + binormal( -`xb1'+`alpha10', ///
        -`xb2'-`alpha2', `rho' ) - binormal( -`xb1'+`alpha10'-///
        `alpha1', -`xb2'-`alpha2', `rho' ) - ///
        binormal( -`xb1'+`alpha10', -`xb2', `rho' ) + ///
        binormal( -`xb1'+`alpha10'-`alpha1', -`xb2', `rho' ) ) ///
        if $ML_y1==0 & $ML_y2==1 & `alpha2'<0

    quietly replace `lnf' = ln( binormal( `xb1', -`xb2' - ///
        `alpha2', -`rho') ) if $ML_y1==1 & $ML_y2==0 & `alpha2'>=0
    quietly replace `lnf' = ln( binormal( `xb1', -`xb2' - ///
        `alpha2', -`rho') + binormal( -`xb1', -`xb2'-`alpha2', ///
        `rho') - binormal( -`xb1'+`alpha10', -`xb2'-`alpha2', ///
        `rho') - binormal( -`xb1', -`xb2', `rho') + ///
        binormal( -`xb1'+`alpha10', -`xb2', `rho') ) ///
        if $ML_y1==1 & $ML_y2==0 & `alpha2'<0

    quietly replace `lnf' = ln( binormal( `xb1' + `alpha1' - ///
        `alpha10', `xb2' + `alpha2', `rho') - ///
        ( binormal( -`xb1'-`alpha1', -`xb2', `rho' ) - ///
        binormal( -`xb1'-`alpha1'+`alpha10', -`xb2', `rho' ) -///
        binormal( -`xb1'-`alpha1', -`xb2'-`alpha2', `rho' ) + ///
        binormal( -`xb1'-`alpha1'+`alpha10', -`xb2'-`alpha2', `rho' ///
        ) ) ) if $ML_y1==1 & $ML_y2==1 & `alpha2'>=0
    quietly replace `lnf' = ln( binormal( `xb1' + `alpha1' - ///
        `alpha10', `xb2' + `alpha2', `rho') ) ///
        if $ML_y1==1 & $ML_y2==1 & `alpha2'<0

end
```

9.3 Stata code for Log-likelihood function for Nash model

```
#delimit ;

capture program drop nash_lf;
program nash_lf;
    version 10.0;
    args lnf xb1 xb2 alpha1 alpha2 rho;
* Note that when a Nash equilibrium does not exist the probability is
evenly distributed over all possibilities;
* When a Nash equilibrium is not unique, the probability is distributed
evenly over all possibilities;


    quietly replace `lnf' = ln( binormal( -`xb1', -`xb2', `rho') –
        0.5 * ( binormal( -`xb1', -`xb2', `rho' ) –
        binormal( -`xb1', -`xb2' -`alpha2', `rho' ) –
        binormal( -`xb1' - `alpha1', -`xb2', `rho' ) +
        binormal( -`xb1' - `alpha1', -`xb2' - `alpha2', `rho' ) )  )
        if `alpha1' >= 0 & `alpha2' >= 0 & $ML_y1==0 & $ML_y2==0 ;
    quietly replace `lnf' = ln( binormal( -`xb1', -`xb2', `rho') +
        0.25 * ( binormal( -`xb1', -`xb2' -`alpha2', `rho' ) -
        binormal( -`xb1', -`xb2', `rho' ) +
```

```
      binormal( -`xb1' - `alpha1', -`xb2', `rho' ) -
      binormal( -`xb1' - `alpha1', -`xb2' - `alpha2', `rho' ) ) )
      if `alpha1' >= 0 & `alpha2' < 0 & $ML_y1==0 & $ML_y2==0 ;
quietly replace `lnf' = ln( binormal( -`xb1', -`xb2', `rho') +
      0.25 * ( binormal( -`xb1', -`xb2' -`alpha2', `rho' ) -
      binormal( -`xb1', -`xb2', `rho' ) +
      binormal( -`xb1' - `alpha1', -`xb2', `rho' ) -
      binormal( -`xb1' - `alpha1', -`xb2' - `alpha2', `rho' ) ) )
      if `alpha1' < 0 & `alpha2' >= 0 & $ML_y1==0 & $ML_y2==0 ;
quietly replace `lnf' = ln( binormal( -`xb1', -`xb2', `rho') )
      if `alpha1' < 0 & `alpha2' < 0 & $ML_y1==0 & $ML_y2==0 ;


quietly replace `lnf' = ln( binormal( -`xb1'-`alpha1', `xb2', -`rho')
       ) if `alpha1' >= 0 & `alpha2' >= 0 & $ML_y1==0 & $ML_y2==1 ;
quietly replace `lnf' = ln( binormal( -`xb1'-`alpha1', `xb2', -`rho')
      + 0.25 * ( binormal( -`xb1', -`xb2' -`alpha2', `rho' ) -
      binormal( -`xb1', -`xb2', `rho' ) +
      binormal( -`xb1' - `alpha1', -`xb2', `rho' ) -
      binormal( -`xb1' - `alpha1', -`xb2' - `alpha2', `rho' ) ) )
      if `alpha1' >= 0 & `alpha2' < 0 & $ML_y1==0 & $ML_y2==1 ;
quietly replace `lnf' = ln( binormal( -`xb1'-`alpha1', `xb2', -`rho')
      + 0.25 * ( binormal( -`xb1', -`xb2' -`alpha2', `rho' ) -
      binormal( -`xb1', -`xb2', `rho' ) +
      binormal( -`xb1' - `alpha1', -`xb2', `rho' ) -
      binormal( -`xb1' - `alpha1', -`xb2' - `alpha2', `rho' ) ) )
      if `alpha1' < 0 & `alpha2' >= 0 & $ML_y1==0 & $ML_y2==1 ;
quietly replace `lnf' = ln( binormal( -`xb1'-`alpha1', `xb2', -`rho')
      - 0.5 * ( binormal( -`xb1', -`xb2', `rho' ) -
      binormal( -`xb1', -`xb2' -`alpha2', `rho' ) –
      binormal( -`xb1' - `alpha1', -`xb2', `rho' ) +
      binormal( -`xb1' - `alpha1', -`xb2' - `alpha2', `rho' ) ) )
      if `alpha1' < 0 & `alpha2' < 0 & $ML_y1==0 & $ML_y2==1 ;


quietly replace `lnf' = ln( binormal( `xb1',
      -`xb2' - `alpha2', -`rho') )
      if `alpha1' >= 0 & `alpha2' >= 0 & $ML_y1==1 & $ML_y2==0 ;
quietly replace `lnf' = ln( binormal( `xb1', -`xb2' - `alpha2',
      -`rho') + 0.25 * ( binormal( -`xb1', -`xb2' -`alpha2', `rho' )
      - binormal( -`xb1', -`xb2', `rho' ) +
      binormal( -`xb1' - `alpha1', -`xb2', `rho' ) -
      binormal( -`xb1' - `alpha1', -`xb2' - `alpha2', `rho' ) ) )
      if `alpha1' >= 0 & `alpha2' < 0 & $ML_y1==1 & $ML_y2==0 ;
quietly replace `lnf' = ln( binormal( `xb1', -`xb2' - `alpha2',
      -`rho') + 0.25 * ( binormal( -`xb1', -`xb2' -`alpha2', `rho' )
      - binormal( -`xb1', -`xb2', `rho' ) +
      binormal( -`xb1' - `alpha1', -`xb2', `rho' ) -
      binormal( -`xb1' - `alpha1', -`xb2' - `alpha2', `rho' ) ) )
      if `alpha1' < 0 & `alpha2' >= 0 & $ML_y1==1 & $ML_y2==0 ;
quietly replace `lnf' = ln( binormal( `xb1', -`xb2' - `alpha2',
      -`rho') - 0.5 * ( binormal( -`xb1', -`xb2', `rho' ) -
      binormal( -`xb1', -`xb2' -`alpha2', `rho' ) –
      binormal( -`xb1' - `alpha1', -`xb2', `rho' ) +
      binormal( -`xb1' - `alpha1', -`xb2' - `alpha2', `rho' ) ) )
      if `alpha1' < 0 & `alpha2' < 0 & $ML_y1==1 & $ML_y2==0 ;


quietly replace `lnf' = ln( binormal( `xb1' + `alpha1',
      `xb2' + `alpha2', `rho') - 0.5 * (
      binormal( -`xb1', -`xb2', `rho' ) -
      binormal( -`xb1', -`xb2' -`alpha2', `rho' ) –
      binormal( -`xb1' - `alpha1', -`xb2', `rho' ) +
```

```
            binormal( -`xb1' - `alpha1', -`xb2' - `alpha2', `rho' ) ) )
            if `alpha1' >= 0 & `alpha2' >= 0 & $ML_y1==1 & $ML_y2==1 ;
      quietly replace `lnf' = ln( binormal( `xb1' + `alpha1',
            `xb2' + `alpha2', `rho') + 0.25 * (
            binormal( -`xb1', -`xb2' -`alpha2', `rho' ) -
            binormal( -`xb1', -`xb2', `rho' ) +
            binormal( -`xb1' - `alpha1', -`xb2', `rho' ) -
            binormal( -`xb1' - `alpha1', -`xb2' - `alpha2', `rho' ) ) )
            if `alpha1' >= 0 & `alpha2' < 0 & $ML_y1==1 & $ML_y2==1 ;
      quietly replace `lnf' = ln( binormal( `xb1' + `alpha1',
            `xb2' + `alpha2', `rho') + 0.25 * (
            binormal( -`xb1', -`xb2' -`alpha2', `rho' ) -
            binormal( -`xb1', -`xb2', `rho' ) +
            binormal( -`xb1' - `alpha1', -`xb2', `rho' ) -
            binormal( -`xb1' - `alpha1', -`xb2' - `alpha2', `rho' ) ) )
            if `alpha1' < 0 & `alpha2' >= 0 & $ML_y1==1 & $ML_y2==1 ;
      quietly replace `lnf' = ln( binormal( `xb1' + `alpha1',
            `xb2' + `alpha2', `rho') )
            if `alpha1' < 0 & `alpha2' < 0 & $ML_y1==1 & $ML_y2==1 ;


end;
```

9.4 Stata code for Log-likelihood function for Pareto Nash model

```
#delimit ;

capture program drop pareto_nash_lf;
program pareto_nash_lf;
      version 10.0;
      args lnf xb1 xb2 alpha10 alpha11 alpha20 alpha21 rho;

      quietly replace `lnf' = ln( binormal( `xb1' - `alpha10' + `alpha11',
            `xb2' - `alpha20' + `alpha21', `rho' ) +
            ( binormal( -`xb1', -`xb2', `rho' ) –
            binormal( -`xb1', -`xb2' -`alpha21', `rho' ) -
            binormal( -`xb1' - `alpha11', -`xb2', `rho' ) +
            binormal( -`xb1' - `alpha11', -`xb2' - `alpha21', `rho' ) ) )
            if `alpha10'-`alpha11'>=0 & `alpha20'-`alpha21'>=0 & $ML_y1==1
            & $ML_y2==1 ;
      quietly replace `lnf' = ln( binormal( `xb1' - `alpha10' + `alpha11',
            `xb2' - `alpha20' + `alpha21', `rho' ) +
            ( binormal( -`xb1', -`xb2', `rho' ) –
            binormal( -`xb1', -`xb2' -`alpha21', `rho' ) -
            binormal( -`xb1' - `alpha11', -`xb2', `rho' ) +
            binormal( -`xb1' - `alpha11', -`xb2' - `alpha21', `rho' ) ) +
            0.5*( binormal( -`xb1' + `alpha10' - `alpha11', -`xb2', `rho' )
            - binormal( -`xb1' + `alpha10' - `alpha11',
            -`xb2' + `alpha20' -`alpha21', `rho' ) -
            binormal( -`xb1', -`xb2', `rho' ) +
            binormal( -`xb1', -`xb2' + `alpha20' - `alpha21', `rho' ) ) )
            if `alpha10'-`alpha11'>=0 & `alpha20'-`alpha21'<0 & $ML_y1==1 &
            $ML_y2==1 ;
      quietly replace `lnf' = ln( binormal( `xb1' - `alpha10' + `alpha11',
            `xb2' - `alpha20' + `alpha21', `rho' ) +
            ( binormal( -`xb1', -`xb2', `rho' ) –
            binormal( -`xb1', -`xb2' -`alpha21', `rho' ) -
            binormal( -`xb1' - `alpha11', -`xb2', `rho' ) +
            binormal( -`xb1' - `alpha11', -`xb2' - `alpha21', `rho' ) ) +
            0.5*( binormal( -`xb1', -`xb2' + `alpha20' - `alpha21', `rho' )
```

```
      - binormal( -`xb1', -`xb2', `rho' ) -
      binormal( -`xb1' + `alpha10' - `alpha11',
      -`xb2' + `alpha20' - `alpha21', `rho' ) +
      binormal( -`xb1' + `alpha10' - `alpha11', -`xb2', `rho' ) ) )
      if `alpha10'-`alpha11'<0 & `alpha20'-`alpha21'>=0 & $ML_y1==1 &
      $ML_y2==1 ;
quietly replace `lnf' = ln( binormal( `xb1' - `alpha10' + `alpha11',
      `xb2' - `alpha20' + `alpha21', `rho' ) +
      ( binormal( -`xb1', -`xb2' + `alpha20' - `alpha21', `rho' ) –
      binormal( -`xb1' - `alpha11', -`xb2' + `alpha20' - `alpha21',
      `rho' ) - binormal( -`xb1', -`xb2' - `alpha21', `rho' ) +
      binormal( -`xb1' - `alpha11', -`xb2' - `alpha21', `rho' ) ) +
      ( binormal( -`xb1' - `alpha11' + `alpha10', -`xb2', `rho' ) -
      binormal( -`xb1' - `alpha11' + `alpha10',
      -`xb2' + `alpha20' - `alpha21', `rho' ) -
      binormal( -`xb1' - `alpha11', -`xb2', `rho' ) +
      binormal( -`xb1' - `alpha11', -`xb2' + `alpha20' - `alpha21',
      `rho' ) ) )
      if `alpha10'-`alpha11'<0 & `alpha20'-`alpha21'<0 & $ML_y1==1 &
      $ML_y2==1 ;

quietly replace `lnf' = ln( binormal( -`xb1' + `alpha10' - `alpha11',
      `xb2', - `rho' ) -
      0.5 * ( binormal( -`xb1' + `alpha10' - `alpha11',
      -`xb2' + `alpha20' - `alpha21', `rho' ) -
      binormal( -`xb1' + `alpha10' - `alpha11', -`xb2', `rho' ) –
      binormal( -`xb1', -`xb2' + `alpha20' - `alpha21', `rho' ) +
      binormal( -`xb1', -`xb2', `rho' ) ) )
      if `alpha10'-`alpha11'>=0 & `alpha20'-`alpha21'>=0 & $ML_y1==0
      & $ML_y2==1 ;
quietly replace `lnf' = ln( binormal( -`xb1' + `alpha10' - `alpha11',
      `xb2', -`rho' ) + 0.5 *
      ( binormal( -`xb1' + `alpha10' - `alpha11', -`xb2', `rho' ) –
      binormal( -`xb1', -`xb2', `rho' ) -
      binormal( -`xb1'  + `alpha10' - `alpha11',
      -`xb2' + `alpha20' - `alpha21', `rho' )
      + binormal( -`xb1', -`xb2' + `alpha20' - `alpha21', `rho' ) ) )
      if `alpha10'-`alpha11'>=0 & `alpha20'-`alpha21'<0 & $ML_y1==0 &
      $ML_y2==1 ;
quietly replace `lnf' = ln( binormal( -`xb1' + `alpha10' - `alpha11',
      `xb2', -`rho' ) )
      if `alpha10'-`alpha11'<0 & `alpha20'-`alpha21'>=0 & $ML_y1==0 &
      $ML_y2==1 ;
quietly replace `lnf' = ln( binormal( -`xb1' + `alpha10' - `alpha11',
      `xb2', -`rho' ) )
      if `alpha10'-`alpha11'<0 & `alpha20'-`alpha21'<0 & $ML_y1==0 &
      $ML_y2==1 ;

quietly replace `lnf' = ln( binormal( `xb1',
      -`xb2' + `alpha20' – `alpha21', -`rho' ) -
      0.5 * ( binormal( -`xb1' + `alpha10' - `alpha11',
      -`xb2' + `alpha20' - `alpha21', `rho' ) -
      binormal( -`xb1' + `alpha10' - `alpha11', -`xb2', `rho' ) –
      binormal( -`xb1', -`xb2' + `alpha20' - `alpha21', `rho' )
      + binormal( -`xb1', -`xb2', `rho' ) ) )
      if `alpha10'-`alpha11'>=0 & `alpha20'-`alpha21'>=0 & $ML_y1==1
      & $ML_y2==0 ;
quietly replace `lnf' = ln( binormal( `xb1',
      - `xb2' + `alpha20' - `alpha21', -`rho') )
      if `alpha10'-`alpha11'>=0 & `alpha20'-`alpha21'<0 & $ML_y1==1 &
      $ML_y2==0 ;
quietly replace `lnf' = ln( binormal( `xb1',
      - `xb2' + `alpha20' - `alpha21', -`rho' ) +
```

```
        0.5*( binormal( -`xb1', -`xb2' + `alpha20' - `alpha21', `rho' )
        - binormal( -`xb1' + `alpha10' - `alpha11',
        -`xb2' + `alpha20' - `alpha21', `rho' ) -
        binormal( -`xb1', -`xb2', `rho' ) +
        binormal( -`xb1' + `alpha10' - `alpha11', -`xb2', `rho' ) ) )
        if `alpha10'-`alpha11'<0 & `alpha20'-`alpha21'>=0 & $ML_y1==1 &
        $ML_y2==0 ;
    quietly replace `lnf' = ln( binormal( `xb1',
        -`xb2' + `alpha20' - `alpha21', -`rho' ) )
        if `alpha10'-`alpha11'<0 & `alpha20'-`alpha21'<0 & $ML_y1==1 &
        $ML_y2==0 ;

    quietly replace `lnf' = ln( binormal( -`xb1', -`xb2', `rho' ) -
        ( binormal( -`xb1', -`xb2', `rho' ) –
        binormal( -`xb1', -`xb2' - `alpha21', `rho' ) -
        binormal( -`xb1' - `alpha11', -`xb2', `rho' ) +
        binormal( -`xb1' - `alpha11', -`xb2' - `alpha21', `rho' ) ) )
        if $ML_y1==0 & $ML_y2==0 ;

end;
```

10 References

Andreoni, J. 1990, 'Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving', *The Economic Journal*, vol. 100, no. 401, pp. 464-477.

Andreoni, J., Brown, E. & Rischall, I. 2003, 'Charitable Giving by Married Couples: Who Decides and Why Does it Matter?', *Journal of Human Resources*, vol. 38, no. 1, pp. 111-133.

Andreoni, J., Gale, W. & Scholz, J. July 1996, 'Charitable Contributions of Time and Money', Working Paper, <http://econ.ucsd.edu/~jandreon/WorkingPapers/ags-v8.pdf>.

Andrews, D.W.K., Berry, S. & Jia, P. May 2004, 'Confidence Regions for Parameters in Discrete Games with Multiple Equilibria, with an application to Discount Store Chain Location', Working Paper, <http://www.econ.yale.edu/~steveb/abj.pdf>.

Australian Bureau of Statistics (ABS). 2001, *Voluntary Work, Australia, 2000* accessed 8 October 2009, <http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/4441.0Explanatory%20Notes12000?OpenDocument>.

Australian Bureau of Statistics (ABS). 2007, *Voluntary Work, Australia, 2006* accessed 30 April 2009, <http://www.abs.gov.au/AUSSTATS/abs@.nsf/Latestproducts/4441.0Main%20Features22006?opendocument&tabname=Summary&prodno=4441.0&issue=2006&num=&view>.

Bajari, P., Hong, H. & Ryan, S. 2007, ‘Identification and Estimation of a Discrete Game of Complete Information’, Working Paper, <http://www.stanford.edu/~doubleh/papers/game5-21-07-stephen.pdf>.

Berry, S. & Reiss, P. 2006, ‘Empirical Models of Entry and Market Structure’, Working Paper, <http://www.stanford.edu/~preiss/hand-entry.pdf>.

Bjorn, P. & Vuong, Q. 1984, 'Simultaneous models for dummy endogenous variables: a game theoretic formulation with an application to household labor force participation', Working paper No. 537, California Institute of Technology.

Bjorn P. & Vuong, Q. 1985, 'Econometric modeling of a Stackelberg game with an application to household labor force participation', Working paper No. 577, California Institute of Technology.

Bowman, W. 2009, ‘The economic value of volunteers to nonprofit organizations’, *Nonprofit Management and Leadership*, vol. 19, no. 4, pp. 491-506.

Bresnahan, T. & Reiss, P. 1991, 'Empirical models of discrete games', *Journal of Econometrics*, vol. 48, no. 1, pp. 57-81.

Brown, E. & Lankford, H. 1992, ‘Gifts of money and gifts of time estimating the effects of tax prices and available time’, *Journal of Public Economics*, vol. 47, no. 3, pp. 321-341.

Bryant, W., Jeon-Slaughter, H., Kang, H. & Tax, A. 2003, ‘Participation in Philanthropic Activities: Donating Money and Time’, *Journal of Consumer Policy*, vol. 26, no. 1, pp. 43-73.

Cameron, T. & Quiggin, J. 1994, ‘Estimation using contingent valuation data from a ‘‘dichotomous choice with follow-up’’ questionnaire’, *Journal of Environmental Economics and Management*, vol. 27, no. 3, pp. 218-234.

Carlin, P. 2001, ‘Evidence on the Volunteer Labor Supply of Married Women’, *Southern Economic Journal*, vol. 67, no. 4, pp. 801-824.

Chao, L. 2002, ‘A comparison of consensus and nonconsensus approaches to modeling contraceptive choice behavior’, *Health Economics*, vol. 11, no. 7, pp. 599-622.

Cugno, F. & Ferrero, M. 2004, ‘Competition among volunteers’, *European Journal of Political Economy*, vol. 20, no. 3, pp. 637-654.

Day, K. & Devlin, R. 1996, ‘Volunteerism and Crowding out: Canadian Econometric Evidence’, *The Canadian Journal of Economics*, vol. 29, no. 1, pp. 37-53.

Day, K. & Devlin, R. 1997, ‘Can volunteer work help explain the male female earnings gap?’, *Applied Economics*, vol. 29, no. 6, pp. 707-721.

Day, K. & Devlin, R. 1998, ‘The Payoff to Work without Pay: Volunteer Work as an Investment in Human Capital’, *The Canadian Journal of Economics*, vol. 31, no. 5, pp. 1179-1191.

Dercon, S. & Krishnan, P. 2000, ‘In Sickness and in Health: Risk Sharing within Households in Rural Ethiopia’, *Journal of Political Economy*, vol. 108, no. 4, pp. 688-727.

Duncan, B. 1999, ‘Modelling charitable contributions of time and money’, *Journal of Public Economics*, vol. 72, no. 2, pp. 213–242.

Freeman, R. 1997, ‘Working for Nothing: The Supply of Volunteer Labor’, *Journal of Labor Economics*, vol. 15, no. 1, pp. 140-166.

Gomez, R. & Gunderson, M. 2003, ‘Volunteer Activity and the Demands of Work and Family’, *Industrial Relations*, vol. 58, no. 4, pp. 573-589.

Haab, T. 1998, ‘Estimation using contingent valuation data from a ‘‘dichotomous choice with follow-up’’ questionnaire: A Comment’, *Journal of Environmental Economics and Management*, vol. 35, no. 2, pp. 190-194.

Hackl, F., Halla, M. & Pruckner, G. 2007, ‘Volunteering and Income – The Fallacy of the Good Samaritan?’, *KYKLOS*, vol. 60, no. 1, pp. 77–104.

Heckman, J. 1978, ‘Dummy Endogenous Variables in a Simultaneous Equation System’, *Econometrica*, vol. 46, no. 4, pp. 931-959.

Ironmonger, D. 2000, ‘Measuring Volunteering in Economic Terms: 2,200 Million Hours Worth $42 Billion a Year’, Chapter 5, pp. 56-72, in J. Warburton and M. Oppenheimer (eds), Volunteers and Volunteering, Federation Press, Leichardt.

Jia, Z. 2005, ‘Labor Supply of Retiring Couples and Heterogeneity in Household Decision-Making Structure’, *Review of Economics of the Household*, vol. 3, no. 2, pp. 215-233.

John Hopkins University, 2008, *Mapping volunteer work around the world*, accessed 8 October 2009, <http://www.ccss.jhu.edu/pdfs/ILO/ILO_2008_12_19_NewsRelease_ICLS_support.pdf>.

Katz, E. & Rosenberg, J. 2005, 'An economic interpretation of institutional volunteering', *European Journal of Political Economy*, vol. 21, no. 2, pp. 429-443.

Keane, M. 1992, 'A Note on Identification in the Multinomial Probit Model', *Journal of Business & Economic Statistics*, vol. 10, no. 2, pp. 193-200.

Kooreman, P. 1994, 'Estimation of Econometric Models of Some Discrete Games', *Journal of Applied Econometrics*, vol. 9, no. 3, pp. 255-268.

Manski, C. 1993, 'Identification of Endogenous Social Effects: The Reflection Problem', *The Review of Economic Studies*, vol. 60, no. 3, pp. 531-542.

Manski, C. & Tamer, E. 2002, 'Inference on Regressions with Interval Data on a Regressor or Outcome', *Econometrica*, vol. 70, no. 2, pp. 519-546.

Menchik, P. & Weisbrod, B. 1987, 'Volunteer Labor Supply', *Journal of Public Economics*, vol. 32, no. 2, pp. 159-183.

Mueller, M. 1975, 'Economic Determinants of Volunteer Work by Women', *Signs*, vol. 1, no. 2, pp. 325-338.

Polidano, C., Harris, M. & Zhao, X. 2009, 'Volunteering for work and pleasure?', Working Paper, <http://www.melbourneinstitute.com/hilda/Biblio/cp/Polidano_ESAM.pdf>.

Pollak, R. & Wales, T. 1991, 'The likelihood dominance criterion: A new approach to model selection', *Journal of Econometrics*, vol. 47, no. 2, pp. 227-242.

Productivity Commission, 2009a, *Contribution of the Not for Profit Sector (Current) – Productivity Commission,* accessed 9 October 2009, <http://www.pc.gov.au/projects/study/not-for-profit>.

Productivity Commission, 2009b, *Reforms Needed for the Not-for-Profit Sector – Productivity* Commission, accessed 16 October 2009, <http://www.pc.gov.au/projects/study/not-for-profit/draft/media-release>.

Prouteau, L. & Wolff, F. 2006, 'Does volunteer work pay off in the labor market?', *Journal of Socio-Economics*, vol. 35, no. 6, pp. 992-1013.

Schram, V. & Dunsing, M. 1981, 'Influences on Married Women's Volunteer Work Participation', *The Journal of Consumer Research*, vol. 7, no. 4, pp. 372-379.

Segal, L. & Weisbrod, B. 2002, 'Volunteer Labor Sorting across Industries', *Journal of Policy Analysis and Management*, vol. 21, no. 3, pp. 427- 447.

Tamer, E. 2003, 'Incomplete Simultaneous Discrete Response Model with Multiple Equilibria', *Review of Economic Studies*, vol. 70, no. 1, pp. 147-165.

Udry, C. 1996, 'Gender, Agricultural Production, and the Theory of the Household', *The Journal of Political Economy*, vol. 104, no. 5, pp. 1010-1046.

Unger, L. 1991, 'Altruism as a motivation to volunteer', *Journal of Economic Psychology*, vol. 12, no. 1, pp. 71-100.

United Nations, 2001, 'Marking end of International Year of Volunteers, General Assembly encourages all people to become more engaged in voluntary activities', in *Press Release GA/9990*, accessed 9 October 2009, <http://www.un.org/News/Press/docs/2001/GA9990.doc.htm>.

Vaillancourt, F. 1994, 'To Volunteer or Not: Canada, 1987', *The Canadian Journal of Economics*, vol. 27, no. 4, pp. 813-826.

Vermeulen, F. 2002, 'Collective Household Models: Principles and Main Results', *Journal of Economic Surveys*, vol. 16, no. 4, pp. 533-564.

Volunteering Australia, 2009, 'Contribution of the Not for Profit Sector', in *Submission – Productivity Commission*, accessed 8 October 2009, <http://www.pc.gov.au/__data/assets/pdf_file/0008/89504/sub047.pdf>.

Watson, N. 2008, 'HILDA User Manual – Release 6', Melbourne Institute of Applied Economic and Social Research, University of Melbourne.