WORKING
PAPER

A Collective Labor Supply Model: Identification and Estimation in the Presence of Externalities by Means of Panel Data

PIERRE-CARL MICHAUD
FREDERIC VERMEULEN

WR-406

July 2006

RAND LABOR AND POPULATION

A collective labor supply model: identification and estimation in the presence of externalities by means of panel data*

Pierre-Carl Michaud† and Frederic Vermeulen‡

June, 2006

Abstract

We study labor supply of elderly couples by means of a collective model. The model allows individuals to enjoy leisure more (or less) in company of their spouse (complementarity/externalities in leisure). Preferences and the intra-household bargaining process are identified by using panel data through the dissolution of the household due to the death of one of the partners. The model does not only look at the extensive margin (working versus being retired), but also at the intensive margin (how many hours are worked). We apply the model to American households coming from the first six waves of the Health and Retirement Study. We compare model simulations with those from a standard unitary model for a set of policy reforms; such as the widely discussed proposals to eliminate the earnings test and the replacement of the spouse benefit with a past earnings sharing mechanism.

Key words: collective model, intra-household allocation, labor supply, retirement, social security, identification.

JEL-classification: D13, H310, J22, J26.

---

*The authors acknowledge the financial support provided through the European Community's Human Potential Programme under contract HPRN-CT-2002-00235, (AGE). We are grateful to seminar participants at CREST, IZA, Louvain-la-Neuve, RAND, Tilburg, Utrecht, the RTN-AGE workshop in Edesheim and ESEM 2004 Madrid for useful comments and suggestions. We also thank Nicole Maestas for helpful discussions on social security rules. Part of this paper was written when the first author was at CentER and the Institute for Fiscal Studies London, whose kind hospitality is gratefully acknowledged.

†RAND Corporation and IZA Bonn, 1776 Main Street, P.O. Box 2138, Santa Monica CA 90407-2138, United States. E-mail: michaud@rand.org.

‡Tilburg University, CentER, Netspar and IZA Bonn, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. E-mail: frederic.vermeulen@uvt.nl.

1 Introduction

A burgeoning literature in microeconomics is that on intra-household decision-making models. The emergence of this field can at least be partly attributed to a growing discomfort with the standard unitary model, which assumes that households behave as single decision makers, irrespective of the number of individuals that form the households under study. One major problem with the unitary approach is that there is much empirical evidence that it does not fit the data very well. Slutsky symmetry and negativity are regularly rejected when confronted with consumption or labor supply data (see Fortin and Lacroix, 1997, Browning and Chiappori, 1998, and Vermeulen, 2005, for some recent examples). Evidence suggests that intra-household bargaining aspects within multi-person households cannot be ignored in general.

One particularly attractive model within the broad literature on intra-household decision-making is the so-called collective model (Chiappori, 1988). This model starts from the basic assumptions that a multi-person household is formed by individuals with own rational preferences, while these individuals are engaged in a bargaining process that results in Pareto-efficient intra-household allocations. Interestingly enough, the collective model entails theoretical implications that seem more difficult to reject when tested on multi-person household data (see Browning and Chiappori, 1998, Chiappori et al., 2002, and Vermeulen, 2005).

A crucial issue within the collective approach is how and when spouses' preferences and the intra-household bargaining process can be identified on the basis of a dataset that consists of only couples. As shown by Chiappori (1988), a so-called sharing rule (which summarizes how household means are allocated to the household members) can be identified up to a constant, whereas individual preferences can be identified up to a translation, by means of observed labor supply of couples, while assuming that preferences are egoistic or of the Beckerian caring type. Similar identification results in the presence of non-participation and/or taxation can be found in Blundell et al. (2001), Donni (2003) and Vermeulen (2006).

Although convenient, the widely used assumption of egoistic or caring preferences is rather restrictive. Both types of preferences imply that an individual's marginal rate of substitution between own leisure and consumption remains unaffected by his or her spouse's labor supply. As soon as more general preference structures are considered, additional assumptions must be introduced to obtain identification (see Chiappori and Ekeland, 2002). For example, Vermeulen et al. (2006) consider a collective model with externalities going beyond those defined by caring preferences. Identification is obtained by assuming that some aspects of individual preferences are the same across singles and couples. Their model is applied to cross-sectional data and involves a calibration stage.

Within a labor supply context, we present a novel approach to obtain identification in the presence of general externalities using longitudinal data. Our identification strategy is built upon the crucial assumption that an individual's preferences do not change over time, except for possible shocks that can be accounted for, when she or he turns into widow(er)hood. At first sight, the equal preferences assumption may appear restrictive. However, if one believes that preferences are individual-specific, as is explicitly the case in the collective approach, then it can be argued that preferences can only change via a clear channel following the death of one's partner. An obvious example is the mental health status that can be controlled for by modelling heterogeneity in preferences that may change over time. Since it can be argued that widow(er)hood is an exogenous shock to a household (conditional on a set of observables such as age and prior health), the equal preferences assumption seems much less controversial than assuming that *all* singles would have the same preferences as individuals in couples, irrespective of the fact that they were never married, were divorced or turned into widow(er)hood.

Clearly, panel data are a prerequisite to pursue this identification strategy because the equal preferences assumption is rather restrictive when comparing couples and singles in a cross-section. Therefore, we will apply our labor supply model to a sample of couples coming from the first six waves of the Health and Retirement Study (HRS). A crucial characteristic of our dataset is that a non negligible number of individuals, who initially were part of a couple, turned into widow(er)hood during the covered period. The application of our identification strategy to the HRS is interesting in its own right. However, it is also important from a policy point of view. For example, there is a lot of evidence that retirement decisions of spouses are coordinated. Indeed, the tendency of husbands and wives to retire together, or very shortly after each other, are well documented (see, e.g., Blau, 1998, Gustman and Steinmeier, 2000, Maestas, 2001; and Michaud, 2005; see Hamermesh, 2002 for a similar observation in non-elderly couples' labor supply). Complementarities in spouses' leisure are a likely explanation for such retirement patterns. Due to its structural approach of externalities with respect to leisure, the model may shed new light on the impact of social security reforms that are likely to have consequences on spouses' labor supply.

The rest of the paper is structured as follows. In Section 2, we present a collective labor supply model with externalities and show how identification is obtained. Section 3 describes the data, while the empirical results are discussed in Section 4. In Section 5, we discuss simulations of social security reforms by means of the structural collective model. One of the simulations concerns the widely discussed reform proposal to abolish the earnings test. Simulation results are compared with those obtained by a standard unitary model, which allows to stress the value added of the collective model. Section 6 concludes.

2 Identifying a collective labor supply model with externalities

2.1 Collective labor supply with externalities

In this section, we focus on the identification problem associated with collective labor supply models with externalities. This discussion proceeds in rather general terms. Specific issues associated with elderly couples' retirement decisions will be dealt with in a next subsection.

We focus on households that consist of two individuals m and f. In what follows, we assume that all consumption in the household is public. We further assume that preferences of each member allow for externalities with respect to the spouse's leisure. Preferences of individual j ($j = m, f$) can thus be represented by the following direct utility functions:

$$u_{it}^{j} = v^{j}(l_{it}^{m}, l_{it}^{f}, c_{it}), \tag{1}$$

where l_{it}^{j} denotes the amount of leisure of individual j in household i (i=1,...,I) at time t ($t = 1$,...,T). The Hicksian public consumption of household i at time t is denoted by c_{it}.

Both household members are involved in a bargaining process that determines the observed allocations over leisure and consumption. Following Chiappori (1988), we assume that this bargaining process results in Pareto efficient allocations. Observed allocations are therefore assumed to result from the maximization of a weighted linear social welfare function, subject to the household's budget constraint:

$$\max_{l_{it}^{m}, l_{it}^{f}, c_{it}} u_{it}^{h} = \mu_{it} v^{m}(l_{it}^{m}, l_{it}^{f}, c_{it}) + (1 - \mu_{it}) v^{f}(l_{it}^{m}, l_{it}^{f}, c_{it}) \tag{2}$$

subject to

$$p_t c_{it} \leq F(w_{it}^{m} h_{it}^{m} + w_{it}^{f} h_{it}^{f} + y_{it}).$$

The price of the public good c_{it} is denoted by p_t; whereas w_{it}^{j} and h_{it}^{j} ($j = m, f$) are respectively equal to individual j's gross wage and labor supply (where $h_{it}^{j} = E - l_{it}^{j}$; E being an individual's time endowment) and y_{it} is the household's non-labor income. The function F is a tax function that converts the household's gross income into net terms.

As is clear from (2), male and female utilities are weighted by respectively μ_{it} and $(1 - \mu_{it})$. Typically, these Pareto weights are assumed to depend on the individuals' wages and other budget set characteristics like the household's non-labor income. Moreover, Pareto weights may also depend on so-called 'distribution factors', which influence individuals' bargaining positions but do not affect their preferences nor the household's budget set (see Chiappori, Fortin and Lacroix, 2002). As a consequence, we have $\mu_{it} = \mu(\mathbf{p}_{it})$, where $\mathbf{p}_{it}$ is a vector of, say, 'bargaining factors' containing wages, other budget set characteristics and distribution factors. If $\mu(\mathbf{p}_{it})$ is increasing in w_{it}^{m}, for example, then the husband's bargaining position improves, following an increase in his wage. This implies that he will be able to claim a higher utility than before, which is produced by an intra-household allocation that is more favorable to the husband.

It must be stressed that the dependence of the above 'household utility function' on wages and non-labor income (via the Pareto weights) is the distinguishing characteristic between the collective model and the traditional unitary model. Exactly this dependence implies that labor supply fails to satisfy Slutsky conditions in a genuine collective labor supply model.

2.2 Identification of preference parameters and Pareto weights

A well-known problem of the above general setting with public goods and externalities is that the model is not identified without additional assumptions, even by making use of bargaining factors in the Pareto weights (see Chiappori and Ekeland, 2002). The problem is essentially that if marginal utilities are strictly positive for all goods, then there exists a continuum of utility functions and Pareto weights that yield the same observed choices. Chiappori and Ekeland (2002) propose restrictions on these marginal utilities to secure identification. They propose to use one exclusive good per spouse (this is a good that is consumed by only one household member; males' and females' clothes serve as examples) that is assumed not to generate any externalities. The marginal utility with respect to the spouse's exclusive good is thus set equal to zero, which entails identification.

Since we only observe (aggregate) disposable income in the data that we use and we explicitly want to consider externalities with respect to leisure, this approach is not immediately applicable. So we must rely on an adapted solution. What we want to show below is that using couples and widow(er)s serves the identification purpose without using exclusive goods. Basically, we restrict the marginal utility of the partner's leisure to be zero when that spouse dies. Essentially, our identification strategy differs from that proposed by Chiappori and Ekeland (2002) in the sense that exclusivity is not strict in our case (i.e., it only applies when a spouse died).

A crucial assumption in our approach is that an individual's preferences do not change, except for possible shocks that we can control for, when she or he turns into widow(er)hood. An obvious candidate of such a shock is the impact on health when somebody loses her or his partner. In other words, we assume that if preferences change due to the loss of a partner, then this only happens via a clear channel.[1] Note that this equal preferences assumption still allows for several sources for differences in behavior between widow(er)s and individuals in couples. Firstly, widow(er)hood may be followed by a health shock that on its turn may influence labor supply and consumption. Secondly, opportunity sets can be very different over household types, especially when a couple ceases to exist due to the death of a spouse. This illustrates that the equal preferences assumption, which is crucial to our identification strategy, does not necessarily straitjacket observed behavior.

Let us now illustrate how identification of the individuals' preferences, as captured by the utility functions $v^j(l^m_{it}, l^f_{it}, c_{it})$ $(j = m, f)$, and Pareto weights $\mu(\mathbf{p}_{it})$ and $(1 - \mu(\mathbf{p}_{it}))$ proceeds. Without losing generality, we will illustrate our identification strategy by means of the preferences postulated in the empirical application.

Preferences are assumed to be of a modified Cobb-Douglas type and are represented by the

[1] It is worth stressing that we do not assume that there is no impact on an individual's utility due to the loss of her or his partner. More broadly, our approach does not exclude a bereavement process (see, Kahneman, 1999).

following direct utility function ($j = m, f$):

$$u_{it}^j = \beta_m^j(\mathbf{z}_{it}^j) \ln l_{it}^j + \beta_c^j(\mathbf{z}_{it}^j) \ln c_{it} + \beta_{mf}^j(\mathbf{z}_{it}^j) \ln l_{it}^m \ln l_{it}^f, \tag{3}$$

where $\mathbf{z}_{it}^j$ is a vector of preference shifters like age and health status of individual j in household i at time t.[2] These utility functions depend on the spouse's leisure via $\beta_{mf}^j(\mathbf{z}_{it}^j)$. We would expect a positive coefficient $\beta_{mf}^j(\mathbf{z}_{it}^j)$ if marginal utility of leisure increases with leisure of the spouse.[3]

Consider next the collective household utility function that is a weighted average of the above individual utilities with appropriate Pareto weights:

$$\begin{aligned} u_{it}^h \;=\;& \mu(\mathbf{p}_{it}) \left(\beta_m^m(\mathbf{z}_{it}^m) \ln l_{it}^m + \beta_c^m(\mathbf{z}_{it}^m) \ln c_{it} + \beta_{mf}^m(\mathbf{z}_{it}^m) \ln l_{it}^m \ln l_{it}^f\right) \\ &+ (1 - \mu(\mathbf{p}_{it})) \left(\beta_f^f(\mathbf{z}_{it}^f) \ln l_{it}^f + \beta_c^f(\mathbf{z}_{it}^f) \ln c_{it} + \beta_{mf}^f(\mathbf{z}_{it}^f) \ln l_{it}^m \ln l_{it}^f\right). \end{aligned} \tag{4}$$

Next, given the above specification for the household utility function, we get the following restrictions from which to identify the parameters from couples' data alone (these are obtained by merging the terms associated with respectively $\ln l_{it}^m$, $\ln l_{it}^f$, $\ln c_{it}$ and $\ln l_{it}^m \ln l_{it}^f$):

$$\begin{aligned} \eta_m(\mathbf{p}_{it}, \mathbf{z}_{it}^m, \mathbf{z}_{it}^f) &= \mu(\mathbf{p}_{it}) \beta_m^m(\mathbf{z}_{it}^m), \\ \eta_f(\mathbf{p}_{it}, \mathbf{z}_{it}^m, \mathbf{z}_{it}^f) &= (1 - \mu(\mathbf{p}_{it})) \beta_f^f(\mathbf{z}_{it}^f), \\ \eta_c(\mathbf{p}_{it}, \mathbf{z}_{it}^m, \mathbf{z}_{it}^f) &= \mu(\mathbf{p}_{it}) \beta_c^m(\mathbf{z}_{it}^m) + (1 - \mu(\mathbf{p}_{it})) \beta_c^f(\mathbf{z}_{it}^f), \\ \eta_{mf}(\mathbf{p}_{it}, \mathbf{z}_{it}^m, \mathbf{z}_{it}^f) &= \mu(\mathbf{p}_{it}) \beta_{mf}^m(\mathbf{z}_{it}^m) + (1 - \mu(\mathbf{p}_{it})) \beta_{mf}^f(\mathbf{z}_{it}^f). \end{aligned} \tag{5}$$

To make the exposition easier, let us assume that the Pareto weights are linear in the set of bargaining factors: $\mu(\mathbf{p}_{it}) = \mathbf{p}_{it}'\boldsymbol{\mu}$; a similar assumption is made with respect to the β-parameters. Let us denote the dimensions of the vectors $\mathbf{p}_{it}, \mathbf{z}_{it}^m$ and $\mathbf{z}_{it}^f$ by respectively $\mathbf{P}$, $\mathbf{Z}^m$ and $\mathbf{Z}^f$, where each vector includes a constant. It can now be easily checked that there are $\mathbf{PZ^m}$ and $\mathbf{PZ^f}$ restrictions for $\eta_m\left(\mathbf{p}_{it}, \mathbf{z}_{it}^m, \mathbf{z}_{it}^f\right)$ and $\eta_f\left(\mathbf{p}_{it}, \mathbf{z}_{it}^m, \mathbf{z}_{it}^f\right)$, while there are $\mathbf{P}(\mathbf{Z}^m + \mathbf{Z}^f - 1)$ restrictions for $\eta_c\left(\mathbf{p}_{it}, \mathbf{z}_{it}^m, \mathbf{z}_{it}^f\right)$ and $\eta_{mf}\left(\mathbf{p}_{it}, \mathbf{z}_{it}^m, \mathbf{z}_{it}^f\right)$, given the interactions between each of the elements of $\mathbf{p}_{it}$ and the diverse β-parameters. Therefore, there are $3\mathbf{P}(\mathbf{Z}^m + \mathbf{Z}^f) - 2\mathbf{P}$ restrictions in total. The number of structural parameters defining preferences and Pareto weights equals $\mathbf{P} + 3(\mathbf{Z}^m + \mathbf{Z}^f)$. Hence, it appears that the model is identified based on the order condition for $(3\mathbf{P} - 3)(\mathbf{Z}^m + \mathbf{Z}^f) - 3\mathbf{P} \geq \mathbf{0}$. For example, if there are no preference shifters (i.e., $\mathbf{Z}^m = \mathbf{Z}^f = 1$), then $\mathbf{P}$ should be greater than 1. This is the case as soon as one

[2] Note that the dimension of the vector of preference shifters may differ across the preference parameters (i.e., some variables may not have any impact on a particular preference parameter). To avoid a notation that is too complex for our purposes, we do not explicitize this.

[3] Gustman and Steinmeier (2000) use a similar specification when it comes to the complementarity effect. They assume that the individual's marginal utility of leisure depends on the retirement decision of the spouse although not in log terms. However, they constrain spouses to have the same marginal utility of consumption. This is done in order to simplify the joint problem to one of within period differences in preferences.

element in $\mathbf{p}_{it}$ is added on top of a constant. However, we can show in this context that the rank condition fails to be satisfied.

Consider, for example, the following transformed individual utility functions and welfare weights (compare with Chiappori and Ekeland, 2002):

$$\begin{aligned} v^{f*}(l_{it}^m, l_{it}^f, c_{it}) &= \epsilon v^f(l_{it}^m, l_{it}^f, c_{it}) + (1-\epsilon)\, v^m(l_{it}^m, l_{it}^f, c_{it}), \\ v^{m*}(l_{it}^m, l_{it}^f, c_{it}) &= (\epsilon - 1)\, v^f(l_{it}^m, l_{it}^f, c_{it}) + \epsilon v^m(l_{it}^m, l_{it}^f, c_{it}), \\ \mu^*(\mathbf{p}_{it}) &= \frac{\mu(\mathbf{p}_{it}) - 1 - \epsilon}{2\epsilon - 1},\; 1 - \mu^*(\mathbf{p}_{it}) = \frac{\epsilon - \mu(\mathbf{p}_{it})}{2\epsilon - 1}, \end{aligned}$$

where $\epsilon > \max\{\mu(\mathbf{p}_{it}), 1 - \mu(\mathbf{p}_{it})\} \setminus \left\{\frac{1}{2}\right\}$. It can now easily be checked that the collective household utility function that is obtained by these transformed utilities and Pareto weights results in the same set of restrictions as in (5) if all marginal utilities are non-zero. In other words, a continuum of structural models (defined by different values of ϵ) give rise to the same reduced-form parameters; and therefore the same predicted behavior from couples' data alone.

Key for this non-identification result is that the individual utility functions are strictly increasing in their arguments. Therefore, Chiappori and Ekeland (2002) propose to use exclusive goods, one for each spouse, to obtain identification (by setting the marginal utility of the partner's exclusive good equal to zero). We will show next that using couples and widow(er)s, with the above equal preferences assumption, serves that purpose without using exclusive goods. To do this, we introduce two dummy variables k_{it}^m and k_{it}^f. If an observation i at time t consists of a couple, then both dummies are equal to one. For widowers (widows), k_{it}^m (k_{it}^f) equals one, while k_{it}^f (k_{it}^m) is equal to zero. We further introduce an exogenously determined leisure level associated with the deceased spouse (essentially a constant; cf. infra), denoted by $\bar{l}_{it}$. We can now write the household utility function (with some abuse of notation) as:

$$\begin{aligned} u_*^h(\mathbf{y}_{it}; \mathbf{p}_{it}, k_{it}^m, k_{it}^f, \mathbf{z}_{it}^m, \mathbf{z}_{it}^f) &= (1 - k_{it}^f) v^m(\mathbf{y}_{it}; k_{it}^f, \mathbf{z}_{it}^m) + k_{it}^f v^f\left(\mathbf{y}_{it}; k_{it}^m, \mathbf{z}_{it}^f\right) \\ &\quad + k_{it}^m k_{it}^f \mu(\mathbf{p}_{it}) (v^m(\mathbf{y}_{it}; k_{it}^f, \mathbf{z}_{it}^m) - v^f(\mathbf{y}_{it}; k_{it}^m, \mathbf{z}_{it}^f)), \end{aligned}$$

where $\mathbf{y}_{it} = (l_{it}^m, l_{it}^f, c_{it})'$.

For widows the utility function equals:

$$\begin{aligned} u_*^h\left(\mathbf{y}_{it}; \mathbf{p}_{it}, 0, 1, \mathbf{z}_{it}^m, \mathbf{z}_{it}^f\right) &= v^f\left(\mathbf{y}_{it}; 0, \mathbf{z}_{it}^f\right) \\ &= (\beta_f^f(\mathbf{z}_{it}^f) + \beta_{mf}^f(\mathbf{z}_{it}^f) \ln \bar{l}_{it}) \ln l_{it}^f + \beta_c^f(\mathbf{z}_{it}^f) \ln c_{it}, \end{aligned}$$

where the marginal utility with respect to the leisure of the deceased husband equals zero. For widowers, we have:

$$\begin{aligned} u_*^h\left(\mathbf{y}_{it}; \mathbf{p}_{it}, 1, 0, \mathbf{z}_{it}^m, \mathbf{z}_{it}^f\right) &= v^m(\mathbf{y}_{it}; 0, \mathbf{z}_{it}^m) \\ &= (\beta_m^m(\mathbf{z}_{it}^m) + \beta_{mf}^m(\mathbf{z}_{it}^m) \ln \bar{l}_{it}) \ln l_{it}^m + \beta_c^m(\mathbf{z}_{it}^m) \ln c_{it}, \end{aligned}$$

in which case the marginal utility of the leisure of the deceased wife equals zero.

Finally, for couples we have

$$\begin{aligned} u_*^h\left(\mathbf{y}_{it};\mathbf{p}_{it},1,1,\mathbf{z}_{it}^m,\mathbf{z}_{it}^f\right) &= \mu(\mathbf{p}_{it})v^m\left(\mathbf{y}_{it};1,\mathbf{z}_{it}^m\right)+(1-\mu(\mathbf{p}_{it}))v^f\left(\mathbf{y}_{it};1,\mathbf{z}_{it}^f\right) \\ &= \eta_m(\mathbf{p}_{it},\mathbf{z}_{it}^m,\mathbf{z}_{it}^f)\ln l_{it}^m+\eta_f(\mathbf{p}_{it},\mathbf{z}_{it}^m,\mathbf{z}_{it}^f)\ln l_{it}^f \\ &\quad +\eta_c(\mathbf{p}_{it},\mathbf{z}_{it}^m,\mathbf{z}_{it}^f)\ln c_{it}+\eta_{mf}(\mathbf{p}_{it},\mathbf{z}_{it}^m,\mathbf{z}_{it}^f)\ln l_{it}^m\ln l_{it}^f. \end{aligned}$$

From widows and widowers, $2(\mathbf{Z}^m+\mathbf{Z}^f)$ parameters can be estimated. Using the estimated parameters from the widow(er)s in the four sets of restrictions for couples, enables to separately identify Pareto weights from preference parameters. This is because there are $(\mathbf{Z}^m+\mathbf{Z}^f)$ remaining preference parameters and $\mathbf{P}$ Pareto weight parameters, while there are $3\mathbf{P}(\mathbf{Z}^m+\mathbf{Z}^f)-2\mathbf{P}$ parameters that can be estimated (assuming that the Pareto weights are not constant).

Intuitively, the identification makes use of labor supply behavior of single-person households (more specifically widows and widowers) to help identifying preferences of spouses in couples and Pareto weights. From standard microeconomic theory, we know that preferences are uniquely identified on the basis of single-decision makers' observed labor supply. It is clear from the above discussion that preferences and Pareto weights are identified up to the constant $\bar{l}_{it}$ that is to be exogenously determined. Although essentially a normalization issue, the interpretation of the model may differ across different choices for $\bar{l}_{it}$. One natural choice for $\bar{l}_{it}$ seems the minimum leisure available to an individual. This level implies that, *ceteris paribus*, a widow is equally well off after the death of her spouse when that spouse worked the maximum number of hours possible. Of course, the *ceteris paribus* clause is important here: as soon as there would be a mental health shock, captured by a change in one of the preference shifters, the female utility level would be affected by the loss (which is thus not ruled out in our specific equal preferences assumption). We can also take into account any changes in financial resources that occur when a spouse dies.

2.3 Further estimation aspects

Let us now turn to the empirical strategy to apply the above structural model. Specific to this strategy is that it is embedded in a discrete choice framework (see van Soest, 1995). This approach assumes that individuals have the choice between only a limited number of labor supply options. The advantage of the approach is that it can easily deal with complex non-linear and non-convex tax schemes; a feature certainly applicable to the tax and social security benefits system applicable to the elderly.

Assuming that both spouses are confronted with the same number of options, let us denote the gender specific discrete set of leisure choices by $L^j=\left\{l_p^j;p=1,...,P\right\}$ $(j=m,f)$. Possible leisure choices l_p^j are equal to $E-h_p^j$, were E is an individual's weekly assignable time endowment (which takes into account time needed to sleep and other maintenance tasks) and

h_p^j are weekly working hours. In the empirical exercise, the number of working hours choices P is set equal to 4. Working hours choices are assumed to be equal to respectively 0, 25, 40 and 50 hours for husbands, and 0, 15, 30 and 40 hours for wives. These choices are based on the empirical distribution in the data used (on the basis of yearly hours divided by 50 weeks).

With discrete choice labor supply in a collective setting, one of the main concerns will be to postulate individual utility functions which impose coherency through concavity, and further will allow for taste variation in all possible states. Restricting coherency conditions to own choice variables, the above described modified Cobb-Douglas utility functions (see (3)) are well-behaved if marginal utilities with respect to own consumption and leisure are positive.

Preference parameters are assumed to be of the following form ($j = m, f$).

$$\beta_c^j \left(\mathbf{z}_{it}^j\right) = \mathbf{z}_{it}^{j\prime}\delta_c^j$$

$$\beta_j^j \left(\mathbf{z}_{it}^j, k\right) = \mathbf{z}_{it}^{j\prime}\delta_j^j + \theta_k^j$$

$$\beta_{mf}^j \left(\mathbf{z}_{it}^j\right) = \mathbf{z}_{it}^{j\prime}\delta_{mf}^j,$$

where $\mathbf{z}_{it}^j$ are observable characteristics like age and health indicators, while θ_k^j is a spouse-specific taste shifter affecting labor supply that is unobserved by the researcher. A higher θ_k^j implies a higher marginal propensity of leisure and thus a lower work effort. Following Hoynes (1996), and in the spirit of Heckman and Singer (1984), we assume that there is only a limited number of pairs $\left(\theta_k^m, \theta_{k'}^f\right)$ that influence spouses' preferences. More specifically, we assume that there are three points of support per spouse that differ across gender ($k, k' = 1, 2, 3$), which produces $3 \times 3 = 9$ different possibilities $\left(\theta_k^m, \theta_{k'}^f\right)$; each associated with a probability φ_n where $\sum_n \varphi_n = 1$.

An issue that was ignored up to now are fixed costs of participation. These are costs that an individual has to pay to get to work. We model these in a way, such that they capture the widely observed state dependence in elderly individuals' labor supply decisions (see Michaud, 2005). We model the budget constraint (when both spouses are alive) as

$$c_{it} \leq F(w_{it}^m h_{it}^m + w_{it}^f h_{it}^f + y_{it}) - \zeta(I(h_{it}^m > 0)I(h_{it-1}^m = 0) + I(h_{it}^f > 0)I(h_{it-1}^f = 0)).$$

The parameter ζ captures the per-spouse monetary costs of returning to work. For simplicity, we assume that they cannot be taxed.[4] Such fixed costs therefore create different reservation wages depending on labor force status in the previous year. Hence, it creates state-dependence in transition probabilities.

[4] We make this assumption in order to avoid computing taxes each time the likelihood is being evaluated. We also consider the costs to be the same for husbands and wives. A likelihood ratio test did not give any evidence against this last assumption.

Let us now focus on the Pareto weights $\mu(\mathbf{p}_{it})$ and $1-\mu(\mathbf{p}_{it})$. Since the latter have to be between zero and one, we opt for the following functional specification:

$$\mu(\mathbf{p}_{it}) = \frac{\exp(\mathbf{p}_{it}'\boldsymbol{\mu})}{1+\exp(\mathbf{p}_{it}'\boldsymbol{\mu})}. \tag{6}$$

Pareto weights theoretically depend on spouses' wages. Since we explicitly take into account the tax and social security benefit system, the choice of which specific wage variables have an impact on spouses' bargaining positions is not entirely innocuous. Gross wages are not interesting since these do not change after a reform in the tax and social security benefit system. Net wages are not attractive neither, since these depend on the specific number of hours worked. Individuals' bargaining positions, however, should be exogenously determined in a genuine collective model. In the empirical exercise, therefore, we will make use of a variable that captures the relative earning potential of spouses and that was earlier applied in Vermeulen et al. (2006). The variable is the quotient of the male's (denoted by y^m) and the female's (denoted by y^f) marginal contributions to the household's consumption when switching from non-participation to full-time participation. More specifically, let $r_{p\prime}^{f}$ $(p\prime = 1, ..., P)$ denote the observed relative sample frequencies of the weekly labor supply choices $h_{p\prime}$ of women. Denote $R_p^{p\prime}$ the household's consumption when the husband works h_p hours and the wife works $h_{p\prime}$ hours. The impact on household consumption if the male switches from non-participation $(p = 1)$ to full-time working hours $(p = P)$ is then measured as:

$$y^m = \sum_{p\prime=1}^{P} r_{p\prime}^{f}\left(R_P^{p\prime} - R_1^{p\prime}\right).$$

In a similar fashion, we can define the variable y^f. The male's relative earning capacity can finally be defined as $\frac{y^m}{y^f}$. We would expect an increase in the male's Pareto weight if his relative earning capacity increases. Other bargaining factors that will be considered in the empirical exercise are the household's non-labor income and male's and female's average indexed monthly earnings, which are directly linked with the calculation of social security benefits.

There are maximum P^2 possible values of the decision variables for married couples. They involve all combinations of both individuals' labor supply. Since the budget constraint is potentially non-convex and choices are discrete we need to evaluate utility at P^2 points:

$$u^h(\mathbf{y}_{it}^s; \mathbf{p}_{it}, \mathbf{k}_{it}, \mathbf{z}_{it}, \boldsymbol{\theta}_n) = u_*^h(\mathbf{y}_{it}; \mathbf{p}_{it}, \mathbf{k}_{it}, \mathbf{z}_{it}, \boldsymbol{\theta}_n) + \varepsilon_{sit},$$

where $\mathbf{k}_{it} = (k_{it}^m, k_{it}^f)'$, $\mathbf{y}_{it}^s$ are the decision variables in the choice set $\left(s = 1, ..., P^2\right)$, $\mathbf{z}_{it}$ captures all other exogenous variables and $\boldsymbol{\theta}_n$ are specific values for the discrete heterogeneity parameters. We add some measurement or optimization error term for each alternative that is assumed to be independent and identically distributed across alternatives with type 1 extreme value distribution (GEV(1)). Note that for widow(er)s only maximum P states are observed.

As noted in van Soest (1995), some options, particularly part-time options, appear not to be chosen by many workers as would be predicted from continuous utility functions. One reason could be that such hours choices are simply not available on the job. Following van Soest (1995) we allow for part-time specific utility costs in the objective function that the couple optimizes (the weighted sum of utilities). In particular, we introduce three parameters, for the three options that involve part-time work (25 hours for husbands; 15 and 30 for wives) to the deterministic part of the utility of each option:

$$u_{**}^h(\mathbf{y}_{it}; \mathbf{p}_{it}, \mathbf{k}_{it}, \mathbf{z}_{it}, \boldsymbol{\theta}_n) = u_*^h(\mathbf{y}_{it}; \mathbf{p}_{it}, \mathbf{k}_{it}, \mathbf{z}_{it}, \boldsymbol{\theta}_n) + \gamma_{m,s} + \gamma_{f,s},$$

where $\gamma_{j,s}$ $(j = m, f)$ is set to zero for options not involving part-time work.

For an observation i, the probability of observing the sequence $\mathbf{y}_i = (\mathbf{y}_{i1}', ..., \mathbf{y}_{iT}')'$, given the history of exogenous variables (including bargaining factors $\mathbf{p}_{it}$ for ease of notation) $\mathbf{z}_i = (\mathbf{z}_{i1}', ..., \mathbf{z}_{iT}')'$, household types $\mathbf{k}_i = (\mathbf{k}_{i1}', ..., \mathbf{k}_{iT}')'$ and some specific value for $\boldsymbol{\theta}_n$, is:

$$\Pr(\mathbf{y}_i|\mathbf{k}_i, \mathbf{z}_i, \boldsymbol{\theta}_n) = \prod_{t=1}^{T} \Pr(\mathbf{y}_{it}|\mathbf{k}_{it}, \mathbf{z}_{it}, \boldsymbol{\theta}_n), \tag{7}$$

which follows from the time-independence assumption on ε_{sit}. If the ε_{sit} follow a GEV(1) distribution, then we have the familiar conditional logit formulation

$$\Pr(d_{it}^s = 1|\mathbf{k}_{it}, \mathbf{z}_{it}, \boldsymbol{\theta}_n) = \frac{\exp(u_{**}^h(\mathbf{y}_{it}^s; \mathbf{k}_{it}, \mathbf{z}_{it}, \boldsymbol{\theta}_n))}{\sum_{s'} \exp(u_{**}^h(\mathbf{y}_{it}^{s'}; \mathbf{k}_{it}, \mathbf{z}_{it}, \boldsymbol{\theta}_n))}, \tag{8}$$

where d_{it}^s is an indicator variable that equals 1 (0) if choice $\mathbf{y}_{it}^s$ is (not) made and s' denotes the available choices.

The maximum likelihood problem can be written as:

$$\max_{\boldsymbol{\varphi}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\zeta}, \boldsymbol{\gamma}} \sum_i \log \sum_n \varphi_n \prod_{t,s} \Pr(d_{it}^s = 1|\mathbf{z}_{it}, \mathbf{k}_{it}, \boldsymbol{\theta}_n)^{d_{it}^s}, \tag{9}$$

where $\boldsymbol{\varphi}, \boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\zeta}, \boldsymbol{\gamma}$ are the parameters to be estimated and which refer to respectively the discrete heterogeneity probabilities, the discrete heterogeneity parameters and the other preference/fixed costs parameters. We estimate parameters using the BFGS algorithm and standard errors using the negative of the inverse of the Hessian computed at the solution.

2.4 Unitary model specification

As mentioned in the introduction, results obtained by means of the above collective model will be compared with those obtained by a standard unitary model. The specification of the latter model is based on that for the individual utility functions in the collective model (see (3)). More specifically, we retain the modified Cobb-Douglas utility function that is assumed to represent preferences of household i at time t:

$$u_{it} = \beta_m(\mathbf{z}_{it}) \ln l_{it}^m + \beta_f(\mathbf{z}_{it}) \ln l_{it}^f + \beta_c(\mathbf{z}_{it}) \ln c_{it} + \beta_{mf}(\mathbf{z}_{it}) \ln l_{it}^m \ln l_{it}^f. \tag{10}$$

It is interesting to compare this unitary utility function with the collective household utility function of equation (4). It is easily seen that the main difference between both functional specifications is the presence of the Pareto weights in the collective model. In fact, two general mechanisms imply that our collective model reduces to a unitary one. The first mechanism is that of a price-independent utility function: if the Pareto weights do not depend on wages nor on non-labor income, then the collective specification can be reduced to a unitary one. The second mechanism is that if both spouses have the same (representation of) preferences, then the (well-behaved) household utility function equals the spouses' common utility function.

Note finally, that we also take into account transition costs and part-time restrictions in the unitary specification and that we allow for discrete unobserved heterogeneity in the form of an equal number of points of support as in the collective model. Hence the unitary model is nested in the collective model we estimate.

2.5 Considerations specific to retirement

In old age, couples not only decide to work or not, but also to claim social security benefits and their private pension if they have such a pension on their current job. A fully realistic model of behavior in old age would recognize that these decisions do not necessarily coincide. We choose to ignore the decision to claim social security, allowing for self-reported income from social security in the budget set, while we make the decision to claim a pension on the current job coincide with labor force exit. We make these choices for two reasons.

First, the Social Security claiming decision is poorly understood and forward-looking models have a hard time rationalizing why so many workers claim at age 62 (the earliest age at which workers can claim) if the pension is in fact actuarily fair (Hurd et al., 2004). We therefore decide to focus on the decision to work for pay. We explicitly allow the budget set to depend on the receipt of social security benefits. Furthermore, conditional on receiving benefits, we allow Social Security institutions such as the spouse benefit and the earnings test to affect resources of couples. Spouses are not restricted in their choices if they decide to claim Social Security benefits. As for private pensions, if a respondent claims to have right to a pension on a current job, we allow the budget set at zero hours of work to depend on his/her estimated pension right at that point. Once he or she has quit the job, we simply include whatever pension income he or she reports receiving in his/her non-labor income. For pensions from past jobs, these should never affect current work decisions and we include those receipts in non-labor income for all work options.[5]

Second, introducing a claiming decision or private pension claiming decision makes it almost necessary to move from a myopic model to a forward-looking model of behavior. Several

[5] This strategy for private pensions works well for most couples who exit the labor force and draw a pension. It will not be adequate for workers who switch job (perhaps to a bridge job) and claim their pension from their previous employer. Income from that pension will only show up in the wave following the transition.

complications, such as commitment, are introduced when one moves from a myopic collective model to a forward-looking model (Mazzocco, 2006). It is also true that the literature on the effects of the earnings test finds labor supply responses that seem to suggest that individuals do not consider the fact that benefits lost due to the earnings test are reimbursed at a nearly actuarily fair rate in the future (the degree of actuarial fairness is increasing for recent cohorts). Since identification issues are at the center of this paper, we decided to side-step from forward-looking considerations and analyze how static collective and unitary models differ when applied to older couples. It is clear that further research on intra-household models should aim to study such models.

3 Data

We apply our model to data that are drawn from the Health and Retirement Study (HRS). The HRS is sponsored by the National Institute on Aging and conducted by the University of Michigan. It consists of a longitudinal study following a cohort of individuals who were born between 1931 and 1941 (both years included), and their spouses if married (regardless of age). The HRS is extremely rich and encompasses lots of socio-economic information like the respondents' demographic background, employment status, job history, income sources, health status, wealth and pension plans. Moreover, the HRS includes a linkage with administrative data of the Social Security Administration which provides respondents' earnings histories from 1950 to 1991, allowing an accurate calculation of the Old Age Social Insurance benefits. We use the public release (version E) of the HRS from the RAND Corporation (StClair et al., 2002), which merges respondents from the first six biennial waves covering the period between 1992 and 2002. These data are complemented with the SSA files and the employer-provided pension information.

The sample selection is for married or cohabiting couples where both individuals were alive in 1992. For identification reasons explained earlier, these initial couples remain in the dataset for the all waves where at least one spouse is observed, irrespective of the fact that some individuals turn to widow(er)hood in the given time span. After deleting observations with important missing information, we have a sample of 2342 households that are potentially observed in six consecutive periods.[6]

Table 1 reports, across all-waves, the numbers of complete couples, widows and widowers along with summary statistics of key variables in the analysis. As is clear from the table, and according to common knowledge, more widows than widowers are observed as time proceeds.

[6] Since we need accurately calculated OASI benefits, the sample selection was for respondents who gave permission to link their earnings histories as reported to the Social Security Administration to the HRS. Earnings histories are available for about 75% of the HRS respondents. According to Haider and Solon (2000), the HRS Social Security earnings sample is reasonably representative for the original sample. In the appendix we give details on how budget sets were constructed using this information.

Widows and widowers are generally in worse health than their counterparts in couples. This may reflect some correlation with their lost spouse's health (see Michaud and van Soest, 2004), but may also be due to the fact that individuals in couples are younger on average than widow(er)s. Note that about 25 to 35% of the individuals have experienced a severe health onset (e.g., cancer, a heart condition or a stroke). This is the case for more than 70% of the sample as far as mild health onsets (e.g., diabetes or arthritis) are concerned. Worthy of note is that relatively more widow(er)s are subject to mental health problems than individuals in couples. This feature allows to relax our equal preferences assumption that does not exclude controllable health shocks after the death of one's spouse.

Focusing on labor supply, it turns out that widow(er)s participate less than (fe)males in couples. The table also indicates that widow(er)s generally earn less than individuals in couples, and that they have lower non-labor income like capital income, pension income and social security benefits.

[Table 1 about here]

Table 2 gives weekly working hours for individuals in couples, widows and widowers. It is clear from the table that a substantive number of individuals work part-time, which indicates that it is indeed worthwhile to look at the intensive margin, in addition to the mere decision whether or not to participate. The table further shows that widow(er)s not only participate less than individuals in couples, but that they also supply less hours on average.

[Table 2 about here]

But these observed differences across household types are not likely to be informative about the degree of externalities in leisure. Couples and widow(ers) are different in many dimensions as seen in Table 1. It is also likely that they are different in unobserved dimensions as well. Hence longitudinal data is crucial in comparing a respondent before and after the event of the spouse's death.

4 Empirical results

4.1 Collective model

Collective model estimates are shown in Table 3 (see equation (4)). Let us first have a closer look at the preference parameters. As one would expect, there is an important age effect for males and females. This age effect runs both via the marginal utility of own leisure (in a positive way) and via the marginal utility of consumption (in a negative way); both implying that, all else equal, an individual works less as she or he gets older. Secondly, health turns out to be multidimensional: different health indicators have their own impact on an individidual's preferences. Since health variables are taken up in the marginal utilities of both own leisure and consumption, and not all indicators have different signs in the marginal utilities, the impact of these indicators cannot always easily be interpreted. It turns out, however, that having experienced onsets of severe or mild conditions implies, *ceteris paribus*, a significant negative impact on labor supply. This is the case for both men and women. Further, a poor self-reported health implies less hours supplied than in the case self-reported health is good to excellent. Finally, having mental health problems is associated with a lower labor supply for males if everything else is held constant. For females, there is no significant impact with respect to mental health.

We can also see from Table 3 that there are significant externalities with respect to leisure. For men, the marginal effect of the spouse's leisure is significantly positive, while there are significant negative externalities with respect to the husband's leisure in female preferences. Note, however, that this somewhat unexpected result for wives is in line with Booth and van Ours (2006). On the basis of time use data (and in a slightly different context), they find that partnered women's life satisfaction is increased if their partners work full-time. For males, they do not find such a result. They claim that this result is consistent with Akerlof and Kranton's (2000) gender identity hypothesis. The latter assumes that individuals experience a welfare loss if they are in a position that is not conform to society's behavioral prescriptions. Our results are indeed consistent with the view, which is probably more widespread among elderly persons, that men should only engage in market work, while women should fully engage in home work (which is part of the amount of time spent on leisure in our model without home production). Interestingly, Gustman and Steinmeier (2004) find a statistically significant positive complementarity in leisure effect for husbands (wife's retirement in husband's marginal utility of leisure) which is in-line with our results. Their model is a non-cooperative model of each spouse's age of retirement which is not directly comparable to the current cooperative model which models both work and hours.

The results in Table 3 further illustrate that there is considerable unobserved discrete heterogeneity between individuals (and between couples). All three gender specific points of support are significantly different from zero and imply quite different marginal utilities of leisure. Note that most combinations of spouses' points of support frequently occur according to our estimates: all but two combinations are associated with more than 5 percent of the couples (see Table 4). The implied correlation coefficient in unobserved heterogeneity across spouses equals 0.35 (see bottom line of Table 4). This generates in itself a correlation in labor supply decisions of households. Gustman and Steinmeier (2004) report a similar estimate for their correlation in unobserved taste for leisure using another dataset. Michaud (2005) finds that this correlation alone explains close to one fourth of the joint transition rate from work to inactivity among couples. The remaining correlation can be explained by joint incentives while our results show that another channel is the husbands positive utility from the wife's consumption of leisure. From these results one could tentatively conclude that externalities are only likely to explain joint retirement for males that follow their wife into retirement. Hence similarly to Gustman and Steinmeier (2000) we could conclude that most of the observed correlation is accounted for by correlation in preferences and differences in opportunity sets.

We find that fixed costs of participation are important, both in economic and statistical terms. Starting to work full-time implies an average weekly fixed cost of 118 dollar, which seems to make sense. Starting to work only part-time implies an additional small cost. This creates state-dependence in transitions out and in the labor force. Such effects can be explained by the transition cost of returning to work after retirement. Because they enter linearily in consumption (except for taxes) one can also interpret this cost as the transitory reduction in wages that a worker must accept when returning to work after retirement. This cost would represent roughly a 15% drop in earnings for a worker earning 20 dollar an hour in a full-time job (40 hours a week). However, to be consistent with the model, this drop would have to vanish in the next wave after returning to work.

A major feature of our model is that it is embedded in the collective approach to household behavior. As the results at the bottom of Table 3 indicate, this structural aspect cannot be ignored. First of all, the husband's relative earning capacity has a positive and highly significant impact on his Pareto weight, which itself has consequences on how leisure and consumption is allocated within the household. Also the household's non-labor income has a negative and significant impact on the husband's Pareto weight. These results (taken together with the rejection of the assumption that spouses have equal utility functions) imply a strong rejection of the standard unitary model, which is characterized by price-independent utility functions (cf. supra). Simulations obtained by means of models not taking into account intra-household bargaining aspects should thus be interpreted with care. Also some specific aspects of the social security system influence a household's bargaining process. This is the case for the husband's AIME, which has a negative impact on his Pareto weight. One explanation for this effect is that if a husband has a high AIME, then his spouse can collect a high spouse benefit or a divorce benefit whatever the husband does. The returns from the spouse benefit accrue to her, independently of her participation status, which may increase her bargaining position.

[Table 3 about here]

[Table 4 about here]

4.2 Unitary model

Let us now discuss the unitary model's estimates (see equation (10)). These are shown in Table 5. Results regarding preference heterogeneity and fixed costs are largely similar except a few expections. The result obtained earlier that health is multidimensional is also confirmed by the unitary model. Generally speaking, worse health conditions imply less hours supplied. Just like above, the null hypothesis of no unobserved discrete heterogeneity with respect to leisure is strongly rejected: all parameters associated with the different points of support are significantly different from each other. Again, heterogeneity in the marginal utility of leisure is positively correlated across spouses.Results about the sources of the jointness of decisions are likely to be similar across models. To assess which models seems to capture better observed pattern in the HRS and how predictions from using such models differ, we use simulations using the estimated parameters.

4.3 Goodness-of-fit results

The goodness-of-fit of both models is illustrated in Table 6. The table shows observed and percentage point differences between the predicted couples' labor supply frequencies and the observed frequencies for the year 1996. This corresponds to a year where spouses in couples are 55 to 65 years old, hence still working considerably while maybe receiving social security benefits. The collective and unitary models' predicted frequencies will serve as the baseline situation in the simulation exercises of the next section. Discrete unobserved heterogeneity has been taken into account by averaging the outcomes of 1000 replications. Each replication is a draw from the estimated discrete heterogeneity distribution.

Both models underpredict the outcome where both spouses do not participate. The unitary model overpredicts the outcome where both spouses choose to work the maximum number of hours; and where the husband works 50 hours while the wife does not participate. It is quite satisfactory that both models allocate couples to all possible choices. This illustrates the models' usefulness to deal with the intensive margin in addition to the mere choice between working and not working. All in all, both models do a relatively good job in terms of fit.

If we compare collective and unitary predictions, then it seems that the collective model obtains a somewhat better fit than the unitary model. On the one hand, chi-square goodness-of-fit tests of both models result in a firm rejection of the null hypothesis that predicted outcomes equal observed ones. Test statistics for the collective and the unitary model are equal to respectively 147.23 and 175.76, which are way above the critical value $\chi^2_{0.05;15} = 25.00$. On the other hand, we can also reject the null hypothesis that predicted outcomes obtained by the unitary model equal those obtained by the collective model (test statistic is equal to 33.27 which is to be compared with $\chi^2_{0.05;15}$). One could tentatively argue that the above results taken together provide evidence that the collective model fits the data slightly better than the unitary model.

[Table 5 about here]

[Table 6 about here]

5 Some illustrative simulations

5.1 Social Security reforms

We now discuss some simulation results by means of the above models. Several proposals to increase labor force participation of the elderly have been made by policy makers. In this section, we discuss two such illustrative proposals. We will in addition consider a third, less realistic, reform to further illustrate what may go wrong when using a unitary model instead of a more appropriate collective model.

The first simulation that we focus on is the abolition of the earnings test. Currently, Old Age Social Insurance benefits are reduced if one has a labor income above some threshold if one is younger than the normal retirement age. Up to 2000, this was also the case for beneficiaries older than the normal retirement age. In order to study possible disincentive effects of this rule, we simulate the complete removal of the earnings test for both beneficiaries below and above the normal retirement age.[7] Earlier results in the literature gave a rather diffuse picture of the labor supply effects of the earnings test. Reimers and Honig (1996), for example, conclude that women are not affected by the earnings test, while there is some evidence that men are deterred from working by this rule. Friedberg (2000) comes to similar results for older men. Gruber and Orszag (2003), on the contrary, claim that there is no robust influence on the labor supply of men, while there is some evidence that the earnings test has an impact on the labor supply decisions of women. All of these papers used implicitly or explicitly static models to characterize results. Using the estimated models here provides a structural interpretation.

The second proposal is the elimination of the spouse benefit or spouse allowance. According to current Social Security rules, a spouse is entitled to the maximum of the own benefit and half of the Primary Insurance Amount of the other spouse, given that he or she is eligible for Old Age Social Insurance benefits, to which own actuarial adjustments are imposed. It is clear that this rule may have an impact on labor supply decisions of the individuals in married couples. Blau (1997), for example, although using data from the 1970s (the Retirement History Survey), concludes that the elimination of the spouse allowance and its replacement by a system of earnings sharing decreases husbands' participation rates, while there is a positive impact for married women. The effects were however rather minimal. Following Blau (1997), we also replace the spouse allowance by income sharing among spouses.

The third reform considered entails a substantial lump sum transfer of 30,000 dollar to each elderly couple. Although unrealistic, this reform can be viewed as a drastic (and not very targeted) policy measure to eliminate poverty among the elderly. Note that this reform, which expands the budget set, should be unambigously positive according to a sound unitary model. In a collective world, this is not necessarily the case: it may be the case that such a transfer

[7] Note that we have taken into account the removal of the earnings test in 2000 for the estimation since our sample covered the period 1992-2002.

substantially reduces the bargaining position of one of the spouses, which itself has an effect on that spouse's utility.

5.2 Simulation results

The baseline situation in the simulation exercises consists of the subsample of married couples, with at least one spouse aged between 62 and 65 in the year 1996. This subsample, which is the target group of the earnings test reform, consists of 881 couples. Starting from these baselines (one for each theoretical model), we proceed discussing the impact of the social security reforms on hours worked.

Table 7 gives the percentage point differences between predicted frequencies after the abolition of the earnings test and baseline frequencies. As is clear from the table, the impact of the reform is rather moderate. Both the collective and the unitary models predict that women, and to a lesser extent men, will increase participation if the earnings test is eliminated. The impact on participation is slightly more pronounced according to the collective model. Male participation increases by about 0.53 (0.42) percentage point on the basis of the collective (unitary) model, while female participation increases by 1.88 (1.39) percentage point according to the collective (unitary) model. Note that women who start working do this on average in part-time terms. All in all, unitary and collective predictions for this reform are qualitatively and quantitatively similar. Both men and women seem deterred from working by the earnings test.

[Table 7 about here]

Simulation results with respect to the abolition of the spouse allowance can be found in Table 8. Contrary to the first reform considered, collective and unitary models do not obtain qualitatively similar results. Although both models predict that male participation slightly decreases, while females increase participation (which points to disincentive effects associated with the allowance for women), they deviate from each other in terms of what happens with specific outcomes. For example, the collective model predicts that some husbands with a full-time working wife will also start working. The unitary model, however, predicts a decrease in male participation when the woman works full-time. Similar deviations can be found for other joint labor supply choices. It is worth stressing that our simulations, as far as the impact on aggregate participation is concerned, are in line with those obtained by Blau (1997), whose model is embedded in the unitary approach. Taking into account intra-household bargaining aspects, however, comes to a different conclusion on how the aggregate impact on participation is obtained. This illustrates that both theoretical approaches may deviate from each other in terms of a more refined positive behavioral analysis. Note that these deviations may entail normatively different conclusions about the welfare economic impact of a reform (cf. infra).

[Table 8 about here]

Important deviations between both models in terms of the impact on hours worked are more clearly illustrated by means of the third reform considered. Results are shown in Table 9. The collective and unitary models' predictions differ both qualitatively and quantitatively from each other; and this in a rather substantial way. The unitary model, for example, predicts that the number of households where no one participates will increase by 3.84 percentage point after the introduction of a lump sum transfer to the elderly. The collective model, on the contrary, predicts a decrease of this choice option by 1.36 percentage point. Further, both models predict qualitatively different gender specific reactions to the reform. According to the unitary model, a substantial number of women (5.05 percentage point) and men (1.58 percentage point) stop working after receiving the lump sum transfer. The collective model, however, predicts that a relatively high number of women will start working after the reform (4.33 percentage point), while some men will leave the labor market (1.86 percentage point in total). Note that this difference can be nicely explained in terms of the theoretical models. In the unitary model, a lump sum transfer implies an unambiguous outerward shift of the budget set. If leisure is a normal good, this implies a decrease in hours supplied. As far as the collective model is concerned, things are more complex. Firstly, as is the case for the unitary model, the lump sum transfer generates an income effect that, *ceteris paribus*, reduces labor supply if leisure is a normal good in the individuals' preferences. Secondly, however, there is the intra-household bargaining aspect. It turns out that the lump sum transfer is in general favorable to the wife. The transfer increases the household's non-labor income, which improves the bargaining position of the female. Moreover, it decreases the husband's relative earning capacity (due to progressive taxation); which further increases the female's Pareto weight. Taken together, the reform improves females' bargaining positions. Given the specific individual preferences (think of the different externalities in males' and females' utility functions), this explains the husbands' decrease in hours supplied and the females' increase in participation.[8]

[Table 9 about here]

An issue that has been ignored up to now is the impact of the reforms on individuals' (in the collective model) and couples' welfare levels. Table 10 summarizes the effects on the different utility levels. It should be stressed that the unitary model does not allow to make a

[8] We also simulated the third reform by holding the Pareto weights constant in the collective model. Results obtained are similar to those obtained by means of the unitary model. E.g., males (females) increase participation by 1.88 (6.10) percentage points, which is qualitatively and quantitatively close to the unitary model's results. This further illustrates the important dimension added by the collective model's non-constant Pareto weights with respect to wages and/or non-labor income.

distinction between households and individuals, since the household is assumed to behave as a single decision maker. As is usually done in welfare economic evaluations based on a unitary model, we will therefore assume that if a household's utility level is increased (decreased) due to a reform, then this implies that both male's and female's utility levels equally increased (decreased). The collective model's ability to distinguish between couples and individuals is a major advantage in welfare economic evaluations. Table 10 clearly illustrates that if a reform is beneficial to a couple, that this is not necessarily the case for both individuals that form the couple. For example, it turns out that for about 3 percent of the couples, there is a divergent impact of the abolition of the earnings test on spouses' utility levels. Some may argue that this difference across models is rather negligible The third reform, however, illustrates that the welfare economic impact can be very different across reform evaluations that are based on a different behavioral approach. Although the lump sum transfer is unambiguously welfare improving for almost all the couples according to the unitary model, this is certainly not the case for the collective model. It turns out that for 8.36 percent of the couples, the wife is better off after the reform, while the husband is worse off. On the other hand, a reverse conclusion is obtained for 1.19 percent of the couples. In our opinion, this is a clear illustration of what can go wrong when using a unitary model instead of a more appropriate collective model.

[Table 10 about here]

6 Conclusion

We presented a structural model to study labor supply behavior of elderly households. Specific to the model is that it is embedded in the collective approach to household behavior. By doing this, we explicitly take into account that spouses may have different preferences over leisure and consumption. Bargaining is involved to reach observed household allocations, which are assumed to be Pareto efficient. We do not only focus on the extensive margin (working versus being retired), but also on the intensive margin (how many hours are worked). The model allows for general externalities with respect to spouses' leisure. We presented a novel identification strategy for such a general model. Preferences and the intra-household bargaining process are identified by making use of panel data with couples and individuals who turned into widow(er)hood in the covered period, along with the assumption that preferences do not change (except for possible health shocks that can be controlled for), in such an event.

We applied our model to a sample of US households coming from the first six waves of the Health and Retirement Study. Many of the model's parameters are significantly estimated and have implications that are according to intuition. Health effects are important. Moreover, they illustrate that health is multidimensional and that different measures of health all have their place in a realistic labor supply model. Important from the model's point of view is that the spouses' Pareto weights in the intra-household allocation process significantly depend on wages and the household's non-labor income. This implies a strong rejection of the standard unitary model, where household preferences are price-independent. Note that this result confirms earlier results obtained in a consumption context (see Browning and Chiappori, 1998) and in a labor supply context with less realistic egoistic preferences (see Fortin and Lacroix, 1997, and Vermeulen, 2005). The estimated externalities of leisure are largely in line with results from other studies (e.g. Gustman and Steinmeier, 2000;2004). Contrary to previous models, the identification of externality effects, and of preferences in general, relied solely on the assumption that couples made Pareto efficient decisions.

Finally, we conducted simulations of three illustrative reforms, among others the widely discussed proposals to eliminate the earnings test and the spouse benefit under Social Security. Results were compared with those obtained by means of a unitary model, which functional specification is closely related to that of our collective model. The three simulations demonstrated that one cannot make an a priori unambiguous statement on whether both models come to different conclusions about the impact of a reform. Simulation results associated with the abolition of the earnings test were quite similar across both structural models; this both in terms of the impact on hours worked and the welfare economic effects of the reform. Collective and unitary models, on the contrary, did not come to qualitatively the same results as far as the abolition of the spouse benefit is concerned. Although both models predict that male participation slightly decreases, while females increase participation, they deviate from each other in terms of how this aggregate effect on participation is obtained. A third reform, which entails a rather unrealistic lump sum transfer to elderly couples, illustrates that deviations in a reform's behavioral and welfare economic impact can be substantial.

What can we now conclude from the above results? Firstly, they illustrate that the collective model is a potentially powerful tool to gain more insight on individual preferences, even when allowing very general externalities within a household. This is an aspect that is entirely ignored by the unitary approach. Secondly, the unitary model is strongly rejected by the data used, which, as indicated earlier, is a recurrent result in the literature. Thirdly, notwithstanding the rejection of the unitary model, it is not necessarily the case that predictions obtained by unitary and collective models always strongly deviate from each other. Our simulation results indicate that this depends on the simulation at hand. On the basis of our results, one could even argue that the representation of household behavior does not matter much if one considers realistic reforms. This conclusion, however, is too preliminary. Part of our results can be explained by the specific data used in this paper. Given the important state dependence in elderly couples' labor supply choices, there is not much room to obtain very different behavioral reactions on realistic social security reforms. As shown by Bargain et al. (2006), distortions due to the use of a unitary model can be very important for realistic tax-benefit reforms that apply to younger couples.

Given the above considerations, it seems interesting to further invest time in the development of collective models to deal with elderly individuals' labor supply decisions. One next, but far from easy, step is to alleviate the assumption that individuals are myopic; any future value from their actions is not considered in the household decision-making process. A dynamic programming approach would be feasible and we intend to follow that line in future research.

References

[1] Akerlof, G. and R. Kranton (2000), "Economics and identity", *Quarterly Journal of Economics*, 115, 715-753.

[2] Bargain, O., M. Beblo, D. Beninger, R. Blundell, R. Carrasco, M.-C. Chiuri, F. Laisney, V. Lechene, N. Moreau, M. Myck, J. Ruiz-Castillo and F. Vermeulen (2006), "Does the representation of household behavior matter for welfare analysis of tax-benefit policies? An introduction", *Review of Economics of the Household*, 4, 99-111.

[3] Blau, D. (1997), "Social security and the labor supply of older married couples", *Labour Economics*, 4, 373-418.

[4] Blau, D. (1998), "Labor force dynamics of older married couples", *Journal of Labor Economics*, 16, 595-629.

[5] Blundell, R., P.-A. Chiappori, T. Magnac and C. Meghir (2001), "Collective labor supply: heterogeneity and nonparticipation", IFS-working paper WP01/19, Institute for Fiscal Studies.

[6] Booth, A. and J. van Ours (2006), "Hours of work and gender identity. Does part-time work make the family happier?", CentER Discussion Paper 2006-02, Tilburg, CentER.

[7] Browning, M. and P.-A. Chiappori (1998), "Efficient intra-household allocations: a general characterization and empirical tests", *Econometrica*, 66, 1241-1278.

[8] Chiappori, P.-A. (1988), "Rational household labor supply", *Econometrica*, 56, 63-89.

[9] Chiappori, P.-A. and I. Ekeland (2002), "The micro economics of group behavior: identification", mimeo, Chicago, University of Chicago.

[10] Chiappori, P.-A., B. Fortin and G. Lacroix (2002), "Marriage market, divorce legislation and household labor supply", *Journal of Political Economy*, 110, 37-72.

[11] de Serres, A. and K.-Y. Yoo (2004), "Tax treatment of private pension savings in OECD countries and the net tax cost per unit of contribution to tax-favoured schemes", OECD Economics Department Working Papers 406, OECD Economics Department.

[12] Donni, O. (2003), "Collective household labor supply: nonparticipation and income taxation", *Journal of Public Economics*, 87, 1179-1198.

[13] Fortin, B. and G. Lacroix (1997), "A test of the unitary and collective models of household labour supply", *Economic Journal*, 107, 933-955.

[14] Friedberg, L. (2000), "The labor supply effects of the social security earnings test", *Review of Economics and Statistics*, 82, 48-63.

[15] Gruber, J. and P. Orszag (2003), "Does the social security earnings test affect labor supply and benefit receipt", *National Tax Journal*, 56, 755-773.

[16] Gustman, A. and T. Steinmeier (2000), "Retirement in dual-career families: a structural model", *Journal of Labor Economics*, 18, 503-545.

[17] Gustman, A. and T. Steinmeier (2004), "Social security, pensions and retirement behavior within the family", *Journal of Applied Econometrics*, 19, 723-737.

[18] Haider, S. and G. Solon (2000), "Nonrandom selection in the HRS social security earnings sample, RAND working paper DRU-2254-NIA, Santa Monica.

[19] Hamermesh, D. (2002), "Timing, togetherness and time windfalls", *Journal of Population Economics*, 15, 601-623.

[20] Heckman, J. and B. Singer (1984), "A method for minimizing the impact of distributional assumptions in econometric models for duration data", *Econometrica*, 52, 271-320.

[21] Health and Retirement Study (2003), (Waves 1-5/Years 1992-2000) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging. Ann Arbor, MI.

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

[23] Hurd, M.D., Smith, J.P. and J. Zissimopoulos (2004): "The Effects of Subjective Survival on Retirement and Social Security Claiming", *Journal of Applied Econometrics*, 19, 761-775.

[24] Kahneman, D. (1999), "Objective happiness", in D. Kahneman, E. Diener and N. Schwarz (eds.), *Well-being: the Foundations of Hedonic Psychology*, Russell Sage Foundation, New York, 3-25.

[25] Maestas, N. (2001), "Labor, love and leisure: complementarity and the timing of retirement by working couples", mimeo, Department of Economics, UC Berkeley.

[26] Mazzocco, M. (2006), "Household intertemporal behavior: a collective characterization and a test of commitment", mimeo, Madison, University of Wisconsin-Madison.

[27] Michaud, P.-C. (2005), "Labor force participation and Social Security claiming decisions of elderly households", mimeo RAND.

[28] Michaud, P.-C. and A. van Soest (2004), "Health and wealth of elderly couples: causality using dynamic panel data models", IZA Discussion Paper 1312, Bonn, IZA.

[29] Reimers, C. and M. Honig (1996), "Responses to social security by men and women. Myopic and far-sighted behavior", *Journal of Human Resources*, 31, 359-382.

[30] StClair, P., D. Bugliari, P. Pantoja, S. Ilchuk, G. Lopez, S. Haider, M. Hurd, D. Loughran, C. Panis, M. Reti and J. Zissimopoulos (2002), *RAND HRS Data Documentation*, RAND Corporation, Santa Monica.

[31] van Soest, A. (1995), "Structural models of family labor supply. A discrete choice approach", *Journal of Human Resources*, 30, 63-88.

[32] Vermeulen, F. (2005), "And the winner is... An empirical evaluation of unitary and collective labour supply models", *Empirical Economics*, 30, 711-734.

[33] Vermeulen, F. (2006), "A collective model for female labour supply with nonparticipation and taxation", *Journal of Population Economics*, 19, 99-118.

[34] Vermeulen, F., O. Bargain, M. Beblo, D. Beninger, R. Blundell, R. Carrasco, M.-C. Chiuri, F. Laisney, V. Lechene, M. Myck, N. Moreau and J. Ruiz-Castillo (2006), "Collective models of household labour supply with non-convex budget sets and non-participation: a calibration approach", *Review of Economics of the Household*, 4, 113-127.

Appendix A Wages

For non-workers, wages need to be imputed. We follow Gustman and Steinmeier (2000) and estimate fixed effect models on the sample of wage earners. We do these regressions for females and males separately. We include age, tenure and experience quadratic forms, time, industry and occupation dummies. We predict wages using information across waves. Starting from the last observation with a non-missing wage available, we impute forward using the age, experience and tenure estimated profiles for non-workers. Then we move to the next observation with a missing wage. If there is a wage available in an earlier wave, we use that wage along with the wage in the year following the missing year to intrapolate the missing wage. We proceed recursively in this way until we meet the earliest wave with an available wage. At that point, if there are earlier waves without a wage, we use again the age, experience and tenure profile to impute for these years. Finally if there is no wage available for an observation over all waves, we use the prediction from the fixed effect regression where we set the fixed effect to the observed mean, along with occupation and industry if not available from longest held job information. Fixed effect regression results can be obtained from the authors upon request.

Appendix B Computation of social security benefits and income tax

We use SSA earning records up to 1991 along with earnings and labor force status information from 1992 to 2002 to compute social security benefits. We attribute 8 quarters of coverage if a worker reports working for pay in a given wave (waves are every two years). This is added to quarters of coverage prior to 1991 found in the earnings record to determine eligibility to benefits. Yearly earnings maxima that enter the calculation of social security benefits are applied. We follow closely the Social Security Handbook in computing benefits. In particular, we take into account the minimum Primary Insurance Amount, changing actuarial adjustments by cohort, the earnings test along with its abolition on the 65-70 portion in 2000, the spouse and survivor benefit. We do not consider benefits paid to divorced respondents or other dependent respondents.

We consider Social Security and Medicare taxes, state and city taxes as well as Federal income tax when computing tax liabilities of households. We use the formula provided in the OECD publication Taxing Wages (OECD, 2005) which takes the tax schedule in Detroit, Michigan to calculate local taxes. Since conditions for the use of restricted access Earnings records preclude us from merging geocode information (to get states and cities), we approximate local taxes using that in Detroit, Michigan. In any case, local taxes are rather low in the U.S. Information found in Pensions at a Glance (OECD, 2004) and De Serres and Yoo (2004) is used to adapt the tax schedule to incorporate specific tax credits available to the elderly along with taxation of income from investment.

Table 1: Summary statistics

Means/proportions reported	Alive in 2002		Died 1992-2002	
	Husband	Wife	Widower	Widow
Age	63.7	60.3	67.2	64.2
Health				
Severe conditions ever	0.35	0.25	0.49	0.31
Mild conditions ever	0.72	0.71	0.79	0.80
Excellent/very good self reported health	0.49	0.55	0.39	0.49
Good self reported health	0.32	0.30	0.29	0.33
Fair/poor self reported health	0.20	0.15	0.32	0.18
Bad mental health (CESD>2)	0.10	0.16	0.41	0.32
Schooling				
Less than high school	0.06	0.04	0.13	0.05
High school/GED	0.52	0.58	0.56	0.71
College and more	0.42	0.38	0.31	0.24
Labor market				
Working for pay	0.53	0.45	0.33	0.34
Net hourly wage	14.2	8.4	10.1	6.5
Non-labor incomes				
Average indexed monthly earnings	2441.0	713.4	2125.9	678.8
Other non-labor income (weekly)	588.6	477.0	329.7	372.4
Observations	9271		138	487

Table 2: Hours worked per week

	Wife				Total	Widower
Husband	40	30	15	0		
50	6.5	1.9	2.4	6.9	17.8	4.4
40	7.2	2.2	2.7	7.7	19.9	16.7
25	3.9	1.5	2.5	7.1	15.0	12.3
0	7.2	2.4	4.1	33.7	47.4	66.7
Total	24.8	7.9	11.8	55.4	100	100
Widow	16.0	6.6	11.5	65.9	100	

Note: Entries are in percentages.

Table 3: Point estimates of the collective model

	Male		Female	
	Est.	S.e.	Est.	S.e.
Own leisure				
First point of support	-7.909*	0.363	-8.204*	0.359
Second point of support	-2.607*	0.454	-1.416*	0.379
Third point of support	5.987*	0.932	17.579*	2.117
Age/10	4.100*	0.313	5.046*	0.345
College	-0.255	0.205	-0.429	0.249
Non white	-0.459	0.298	-0.170	0.403
Good self reported health (Excel./very good ref.)	0.617*	0.185	0.214	0.196
Fair/poor self reported health (Excel./very good ref.)	1.579*	0.276	1.399*	0.308
Onset of severe health condition	0.557*	0.207	1.010*	0.256
Onset of mild health condition	0.667*	0.199	0.180	0.223
Bad mental health	0.770*	0.276	-0.036	0.232
Partner's leisure	0.455*	0.082	-0.335*	0.060
Consumption				
Constant	2.506*	0.569	3.903*	0.470
Age/10	-0.328	0.273	-1.493*	0.273
College	-0.963*	0.318	0.333	0.296
Non white	1.019*	0.309	0.326	0.239
Good self reported health (Excel./very good ref.)	0.064	0.374	0.135	0.326
Fair/poor self reported health (Excel./very good ref.)	-0.791*	0.287	0.667*	0.276
Onset of severe health condition	0.325	0.291	-0.446	0.267
Onset of mild health condition	-0.027	0.379	-0.611*	0.267
Bad mental health	-0.171	0.437	-0.161	0.362
Fixed cost parameters				
Participation costs (weekly)	-118.64*		1.034	
Part-time male	-0.295*		0.035	
Part-time female (15 hours)	-0.690*		0.045	
Part-time female (30 hours)	-0.396*		0.041	
Husband's Pareto weight				
Constant	-0.196*	0.132		
Relative earning capacity	0.133*	0.013		
Non labor income	-0.135*	0.014		
AIME husband	-0.082*	0.024		
AIME wife	0.051	0.032		

Note: Coefficients with an asterisk are significant at the 5 percent significance level.

Table 4: Discrete heterogeneity probabilities

		Female			Male marginal probability
		1	2	3	
Male	1	0.11	0.08	0.12	0.31
	2	0.16	0.28	0.04	0.48
	3	0.01	0.08	0.12	0.21
Female marginal probability		0.28	0.44	0.28	1.00
Correlation in unobserved heterogeneity across spouses					0.35

Table 5: Point estimates of the unitary model

	Leisure male		Leisure female	
	Est.	S.e.	Est.	S.e.
First point of support	-5.607*	0.073	-6.692*	0.336
Second point of support	-3.123*	0.268	-2.812*	0.118
Third point of support	1.529*	0.477	6.834*	1.206
Age/10	1.834*	0.093	2.694*	0.154
College	-0.268*	0.092	-0.145	0.152
Non white	-0.079	0.135	0.149	0.235
Good self reported health (Excel./very good ref.)	0.248*	0.082	0.108	0.117
Fair/poor self reported health (Excel./very good ref.)	0.670*	0.119	0.777*	0.178
Onset of severe health condition	0.284*	0.093	0.688*	0.156
Onset of mild health condition	0.288*	0.090	0.196	0.135
Bad mental health	0.324*	0.132	0.002	0.140
	Leisure interaction parameter			
	0.553*		0.067	
	Consumption			
Constant	3.093*		0.260	
Non white couple	0.036		0.207	
	Male characteristics		Female characteristics	
Age/10	-0.563*	0.162	-0.519*	0.159
College	-0.522*	0.162	-0.030	0.149
Good self reported health (Excel./very good ref.)	0.476*	0.153	0.073	0.128
Fair/poor self reported health (Excel./very good ref.)	0.071	0.188	-0.087	0.182
Onset of severe health condition	-0.566*	0.148	0.721*	0.156
Onset of mild health condition	0.192	0.156	-0.079	0.138
Bad mental health	-0.297	0.203	-0.189	0.154
Fixed cost parameters				
Participation Cost	-121.98*		0.811	
Part-time male	-0.247*		0.035	
Part-time female (15 hours)	-0.646*		0.046	
Part-time female (30 hours)	-0.347*		0.042	

Note: Coefficients with an asterisk are significant at the 5 percent significance level.

Table 6: Observed and predicted distributions

Observed outcomes	Wife			
Husband	40	30	15	0
50	8.00	8.05	4.15	7.33
40	2.00	2.46	1.08	2.31
25	3.02	3.08	2.15	3.38
0	7.89	9.33	6.61	29.17
Predicted outcomes collective model	Wife			
Husband	40	30	15	0
50	0.94	-1.76	0.50	1.84
40	0.58	-0.38	0.60	1.39
25	0.12	-0.24	0.30	2.33
0	2.26	0.00	1.64	-10.13
Predicted outcomes unitary model	Wife			
Husband	40	30	15	0
50	3.33	-0.72	1.30	3.15
40	0.57	-0.57	0.56	1.47
25	-0.41	-0.86	-0.07	2.12
0	0.69	-1.46	1.04	-10.14

Note: Observed outcomes are in percentages. Predicted outcomes are in percentage point differences with observed outcomes.

Table 7: Simulation 1: Abolition earnings test

Predicted outcomes collective model	Wife			
Husband	40	30	15	0
50	0.07	0.13	0.16	0.02
40	0.03	0.06	0.09	-0.03
25	0.03	0.09	0.17	-0.30
0	0.10	0.36	0.58	-1.57
Predicted outcomes unitary model	Wife			
Husband	40	30	15	0
50	0.06	0.12	0.19	-0.03
40	0.02	0.05	0.08	-0.05
25	0.02	0.06	0.11	-0.20
0	0.05	0.22	0.42	-1.11

Note: Predicted outcomes are in percentage point differences with collective and unitary predicted baseline outcomes.

Table 8: Simulation 2: Abolition spouse benefit

Predicted outcomes collective model	Wife			
Husband	40	30	15	0
50	-0.04	-0.03	-0.01	-0.18
40	0.00	0.01	0.02	0.02
25	0.00	0.02	0.05	0.04
0	-0.07	0.02	0.15	-0.02
Predicted outcomes unitary model	Wife			
Husband	40	30	15	0
50	0.04	0.05	0.04	-0.20
40	0.02	0.03	0.03	-0.06
25	0.03	0.04	0.05	-0.12
0	0.09	0.19	0.23	-0.46

Note: Predicted outcomes are in percentage point differences with collective and unitary predicted baseline outcomes.

Table 9: Simulation 3: Lump sum transfer

Predicted outcomes collective model	Wife			
Husband	40	30	15	0
50	0.53	0.36	0.03	-1.15
40	0.07	-0.01	-0.14	-0.90
25	0.26	0.14	-0.11	-0.93
0	1.69	1.24	0.28	-1.36
Predicted outcomes unitary model	Wife			
Husband	40	30	15	0
50	-0.69	-0.71	-0.40	0.05
40	-0.19	-0.21	-0.11	0.16
25	-0.18	-0.23	-0.06	0.99
0	-0.77	-1.02	-0.47	3.84

Note: Predicted outcomes are in percentage point differences with collective and unitary predicted baseline outcomes.

Table 10: Welfare effects

	Collective model		Unitary model	
Simulation 1: Abolition earnings test				
	Female loses	Female gains	Female loses	Male gains
Male loses	0.46	2.33	1.65	0.00
Male gains	0.48	96.73	0.00	98.35
Simulation 2: Abolition spouse benefit				
	Female loses	Female gains	Female loses	Male gains
Male loses	40.21	0.82	41.62	0.00
Male gains	1.26	57.71	0.00	58.38
Simulation 3: Lump sum transfer				
	Female loses	Female gains	Female loses	Female gains
Male loses	0.27	8.36	1.61	0.00
Male gains	1.19	90.17	0.00	98.39

Note: Entries are in percentages.