Informal care and labour force participation among middle-aged women in Spain

**David Casado-Marín · Pilar García-Gómez ·
Ángel López-Nicolás**

Abstract   Informal care is today the form of support most commonly used in Spain by those who need help in order to carry out basic daily activities. The potential labour opportunity costs incurred by Spanish informal carers have not as yet been quantified. In this paper we use the Spanish subsample of the European Community Household Panel (1994–2001) to estimate an econometric model which we exploit to examine the effects of various types of informal care on labour market outcomes. Our results reveal the existence of non-negligible costs in terms of foregone employment for carers who live with the dependent person and/or provide more than 28 h of care/week. We also find that providing care for more than a year has negative effects on employment. Nonetheless, there seems to be no contemporaneous employment effects associated to either starting or ending an episode of care.

D. Casado-Marín
Institut d'Avaluació de Polítiques Públiques (IVALUA), Barcelona, Spain
e-mail: david.casado@ivalua.cat

P. García-Gómez (✉)
Department of Applied Economics. Erasmus School of Economics, Erasmus University Rotterdam,
P.O. Box 1738, 3000 DR, Rotterdam, The Netherlands
e-mail: garciagomez@ese.eur.nl

Á. López-Nicolás
Department of Economics, Technical University of Cartagena, Cartagena, Spain
e-mail: angel.lopez@upct.es

D. Casado-Marín · P. García-Gómez · Á. López-Nicolás
Centre de Recerca en Economia i Salut (CRES), Pompeu Fabra University, Barcelona, Spain

1 Introduction

The process of population ageing that developing countries are likely to undergo in the coming decades is one of the phenomena whose social and economic consequences cause most concern. The already classical debates on the future sustainability of the systems of public pensions (Jimeno et al. 2008) and health care (Ahn et al. 2005) have extended more recently to the discussion on how to provide and fund the care required by older people who cannot look after themselves.

To date, in Spain and throughout southern Europe, the family has characteristically been the main source of support to meet the needs of dependent people (OECD 2005). Thus, in the particular case of Spain (Casado 2006), the needs of 74% of all dependent people are met solely by informal carers,[1] and the figure rises to 85% if we include those who combine informal care with some other source of support of a formal nature (for example, home care). The extraordinary vigour of this family model, undoubtedly made possible by the low labour force participation rates of current cohorts of middle-aged women and their predecessors, has until now enabled the public sector to take on a subsidiary role: only when the family is unable or unwilling to help or does not exist, and always depending on the economic capacity of the older person concerned, is the required care publicly funded (Fundació Institut Català de l'Envelliment 2004).

Following the lead of other European countries, which have had universal public long-term care systems for some years, Spain is now developing a similar scheme known as the National Long-Term Care System (Sistema para la Autonomía y Atención a la Dependencia or SAAD) over the period 2007–2015.[2] One of the main goals pursued through the SAAD, in addition to eliminating means testing for access to public long-term care services, is to strike a new balance between formal and informal care that is compatible with the higher labour force participation rates of future cohorts of middle-aged women. Specifically, given that a steep rise is expected in the percentage of women that will be in employment when someone in their family becomes dependent, the development of community services (home care, day centres and so on) through the SAAD seeks to make providing a certain amount of informal care compatible with having a paid job. This would not only avoid the negative consequences at an individual level associated with leaving the labour market (loss of income, smaller future pension, etc.) but would also make it possible to take on family responsibilities without jeopardising the macroeconomic objective, enshrined in the Lisbon Agenda, of increasing the female labour force participation rate to 60% over the next decade.

However, if the SAAD is really to reach the goals that have been set, the design of the new benefits must be based on a profound knowledge of how today's middle-aged women combine (or fail to combine) informal caregiving with doing paid work. Although several studies have been published that examine the existence of

---

[1] Carers are considered to be informal when they receive no financial remuneration for the help they give.

[2] See OECD (2005) for an up-to-date description of the long-term care systems of EU countries.

labour opportunity costs associated with informal care in other countries, to our knowledge there is no specific study on this issue for the Spanish case.

Thus, in view of the above, the main aim of this paper is to analyse to what extent women who give informal care today incur labour opportunity costs as a result of doing so. To this end, we use the eight waves of the European Community Household Panel (1994–2001) to estimate a dynamic ordered probit model that enables us to examine the effects of various types of informal care on labour behaviour. The results obtained indicate the existence of labour opportunity costs for those women who live with the dependent person they care for, but not for those who care for someone outside the household. Furthermore, whereas providing care for more than a year has negative effects on labour force participation, the same cannot be said of those who just "start caregiving" or just "stop caregiving". That is, there seem to be no contemporaneous employment effects associated to starting or ending an episode of care. In addition, the results also show that the labour opportunity costs occur when women are providing more than 28 h/week of care.

2 Informal care and labour market outcomes

The main methodological challenge faced when analysing the relationship between informal care and labour behaviour is that informal care is usually endogenous to the process determining labour outcomes. This endogeneity may arise from either of two types of elements. First, considering that the two activities compete for the potential carer's time, allocations to one or the other will be the result of a simultaneous choice process in which other factors also come into play: the use of formal services, the previous employment status of the potential carer, the availability of other informal carers, etc. And second, even in the event of being able to model the simultaneity of the choices and the influence of the factors mentioned above, we may still be faced with a problem of endogeneity if the individuals possess unobserved characteristics correlated with both the propensity to care for a dependent relative and the propensity to participate in the labour market.

On the basis of the definition of the two problems described above, henceforth referred to as the simultaneity problem and the unobserved individual heterogeneity problem, previous studies examining the relationship between informal care and labour force participation can be classified according to whether they deal with both problems, only one of them, or neither. Starting with the last of these groups of studies, the two papers by Carmichael and Charles (1998, 2003) analyse the relationship between informal care and labour behaviour in the UK, using cross-section data from the General Household Survey of 1985 and 1990, respectively. The results obtained by these authors, undoubtedly the least robust from a methodological point of view in that they assume informal care to be exogenous in both cases, show this variable to have negative effects on both the probability of being employed and the number of hours worked.

A second group of studies have attempted, despite their use of cross-section data, to tackle the possible endogeneity of informal care by estimating the labour equations of interest with instrumental variables (Wolf and Soldo 1994; Ettner 1995,

1996; Heitmueller 2007; Bolin et al. 2008). The instruments used in these studies typically include the health status of the parents of caregiving and non-caregiving women (as worse health status is assumed to require more intensive care) and the number of siblings these women have (as the intensity of the informal care to be given will be lower if there are alternative carers).

The results obtained by this second group of studies tend to confirm the existence of labour opportunity costs associated with informal care. Thus, with the exception of Wolf and Soldo (1994), who find no effect either on the probability of being employed or on the number of hours worked, the rest of the papers mentioned above point to the existence of considerable labour effects for women who provide care, despite using databases referring to different countries and different time periods. For instance, Ettner (1995, 1996) obtains different results than Wolf and Soldo (1994) for the US: firstly, a significantly lower participation rate is detected for women providing informal care to a live-in dependent person; and secondly, although women providing care to someone outside the household do not seem to have a lower participation rate, their number of working hours worked is lower than that of the rest of women all else held equal.

Ettner's results have been confirmed in part by more recent studies conducted using European data. Heitmueller (2007) uses instrumental variables to estimate, on the basis of the 2002 wave of the British Household Panel Survey (BHPS), the effect on labour force participation of providing care both inside and outside the household. His results show that only in the first instance there is a statistically significant decrease in the probability of being employed. Within the same empirical framework, Bolin et al. (2008) use data from the Survey of Health, Ageing and Retirement in Europe (SHARE) to analyse the associations between hours of informal care outside the household provided to an elderly parent and the probability of employment, hours worked and wages. Their results suggest that providing informal care to one's elderly parents is associated with significant costs in terms of foregone labour market opportunities and that these effects vary between European countries. They cannot reject the null hypothesis of informal-care giving being exogenous in any of their IV-estimations.

Crespo (2007) uses data from the first available wave (2004) of the SHARE to calculate the effects of informal care on female labour force participation in two triplets of countries in southern Europe (Spain, Italy and Greece) and northern Europe (Sweden, Denmark and The Netherlands), by estimating a bivariate probit model that controls for the endogeneity of the caregiving decision. Her results indicate that women who provide an "intense" level of care—i.e., live in the same household as the dependent person, or give daily care elsewhere—have a lower probability of participating in the labour force in the three southern European countries as well as in the three northern ones.

A third group of studies is characterised by concentrating on unobserved individual heterogeneity using longitudinal data. In particular, using the first three waves of the European Community Household Panel (1994–1996), Spiess and Schneider (2003) employ a difference-in-difference model to examine the impact on number of hours worked of three "stages" of informal care: starting caregiving, continuing caregiving, and stopping caregiving. Their results, which cannot be

broken down into countries due to the small sample size, show that in the southern European (Mediterranean) group of countries it is the continuation of care provision—not the fact of starting—what affects the number of hours worked. Conversely for the rest of the countries analysed (non-Mediterranean Europe) the results show exactly the opposite.

In turn, Viitanen (2005) uses all eight waves of the ECHP (1994–2001) to examine the effects of informal care on the labour behaviour of women aged 20–59, with the aid of dynamic probit models that take into consideration unobserved individual heterogeneity (random effects), state dependence, and the attrition biases that tend to appear when working with panel data. The results obtained by this author, which unlike those of Spiess and Schneider are country-specific, indicate that informal caregiving only has a negative influence on the probability of being employed in the case of Germany. However, when replicating the study taking specific subgroups of women into consideration, Viitanen detects significant effects in several countries among middle-aged women (Belgium, Finland and Germany) and among single women (Greece, The Netherlands, Italy and Germany).

More recently, Heitmueller (2007) also analyses the relationship between caregiving and labour market participation by estimating fixed effects models using the first 12 waves of the BHPS (1991–2002). The results in this case are similar to the ones obtained using an instrumental variable approach using only data from 2002 (see above).

The last group of studies that have examined the relationship between informal care and labour behaviour have tackled the two issues of simultaneity and unobserved heterogeneity. Thus, on the basis of two waves of the Health and Retirement Study, Johnson and Lo Sasso (2000) estimate a simultaneous equation model with panel data to analyse the impact of caring for a dependent parent for more than 100 h a year on the annual number of hours worked. Their results indicate that the annual labour supply of middle-aged (aged 53–63) carers is 23 and 28% lower (among men and women, respectively) than that of non-carers. In a recent paper based on 13 waves of the British Household Panel Survey (1991–2003), Heitmueller and Michaud (2006) estimate a dynamic bivariate probit that adjusts for reverse causality, state dependence and individual heterogeneity. The model is estimated separately for two distinct samples of carers, which yield different results in each case: when they consider everyone who cares for another person, whether at home or elsewhere, labour force participation does not appear to be lower than that of non-carers; but when the model is estimated for the subsample of co-resident carers, the results show a lower labour force participation, both among women (−6%) and among men (−4.7%).

We contribute to the existing literature by looking at the effect of various characteristics of caregiving (location, whether there is a transition into or out of caregiving or simply a continuation and the number of hours of care) on the probability of being employed full-time or part-time. We exploit the panel structure of the Spanish sample of the ECHP, allowing for the presence of individual specific unobserved heterogeneity and state dependence, to estimate a dynamic ordered probit model. In this regard, we follow perspective similar to Viitanen (2005). However, our approach carefully considers different caregiving states, as we suspect

that Viitanen's (2005) results suggesting that caregiving affects the probability of employment only in the case of German women can be explained by having considered co-residents and non-coresidents together. Thus, we conceive caregiving at home or elsewhere as possibly having different effects on labour market outcomes, as the results obtained by Heitmueller (2007) and Heitmueller and Michaud (2006) for the UK suggest. In addition, we exploit the dynamics of caregiving, i.e., not only do we allow the effects of caregiving to be different depending on the number of hours of care provision, but also on whether it is a recent situation or a mere continuation (first years vs. subsequent years). We also analyse the propensity to be employed once the individual stops giving care. One further distinguishing feature of our approach consists in testing for the assumption of (conditional) exogeneity of caregiving status in our equation for labour outcomes.

3 Data

3.1 Sample analysed and selection of variables

The ECHP has a series of characteristics that make it an interesting database for analysing possible relationships between informal care and labour behaviour. First of all, while subjects remain in the panel, the survey provides ample information on their labour behaviour (employment status, whether full or part-time, number of hours worked, salary, etc.). Also, the ECHP enables us to characterise informal care fairly precisely: subjects are asked not only whether they care for a dependent adult, but also how many hours of care they provide per week, whether or not the dependent lives in the same household, and so on.[3] And lastly, with a view to controlling for the influence of other variables on the labour behaviour of carers and non-carers, the survey contains ample socioeconomic information not just on the interviewee (age, gender, educational level, health status, employment record, income from labour and property, etc.) but also on the rest of the members of the household.

For our analysis we took the subsample of women residing in Spain and aged 30–60 who were in the panel in 1994 and participated in at least three consecutive years of the eight waves of the ECHP and supplied complete information on the variables that appear in Table 1 in all the waves in which they participated. We were able to use up to a maximum of 15,247 observations, corresponding to 3,859 individuals. Nonetheless the size of the estimating samples varies across models, as indicated at the bottom of Table 6 in Sect. 5.

As can be seen in Table 1, caregiving was characterised according to three alternative classifications: first, given that several research studies have detected different effects on employment depending on whether or not the dependent co-resides with the carer (Heitmueller 2007), we divided carers in our sample into

---

[3] The respondents are asked whether their present daily activities include, without pay, looking after children or other persons who need special help because of old age, illness or disability. We classify as informal carers women that were looking after other persons, so excluding children care.

Table 1 Variables included in the analysis

Labour behaviour	
No work	1 if not working, 0 otherwise
Part-time	1 if working part-time (working between 15 and 30 h/week), 0 otherwise
Full time	1 if working full-time (working at least 30 h/week), 0 otherwise
Informal care	
carer	1 if engaged in unpaid care of an adult dependent, 0 otherwise
carer_household	1 if engaged in unpaid care of an adult dependent within the household, 0 otherwise
carer_elsewhere	1 if engaged in unpaid care of an adult dependent outside the household, 0 otherwise
starting caregiving	1 if starting unpaid care of an adult dependent, 0 otherwise
continuing caregiving	1 if this is at least the second consecutive year of unpaid care of an adult dependent, 0 otherwise
stopping caregiving	1 if engaged last year in unpaid care of an adult dependent to whom she no longer gives care this year, 0 otherwise
carer <14	1 if engaged in unpaid care of an adult dependent less than 14 h/week, 0 otherwise
carer 14–28	1 if engaged in unpaid care of an adult dependent between 14 and 28 h/week, 0 otherwise
carer >28	1 if engaged in unpaid care of an adult dependent more than 28 h/week, 0 otherwise
Sociodemographic	
d3539	1 if aged 35–39, 0 otherwise
d4044	1 if aged 40–44, 0 otherwise
d4549	1 if aged 45–49, 0 otherwise
d5054	1 if aged 50–54, 0 otherwise
d5560	1 if aged 55–60, 0 otherwise
single	1 if single, 0 otherwise
widow	1 if widowed, 0 otherwise
sepdiv	1 if separated or divorced, 0 otherwise
nch04	Number of children in household aged 0–4
nch511	Number of children in household aged 5–11
sizehshd	Size of household
sahvgood	1 if self-assessed health status is very good, 0 otherwise
sahgood	1 if self-assessed health status is good, 0 otherwise
sahbad	1 if self-assessed health status is bad, 0 otherwise
sahvbad	1 if self-assessed health status is very bad, 0 otherwise
Northwest	1 if region of residence is Northwest, 0 otherwise
Northeast	1 if region of residence is Northeast, 0 otherwise
Madrid	1 if region of residence is Madrid, 0 otherwise
Centre	1 if region of residence is Centre, 0 otherwise
South	1 if region of residence is South, 0 otherwise
Canaries	1 if region of residence is Canaries, 0 otherwise
isced57	1 if highest completed educational level is tertiary, 0 otherwise
isced3	1 if highest completed educational level is secondary, 0 otherwise
lnincome_ot	Logarithm of the total income of other household members

Source authors, based on the ECHP

those who provide care at home and those who do so elsewhere; and then, since there is also some evidence that the effects of informal care on labour behaviour change over time (Spiess and Schneider 2003), women in our sample were classified into four possible dynamic states between t and $t + 1$: "starting caregiving", "continuing caregiving", "stopping caregiving" and "no caregiving in either period". In addition, we exploit the information on hours of care as the labour opportunity costs might appear only above a threshold of hours of care (Heitmueller 2007). In this respect, we define three dummy variables depending on whether the caregiver provides less than 14 weekly hours of care, between 14 and 28 or more than 28 h/week of informal care, as these are the categories available to respondents to the ECHP.

Furthermore, although the ECHP contains information on the number of hours worked, the employment status of the women that make up the sample is coded by means of a categorical variable that takes three possible values: "no work", "part-time" and "full-time". This is because we are interested in assessing the impact of caregiving on women's degree of integration into the labour market, and we believe that this impact—apart from the implicit change in hours worked—will tend to manifest itself as a transition between these three states. In addition, the number of hours worked in full-time employment tends to vary from job to job, and a woman will declare herself to be working part-time whenever her working day is shorter than the standard working day for that job. Therefore, using this categorisation enables us to work with a measure of employment status that implicitly takes into account the characteristics of the job as regards the length of the working day.

3.2 Descriptive analysis

The relevance of focusing on the subsample of women aged 30–60 is quite clear in view of the information contained in the graphs below. Specifically, on calculating the percentages of total men and women who stated that they were caring for an adult dependent in 1994 (Fig. 1), we find that the average prevalence among women was three times that of men (12% vs. 4%). Furthermore, with regard to the age groups that concentrate the largest proportion of women carers, it should be noted that middle-aged cohorts show prevalence rates of above 15% in all cases. Thus, the exclusion from our analysis of women younger than 30 and older than 60 is justified not only because additional factors are involved in determining the labour behaviour of both these groups (uncompleted education, abandonment of the labour market due to retirement, etc.) but also because most carers are not to be found in these two age groups.

If we look at the dynamic incidence of the event "starting caregiving", again notable differences can be seen both between men and women and between age groups (Fig. 2). Specifically, the cohorts of middle-aged women display the highest incidence rates, for two reasons: first, since dependency problems are concentrated in older people, mostly widows, the cohorts of individuals with the greatest probability of having a dependent parent are precisely those aged between 45 and 65; and second, owing to the gender bias that characterises the adoption of a
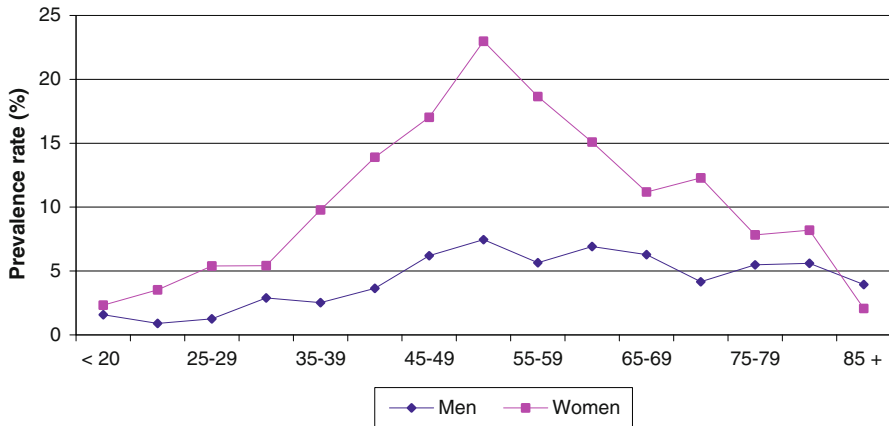
Fig. 1 Percentage of informal carers: Spain, 1994. *Source* authors, based on the ECHP
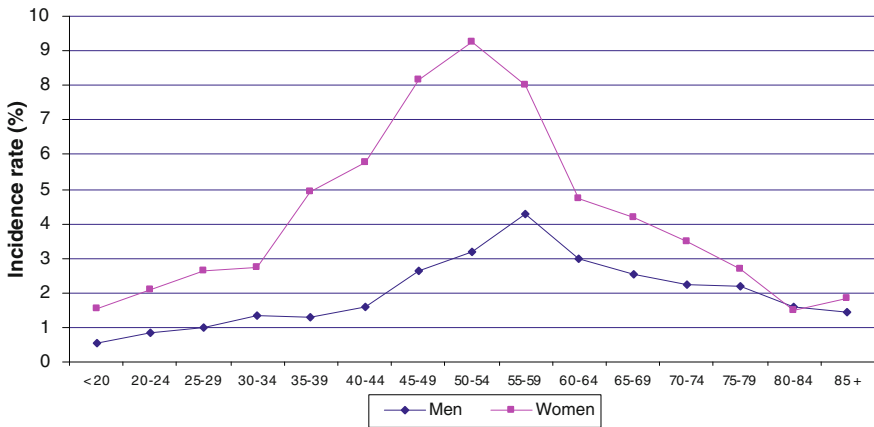
Fig. 2 Incidence rate of new carers (%): Spain, 1995–2001. Average incidence rates during the period. *Source* authors, based on the ECHP

caregiving role, it is generally the daughters and daughters-in-law of these dependent people who provide the required help. However, when we consider older cohorts (65 plus), the differences between men and women narrow, as the carers that appear in these cases are usually the dependents' spouses.

Table 2 presents a sample breakdown of the number of caregivers by type of care (within the household or elsewhere), and temporal sequence of care (start caregiving, continue caregiving or stop caregiving). For each category we also report the average number of hours of care provided per week. Note that the number of women that provide care at home roughly doubles the number of carers outside the household, and that the former seem to provide twice as many hours of care on

Table 2 Sample breakdown of caregivers by type of care and sequence of care (in percentages)

Type of care/sequence	Years						
	1995	1996	1997	1998	1999	2000	2001
Carer household	0.610 (46.48)	0.632 (44.7)	0.613 (48.95)	0.663 (50.22)	0.645 (44.53)	0.718 (40.76)	0.660 (45.87)
Carer elsewhere	0.390 (26.65)	0.368 (23.53)	0.387 (26.4)	0.337 (26.8)	0.355 (24.54)	0.282 (25.17)	0.340 (25.41)
Start caregiving	0.276 (33.56)	0.327 (31.93)	0.352 (37.44)	0.264 (36.21)	0.305 (31.23)	0.285 (31.38)	0.423 (33.47)
Continue caregiving	0.364 (41.65)	0.309 (42.14)	0.355 (43.01)	0.389 (45.46)	0.361 (43.24)	0.315 (41.84)	0.335 (44.91)
Stop caregiving	0.360 (0)	0.365 (0)	0.293 (0)	0.347 (0)	0.334 (0)	0.400 (0)	0.242 (0)

Average weekly hours of care in *parentheses*

average than the latter, and in both cases carers provide informal care on average more than 20 h/week. Note also that around 50% of the women that provide care in each period have started providing care at least 1 year earlier. The observed average number of years of care, conditional on being a caregiver, is 2.4.

Table 3 shows the descriptive statistics of the variables used in the analysis, calculated separately for women who provide care for a dependent adult over the various waves of the ECHP and for women who do not. The main features characterising the carers are as follows: the labour force participation rate is 11% points lower than non-carers; carers predominantly belong to middle-aged cohorts and lower educational levels.

As a way of providing preliminary evidence on the relationship between caregiving and employment, we examine the correlation between changes in caregiving status and changes in employment status. Table 4 shows that both transitions into and out of caregiving seem to be positively correlated with non-working at t, regardless of the working status at $t - 1$. Also, remaining out of work is positively correlated with remaining in caregiving, and vice versa, that is, remaining in work is negatively correlated to remaining in caregiving.

The purpose of our exercise, as we will explain below, is to ascertain the extent to which this negative relationship between informal care and labour force participation is maintained when we control for: (1) the differences between carers and non-carers as regards observable characteristics (age, marital status, educational level, etc.), (2) the existence of unobservable fixed factors (individual heterogeneity), (3) the state dependence that tends to characterise the labour behaviour of individuals over time, and (4) the attrition problems that tend to arise when working with panel data.

4 Methods

4.1 Econometric model

The econometric model we use to estimate the impact of informal care is an ordered probit.[4] This model specifies the relationship between a latent index of linkage to the labour market, l_{it}^*, and the explanatory variables according to the following expression:

---

[4] This is not the only possibility for modelling. For example, Bingley and Walter (1997) use a multinomial probit to assess the impact of a labour force participation incentive programme for single mothers. However, the identification of the multinomial probit requires the existence of variables with different values for each alternative. The obvious case is that of the variable "labour income", which habitually requires us to predict counterfactual values, as income is only observed in the declared employment status. Another possibility would be the multinomial logit model, which is identified even when its specification only includes variables that do not vary between alternatives. However, the clearly ordered nature of the three categories of employment status justifies our choice of the ordered probit, which has also been used in previous studies of the employment status of women (e.g., Ermisch and Wright 1993). Furthermore, its simple structure facilitates the introduction of delayed employment status variables capturing state dependence in the labour supply.

Table 3 Descriptive statistics of the variables included

	Non carers		Carers	
	Mean	Standard deviation	Mean	Standard deviation
No work	0.631	0.483	0.743	0.437
Full time	0.317	0.465	0.219	0.413
Part-time	0.052	0.222	0.038	0.192
carer household	0.000	0.000	0.637	0.481
carer elsewhere	0.000	0.000	0.363	0.481
starting caregiving	0.000	0.000	0.466	0.499
continuing caregiving	0.000	0.000	0.534	0.499
stopping caregiving	0.067	0.251	0.000	0.000
care <14 h	0.000	0.000	0.172	0.378
care 14–28 h	0.000	0.000	0.295	0.456
care >28 h	0.000	0.000	0.533	0.499
d3539	0.200	0.400	0.108	0.310
d4044	0.185	0.389	0.177	0.382
d4549	0.169	0.375	0.217	0.413
d5054	0.140	0.347	0.245	0.430
d5560	0.134	0.341	0.192	0.394
single	0.078	0.269	0.092	0.290
widow	0.029	0.166	0.045	0.208
sepdiv	0.036	0.186	0.031	0.173
nch04	0.185	0.455	0.154	0.414
nch511	0.339	0.636	0.336	0.624
sizehshd	3.944	1.240	4.318	1.353
sahvgood	0.154	0.361	0.094	0.292
sahgood	0.563	0.496	0.483	0.500
sahbad	0.070	0.255	0.097	0.297
sahvbad	0.011	0.106	0.010	0.097
lincome_ot	13.573	0.941	13.584	0.664
Northwest	0.146	0.353	0.179	0.384
Northeast	0.158	0.365	0.130	0.336
Madrid	0.100	0.300	0.082	0.275
Centre	0.150	0.357	0.185	0.388
South	0.185	0.388	0.191	0.393
Canaries	0.059	0.236	0.049	0.216
isced57	0.230	0.421	0.106	0.308
isced3	0.189	0.392	0.149	0.356
Number of observations	13,472		1,775	

Source authors, based on the ECHP

Table 4 Correlation matrix between work and caregiving transitions

$t - 1/t$	No caregive/no caregive	No caregive/ caregive	Caregive/no caregive	Caregive/ caregive
No work/no work	0.2943*	0.0770*	0.0872*	0.1281*
No work/work	0.0918*	0.0052	0.0077	−0.0089
Work/no work	0.0745*	0.0240*	0.0130*	0.0021
Work/work	0.2836*	0.0067	0.0025	−0.0303*

* P value <0.05

$$l_{it}^* = \delta' C_{it} + \beta' X_{it} + \gamma' l_{it-1} + \alpha_i + \varepsilon_{it}$$
$$i = 1 \ldots N \tag{1}$$
$$t = 2 \ldots T$$

where i represents individuals and t years, C_{it} contains dummy variables denoting that woman i is engaged in caregiving in period t, X_{it} contains observable characteristics potentially associated with the decision to work, such as age, marital status, region of residence, etc., l_{it-1} contains dummy variables capturing the employment status in the previous period, α_i is an individual fixed effect denoting the effect of the unobserved systematic heterogeneity inherent to microeconomic data, and ε_{it} represents the purely random variation around the expected value of l_{it}^* (conditional on the value of the observed explanatory variables and the individual fixed effect). Furthermore, whereas α_i can be correlated with the explanatory variables and generates intra-individual autocorrelation in the composite error term $(\alpha_i + \varepsilon_{it})$, ε_{it} is independent of the explanatory variables and is not autocorrelated.

In order to model the correlation between the observed variables and the unobserved individual fixed effect, we specify a parametric relationship between the latter and the former along the lines of those proposed by Mundlak (1978) and Chamberlain (1984). That is,

$$\alpha_i = \eta' \bar{X}_i + \kappa' \bar{C}_i + \lambda' l_{i0} + u_i$$
$$i = 1 \ldots N \tag{2}$$

where $\bar{X}_i$ contains the mean of vector X_i over T time periods for individual i, $\bar{C}_i$ is the mean over time for individual i of the variables denoting the caregiving status, l_{i0} contains the values of the employment status variables in the initial period, and u_i is a random term uncorrelated to the observed explanatory variables. Equation 2 also enables us to solve the initial conditions problem that arises in dynamic models for discrete dependent variables with unobserved heterogeneity (Heckman 1981), as it incorporates the proposal made by Wooldridge (2005) which consists in conditioning α_i to the initial employment status values.

Thus, substituting (2) into (1), we get:

$$l_{it}^* = \delta' C_{it} + \beta' X_{it} + \gamma' l_{it-1} + \eta' \bar{X}_i + \kappa' \bar{C}_i + \lambda' l_{i0} + u_i + \varepsilon_{it}$$
$$i = 1\ldots N \tag{3}$$
$$t = 2\ldots T$$

where u_i is independent of the explanatory variables, but the composite error $(u_i + \varepsilon_{it})$ presents intra-individual temporal autocorrelation.

The latent variable l_{it}^* is not observed, but we do observe whether woman i in period t falls into one of the three categories "no work", "part-time" or "full-time". The rule that governs the relationship between the latent variable and the information on employment status in models of the ordered multinomial family is that as l_{it}^* exceeds certain thresholds we observe alternatives in ascending order. That is, for the three possible ordered alternatives (for $k = 1, 2, 3$), which in our case correspond to the employment statuses mentioned above, we observe $l_{it} = k$ if $\mu_{k-1} < l_{it}^* \leq \mu_k$, where $\mu_0 = -\infty$ and $\mu_m = \infty$, with m being the number of alternatives. Therefore, the basis for the maximum likelihood estimation of the model is given by the expression:

$$\mathrm{Pr}_{itk} = \mathrm{Pr}(l_{it} = k) = \mathrm{Pr}(\mu_{k-1} < l_{it}^* \leq \mu_k) \tag{4}$$

where Pr denotes probability.

There are two alternatives to consistently estimate the parameters of Eq. 3. First, under the assumptions $u_i \sim N(0, \sigma_u^2)$ and $\varepsilon_{it} \sim N(0,1)$, it is possible to integrate (throughout the distribution of u_i) the probabilities of expression (4) conditioned in realisations of u_i. The log-likelihood function would be as follows:

$$\ln L = \sum_{i=1}^{N} \ln \int_{-\infty}^{\infty} \prod_{t=2}^{T} l_{itk}(\mathrm{Pr}_{itk}|C_{it}, X_{it}, l_{it-1}, \bar{X}_i, \bar{C}_i, l_{i0}, u_i)du \quad \text{where } l_{itk} = 1 \text{ if } l_{it} = k \tag{5}$$

Expression (5) is the log-likelihood function for the random effects ordered probit model. Thus, under the assumptions of the model, the maximisation of (5) yields consistent and efficient estimates.

Alternatively, we can consider the composite error $v_i = u_i + \varepsilon_{it}$, and make the assumptions $\mathbf{v} \sim N(\mathbf{0}, \mathbf{I})$ so as to maximise the following function:

$$\ln L = \sum_{i=1}^{N} \sum_{t=2}^{T} l_{itk} \ln(\mathrm{Pr}_{itk}|C_{it}, X_{it}, l_{it-1}, \bar{X}_i, \bar{C}_i, l_{i0}) \quad \text{where } l_{itk} = 1 \text{ if } l_{it} = k. \tag{6}$$

Expression (6) is the log-likelihood function for the pooled ordered probit model. Although expression (6) is an incorrect specification of the likelihood function of the model we are using, since the assumption $\mathbf{v} \sim N(\mathbf{0}, \mathbf{I})$ ignores the existence of intra-individual correlation induced by u_i, its maximisation yields consistent but inefficient estimates of the parameters of interest. In fact, the estimate based on (6) corresponds to the estimate of the model by quasi-maximum likelihood (or partial maximum likelihood). As shown by Cameron and Trivedi (2005, p. 150), the consistency of quasi-maximum likelihood estimation does not require the correct specification of the joint density of the vector $\mathbf{l_i} = (l_{i2}, l_{i3}, \ldots, l_{iT})$ as performed in

expression (5); it is sufficient to correctly specify the marginal density of each of its elements l_{it}. It is important to note, however, that the standard error estimate based on (6) is not consistent, and therefore we use an estimator of the matrix of variances and covariances that is robust to the autocorrelation in the composite error term v_i. Obviously, the preferred way to estimate the model is the one that uses expression (5), since it yields consistent and efficient estimates. However, for reasons associated with the problem of attrition bias which we will elucidate below, we will use expression (6) for the set of final results.

4.2 Treatment of attrition

In the ECHP, and generally in all panel data sets, we encounter the problem of attrition (Peracchi 2002). Insofar as attrition is related to the variable that we are modelling, the parameters estimates—if obtained by either of the methods discussed above—will be biased. With the aim of analysing the presence of attrition bias, we perform the variable addition test proposed by Verbeek and Nijman (1992), whereby we add a dummy variable indicating whether the individual has responded in the following wave to the estimated model. The null hypothesis of no attrition bias is rejected if this variable is significant.

As we will show below, in our case attrition bias cannot be rejected. It is nevertheless possible to obtain consistent estimates using the inverse probability weighting estimator, as suggested by Wooldridge (2007). In order to implement this estimator, first we use binomial probit models to estimate the probability of individual i being present in the sample in period t, $\hat{p}_{it}$, as a function of a set of characteristics. These models are estimated for each wave of the ECHP (2–7) using the whole sample of individuals observed in the first wave. Two different specifications, yielding to alternative weighting schemes, are considered. The first, to which we shall refer as IPW-1, conditions on the first wave (1994) values of the explanatory variables to estimate a binary response model for each wave to model the probability that the individual is in this year in the sample. The second, referred to as IPW-2, conditions on the $t - 1$ values of the explanatory variables to estimate the same binary response model. In this latter case, since the sample in $t - 1$ is potentially unrepresentative of the sample in the first year of the survey, it is necessary to update the predicted probability such that $\hat{p}_{it} = \hat{\Pi}_{i2}\hat{\Pi}_{i3}, \ldots, \hat{\Pi}_{it}$, where $\hat{\Pi}_{it}$ represents the response probabilities estimated for each year (Wooldridge 2007). Lastly, we use the inverse of the predicted probabilities for each individual $(1/\hat{p}_{it})$, to weight the contributions of each observation to the log-likelihood function. In this respect, as mentioned by Contoyannis et al. (2004), the IPW estimator can be applied in situations where the objective function is additive in the contribution of each observation. This is why this estimator cannot be used in models such as the random effects ordered probit model, for which—as can be seen in expression (5)—there is a term consisting of the product of the contributions of the observations of any given individual for different time periods. This limitation does not affect the pooled ordered probit model, in which the log-likelihood function to maximise is:

$$\ln L = \sum_{i=1}^{N} \sum_{t=2}^{T} \left(\frac{R_{it}}{\hat{p}_{it}} \right) \ln(\Pr_{itk} | C_{it}, X_{it}, l_{it-1}, \bar{X}_i, \bar{C}_i, l_{i0}) \tag{7}$$

where R_{it} is a dummy variable which takes the value 1 if individual i is present in the sample for period t and 0 otherwise.

5 Results

We consider three different models to assess the impact of caregiving on employment status. In all models the dependent variable is the ordered categorical variable $l_{it} = 1, 2, 3$, corresponding to whether the woman declares herself to be in the "no work", "part-time" or "full-time" status, respectively. The models differ, however, in the specification of caregiving. In Model 1 we use three categories which are intended to capture whether the place in which the care is given is relevant: "caregiving at home", "caregiving elsewhere", and "non-caregiving". In Model 2 we use four categories that are intended to capture whether the moment at which the transition to (or from) caregiving occurs is important: "start caregiving" (did not provided care in $t - 1$ but does so in t), "continue caregiving" (provided care in $t - 1$ and also does so in t), "stop caregiving" (provided care in $t - 1$ but does not do so in t) and "continue not caregiving" (did not give care in $t - 1$ and does not do so in t). In Model 3 we consider the number of weekly hours of care provision distinguishing four possible categories: less "than 14 h", "between 14 and 28 h", "more than 28 h" and "0 h". For all models, the different categories of caregiving are parameterised with dummy variables. The omitted categories are "non-caregiving", "continue not caregiving" and "zero hours" in Models 1, 2 and 3, respectively. As mentioned earlier, all models include a broad set of control variables: age, educational level, marital status, etc. (see Table 1).

5.1 Attrition test

The results of the tests for attrition bias using a dummy variable that captures whether the individual is in the sample in the following year as proposed by Verbeek and Nijman (1992), in which the null hypothesis is no bias, are shown in Table 5 for both the random effects specifications and the pooled specifications. The null hypothesis of no bias is rejected for all the random effects models and the pooled Model number 2. With p-values near 0.20, it cannot be rejected at conventional significance levels for pooled models 1 and 3.

5.2 Model estimates

The rejection of the null of no attrition bias in four out of the six models considered suggests that it is necessary to use the IPW estimator. For the reasons discussed in Sect. 4.2, it can only be applied in the case of the pooled ordered probit model; hence the results we present below correspond to this specification. As we

Table 5 Attrition test results

	Ordered probit (pooled)			Ordered probit (random effects)		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Following year	1.66 (0.1980)	2.97 (0.0849)	1.81 (0.1789)	4.09 (0.0431)	3.96 (0.0467)	3.97 (0.0462)

Values of χ^2 and probabilities (*in parentheses*)

Model 1 refers to the specification that distinguishes between caregiving at home and caregiving elsewhere

Model 2 refers to the specification that distinguishes between starting, continuing and stopping caregiving

Model 3 refers to the specification that distinguishes the number of hours of caregiving

mentioned earlier, we estimate the labour supply models using two alternative weighting schemes and, for the sake of comparison, we also report estimates using no weights. Table 6 presents the estimates for the nine resulting alternative specifications.

The first three columns of coefficients contain the results for the models that distinguish between caregiving at home and caregiving elsewhere. The three columns in the middle show the results for the models that consider the dynamics of the care provision. And the three columns on the right side of the table correspond to the estimates for the models that consider the number of weekly hours of care provision.

Although the scale of the ordered probit is arbitrary, the estimates in Table 6 enable us to know the direction of the effect of the different explanatory variables.[5] The first outstanding feature is the importance of state dependence, as women who work in the previous period, whether full-time or part-time, have a higher probability of working in the following period. Regarding the rest of the explanatory variables, note the positive effect on the probability of working of being single, having a higher educational level and having very good health status. Major regional differences are also found: women living in the autonomous communities of the centre and south have a lower probability of working than those living in the rest of Spain. Finally, we should mention the robustness of the results to the use of the different types of weights (IPW-1 vs. IPW-2) for all the specifications considered.

5.3 Average effects

As we mentioned earlier, the scale of the ordered probit model is arbitrary. In order to obtain an indicator of the magnitude of the relationship between the various caregiving conditions and employment status, we have calculated the average effect

---

[5] The reference category is a woman aged 30–34, married, with fair health status, with education attainment lower than secondary education, and lives in the East region.

Table 6 Dynamic ordered probit models for employment status

Variables	Caregiving at home or elsewhere			Dynamics of caregiving			Number of hours of caregiving		
	IPW-1 (1994 values)	IPW-2 (t − 1 values)	No IPW	IPW-1 (1994 values)	IPW-2 (t − 1 values)	No IPW	IPW-1 (1994 values)	IPW-2 (t − 1 values)	No IPW
Full time (1994)	0.845 (0.056)*	0.889 (0.055)*	0.850 (0.047)*	0.849 (0.056)*	0.888 (0.054)*	0.853 (0.047)*	0.851 (0.056)*	0.885 (0.054)*	0.850 (0.047)*
Part time (1994)	0.350 (0.069)*	0.390 (0.060)*	0.343 (0.053)*	0.353 (0.070)*	0.388 (0.060)*	0.344 (0.054)*	0.339 (0.069)*	0.382 (0.061)*	0.336 (0.054)*
Full time (t − 1)	2.045 (0.058)*	2.076 (0.060)*	2.035 (0.050)*	2.039 (0.058)*	2.076 (0.060)*	2.029 (0.050)*	2.044 (0.058)*	2.077 (0.060)*	2.035 (0.050)*
Part time (t − 1)	1.303 (0.067)*	1.234 (0.064)*	1.223 (0.053)*	1.304 (0.068)*	1.243 (0.064)*	1.224 (0.054)*	1.307 (0.067)*	1.233 (0.063)*	1.225 (0.053)*
Carer_household	−0.233 (0.091)*	−0.158 (0.099)	−0.124 (0.088)						
Carer_elsewhere	−0.204 (0.190)	0.122 (0.113)	0.064 (0.089)						
Starting caregiving				−0.060 (0.117)	0.022 (0.089)	0.024 (0.078)			
Continuing caregiving				−0.208 (0.112)**	−0.147 (0.108)	−0.113 (0.098)			
Stopping caregiving				0.170 (0.101)**	−0.030 (0.082)	0.012 (0.072)			
Caregive <14							−0.248 (0.263)	0.144 (0.123)	0.073 (0.105)
Caregive 14–28							0.005 (0.118)	0.146 (0.112)	0.124 (0.093)
Caregive >28							−0.381 (0.101)*	−0.280 (0.102)*	−0.233 (0.089)*
D3539	−0.070 (0.094)	0.000 (0.082)	0.014 (0.070)	−0.069 (0.095)	0.004 (0.081)	0.020 (0.070)	−0.074 (0.094)	−0.002 (0.082)	0.014 (0.070)
d4044	−0.021 (0.138)	0.172 (0.128)	0.150 (0.110)	−0.022 (0.139)	0.175 (0.127)	0.154 (0.110)	−0.026 (0.139)	0.170 (0.129)	0.147 (0.111)
d4549	0.038 (0.193)	0.221 (0.174)	0.192 (0.150)	0.030 (0.195)	0.224 (0.173)	0.188 (0.149)	0.027 (0.194)	0.216 (0.175)	0.185 (0.150)
d5054	−0.016 (0.235)	0.175 (0.215)	0.089 (0.183)	−0.030 (0.237)	0.170 (0.214)	0.077 (0.183)	−0.031 (0.238)	0.171 (0.215)	0.082 (0.183)
d5560	−0.352 (0.292)	0.011 (0.253)	−0.108 (0.218)	−0.370 (0.295)	0.000 (0.252)	−0.120 (0.218)	−0.381 (0.296)	−0.003 (0.254)	−0.123 (0.220)
Single	0.790 (0.348)*	0.720 (0.359)*	0.600 (0.246)*	0.819 (0.351)*	0.712 (0.359)*	0.620 (0.245)*	0.782 (0.345)*	0.723 (0.359)*	0.601 (0.246)*
Widow	−0.371 (0.383)	−0.812 (0.493)**	−0.394 (0.381)	−0.359 (0.381)	−0.806 (0.488)	−0.369 (0.376)	−0.342 (0.385)	−0.834 (0.491)**	−0.394 (0.385)
Sepdiv	−0.114 (0.165)	0.075 (0.203)	−0.082 (0.154)	−0.109 (0.164)	0.069 (0.200)	−0.082 (0.153)	−0.107 (0.166)	0.074 (0.203)	−0.076 (0.155)
nch04	0.093 (0.060)	0.037 (0.049)	0.075 (0.044)**	0.094 (0.061)	0.046 (0.048)	0.079 (0.044)**	0.099 (0.060)	0.043 (0.049)	0.080 (0.044)
nch511	−0.061 (0.042)	−0.050 (0.043)	−0.049 (0.035)	−0.056 (0.042)	−0.041 (0.043)	−0.042 (0.035)	−0.058 (0.043)	−0.049 (0.043)	−0.047 (0.035)
sizehshd	−0.064 (0.040)	−0.011 (0.042)	−0.025 (0.036)	−0.064 (0.040)	−0.013 (0.042)	−0.027 (0.036)	−0.061 (0.040)	−0.011 (0.042)	−0.024 (0.036)
sahvygood	0.166 (0.077)*	0.143 (0.070)*	0.125 (0.058)*	0.159 (0.077)*	0.139 (0.070)*	0.118 (0.058)*	0.169 (0.077)*	0.140 (0.071)*	0.123 (0.058)*
sahgood	0.104 (0.060)**	−0.010 (0.050)	0.021 (0.044)	0.099 (0.060)	−0.011 (0.050)	0.016 (0.044)	0.107 (0.060)**	−0.009 (0.051)	0.022 (0.044)

Table 6 continued

Variables	Caregiving at home or elsewhere			Dynamics of caregiving			Number of hours of caregiving		
	IPW-1 (1994 values)	IPW-2 (t − 1 values)	No IPW	IPW-1 (1994 values)	IPW-2 (t − 1 values)	No IPW	IPW-1 (1994 values)	IPW-2 (t − 1 values)	No IPW
sahbad	0.232 (0.114)*	−0.046 (0.087)	−0.019 (0.080)	0.219 (0.112)**	−0.047 (0.086)	−0.026 (0.080)	0.241 (0.115)*	−0.042 (0.087)	−0.021 (0.080)
sahvbad	−0.116 (0.298)	−0.121 (0.273)	−0.051 (0.209)	−0.130 (0.298)	−0.129 (0.272)	−0.056 (0.208)	−0.101 (0.302)	−0.118 (0.276)	−0.049 (0.211)
lincome_ot	0.004 (0.027)	0.025 (0.031)	0.006 (0.025)	0.004 (0.027)	0.025 (0.031)	0.007 (0.025)	0.005 (0.027)	0.025 (0.031)	0.006 (0.025)
Northwest	0.138 (0.067)*	0.028 (0.062)	0.099 (0.053)**	0.138 (0.068)	0.030 (0.062)	0.092 (0.053)**	0.136 (0.067)*	0.029 (0.062)	0.100 (0.053)**
Northeast	−0.019 (0.065)	−0.091 (0.058)	−0.057 (0.049)	−0.021 (0.065)	−0.095 (0.058)	−0.058 (0.049)	−0.025 (0.065)	−0.097 (0.058)**	−0.060 (0.049)
Madrid	−0.059 (0.082)	−0.085 (0.068)	−0.086 (0.059)	−0.061 (0.083)	−0.089 (0.068)	−0.088 (0.059)	−0.069 (0.082)	−0.093 (0.067)	−0.091 (0.059)
Centre	−0.192 (0.067)*	−0.187 (0.063)*	−0.196 (0.054)	−0.195 (0.066)*	−0.191 (0.063)*	−0.198 (0.053)*	−0.202 (0.064)*	−0.191 (0.063)*	−0.199 (0.054)*
South	−0.175 (0.071)*	−0.231 (0.062)*	−0.211 (0.052)*	−0.175 (0.071)*	−0.229 (0.062)*	−0.213 (0.052)*	−0.174 (0.069)*	−0.235 (0.062)*	−0.211 (0.052)*
Canaries	0.004 (0.087)	−0.035 (0.081)	−0.004 (0.073)	0.003 (0.087)	−0.040 (0.081)	−0.005 (0.073)	0.004 (0.087)	−0.034 (0.081)	0.000 (0.073)
isced57	0.411 (0.053)*	0.354 (0.051)*	0.364 (0.044)*	0.413 (0.053)*	0.357 (0.051)*	0.371 (0.044)*	0.404 (0.054)*	0.348 (0.051)*	0.358 (0.044)*
isced3	0.168 (0.056)*	0.152 (0.047)*	0.180 (0.041)*	0.168 (0.056)*	0.149 (0.047)*	0.180 (0.041)*	0.157 (0.056)*	0.147 (0.047)*	0.174 (0.041)*
cut1	−0.141 (0.495)	−0.176 (0.461)	−0.316 (0.365)	−0.131 (0.501)	−0.169 (0.459)	−0.326 (0.364)	−0.129 (0.483)	−0.179 (0.460)	−0.303 (0.363)
cut2	0.194 (0.493)	0.186 (0.460)	0.025 (0.364)	0.203 (0.500)	0.191 (0.458)	0.016 (0.363)	0.207 (0.482)	0.184 (0.459)	0.039 (0.363)
N observations	15,235	12,419	15,235	15,247	12,544	15,247	15,242	12,446	15,242
N individuals	3,858	2,957	3,858	3,859	2,966	3,859	3,859	2,957	3,859
Log-likelihood	−60,456	−4,780	−6,117	−6,054	−4,815	−6,125	−6,044	−4,767	−6,119

Coefficients (Standard Errors)

The standard errors are robust to autocorrelation in the composite error term

* P value <0.05

** P value <0.10

on the subsample of carers.[6] This measures the average effect on the probability of each employment status (no-employment, part-time or full-time work) of entering into a particular caregiving category (i.e. caring at home, caring elsewhere, starting to care, caring less than 14 h/week, …), for those women observed to do so in the sample (this implies that the effect is computed for women who belong to the caregiver category at some point throughout the sampling period). We estimate the effect of interest on the subsample of caregivers because they constitute a group of utmost relevance, as these are the women who face potential opportunity costs in terms of foregone labour market opportunities associated to providing care. The change in labour outcomes for this group, from/to a counterfactual situation where they do not supply care, seems more policy relevant than the corresponding change for the overall population.

Table 7 shows the average effects for each of the models estimated, with standard errors obtained by bootstrapping (500 replications).

The results show, firstly, that the absolute effects of informal care on employment status are mainly restricted to the decision between working full-time and not working, since the estimated effects on the probability of working part-time are in all cases lower than 0.7% points. However, as we show below the relative effects on working part-time are non-negligible. Secondly, we find that caring for someone at home reduces the probability of working full-time by 2.7% points, yet caregiving elsewhere does not appear to have any effect. Similar qualitative results have been reported in the related literature (Ettner 1995; Heitmueller and Michaud 2006; Heitmueller 2007), although the size of our estimates is smaller than the estimates of Heitmueller (2007) and Heitmueller and Michaud (2006), who analyse British women within similar age ranges. Heitmueller (2007) in fact estimates that providing informal care for a co-resident decreases the probability of employment in the overall population by 15%, while Heitmueller and Michaud (2006) obtain estimates around 5.9% in a model that controls for state dependence. The differences with respect to these studies suggest that the labour opportunity costs of providing informal care are smaller for Spanish caregivers. This explanation is not inconsistent with the lower rate of female participation found in Spain.

Concerning the moment at which the possible change in employment status occurs, we find that the probability of working full-time does not diminish significantly on the first year. Rather, it decreases on subsequent years. The size of this decrease is similar to the effect found for caregiving at home. This is not surprising, in view of the number of hours dedicated to caregiving in each case. Stopping care is found to exert a significant effect (at the 10% level, and only in the IPW-1 specification) on the chances of leaving inactivity and entering either part-time or full time employment.

The magnitude of the effects from the three representations of informal care points towards the number of hours of informal care provision as the crucial factor

---

[6] Average effect$_{ks} = E\left[\Phi(\hat{\mu}_k - P_{it}\hat{\beta} - C_{it}^s\hat{\delta}) - \Phi(\hat{\mu}_{k-1} - P_{it}\hat{\beta} - C_{it}^s\hat{\delta}) - \Phi(\hat{\mu}_k - P_{it}\hat{\beta}) + \Phi(\hat{\mu}_{k-1} - P_{it}\hat{\beta})\right|$ $C_{it}^s = 1]$, where the vector P contains all the variables included in the model, except for the caregiving dummies, and C_{it}^s is a dummy variable that equals one when the woman is in the caregiving status of interest.

Table 7 Average effects on the subsample of carers

	Prob (no work)			Prob (part-time)			Prob (full-time)		
	IPW-1	IPW-2	No IPW	IPW-1	IPW-2	No IPW	IPW-1	IPW-2	No IPW
At home	0.031 (0.011)*	0.020 (0.012)**	0.017 (0.011)	−0.004 (0.002)*	−0.003 (0.002)**	−0.002 (0.002)	−0.027 (0.010)*	−0.017 (0.011)**	−0.014 (0.010)
Elsewhere	0.026 (0.024)	−0.016 (0.015)	−0.009 (0.012)	−0.004 (0.003)	0.003 (0.003)	0.001 (0.002)	−0.022 (0.021)	0.013 (0.013)	0.007 (0.010)
Start caregiving	0.008 (0.015)	−0.003 (0.012)	−0.003 (0.011)	−0.001 (0.001)	0.000 (0.002)	0.000 (0.001)	−0.007 (0.014)	0.003 (0.011)	0.003 (0.010)
Continue caregiving	0.026 (0.014)**	0.018 (0.014)	0.014 (0.013)	−0.004 (0.003)**	−0.004 (0.003)	−0.003 (0.002)	−0.022 (0.012)**	−0.015 (0.011)	−0.012 (0.011)
Stop caregiving	−0.025 (0.014)**	0.004 (0.011)	−0.002 (0.010)	0.003 (0.002)**	−0.001 (0.002)	0.000 (0.001)	0.022 (0.012)**	−0.004 (0.009)	0.001 (0.009)
Caregive <14 h	0.033 (0.037)	−0.020 (0.018)	−0.010 (0.016)	−0.001 (0.002)	0.002 (0.002)	0.001 (0.001)	−0.032 (0.035)	0.018 (0.016)	0.009 (0.014)
Caregive 14–28 h	−0.001 (0.017)	−0.022 (0.017)	−0.019 (0.014)	0.000 (0.003)	0.004 (0.003)	0.003 (0.002)	0.001 (0.014)	0.018 (0.014)	0.016 (0.012)
Caregive >28 h	0.045 (0.013)*	0.032 (0.012)*	0.028 (0.012)*	−0.007 (0.002)*	−0.006 (0.002)*	−0.005 (0.002)*	−0.038 (0.011)*	−0.026 (0.010)*	−0.023 (0.010)*

Bootstrapped standard errors in *parenthesis*

* P value <0.05

** P value <0.10

**Table 8** Correlation between the different caregiving variables

	At home	Elsewhere	Start caregiving	Continue caregiving	Stop caregiving
Elsewhere	−0.057				
Start caregiving	0.427	0.510			
Continue caregiving	0.639	0.293	−0.059		
Stop caregiving	−0.067	−0.050	−0.058	−0.061	
Caregive <14 h	0.142	0.422	0.354	0.174	−0.033
Caregive 14–28 h	0.329	0.396	0.378	0.334	−0.045
Caregive >28 h	0.702	0.240	0.402	0.584	−0.062

All the correlations are significantly different from zero at 95%

affecting labour supply. In particular note, at the bottom of Table 7, the differing effects of caregiving less than 14 h, between 14 and 28 h and more than 28 h. These results suggest the absence of labour opportunity costs when women provide less than 28 h/week of informal care. However, when the 28 h threshold is surpassed, the probability of not working rises by as much as 4.5% points, all else held equal, according to one of our specifications. Moreover, women do not seem to transit from full-time employment to part-time employment, rather, the estimated effects suggest that women transit from employment (either full-time or part-time) to non-employment.

Considering the rest of the reported estimated average effects, a picture emerges suggesting that the labour opportunity costs of informal care affect co-resident carers, those who provide care for long periods, and those who provide care for more than 28 h/week. However, these three variables are highly correlated, as shown in Table 8, where it can also be seen that individuals who continue caregiving mostly do so at home. Table 8 also suggests that women who provide care for more than 28 h/week, do so at home and are either in the second or subsequent years of the caregiving episode. This suggests that these three variables could be proxies for the cared for subject's level of dependency. Specifically, evidence exists for other countries to the effect that older people move in with their adult children when dependency problems prevent them from living in their own homes (Pezzin and Schone 1999), so co-residence of carer and dependent would be a proxy for serious dependency when the former is a middle-aged woman. In addition, since in most cases dependency problems get worse as time passes (because they tend to stem from chronic processes of a degenerative nature such as Alzheimer's, cancer, etc.), continuity of care might also proxy the degree of dependency.

The figures in Table 7 could suggest at a first glimpse that the labour effects of caregiving are relatively small. However, when we compare the relative change in the probability of being in full-time employment or in part-time employment for women who caregive more than 28 h or caregive at home with the probability of being in part-time or full-time employment in a counterfactual situation where they do not provide care, the relative decrease in the probability of employment due to caregiving is 17.5% for part-time employment and 20.5% for full-time employment
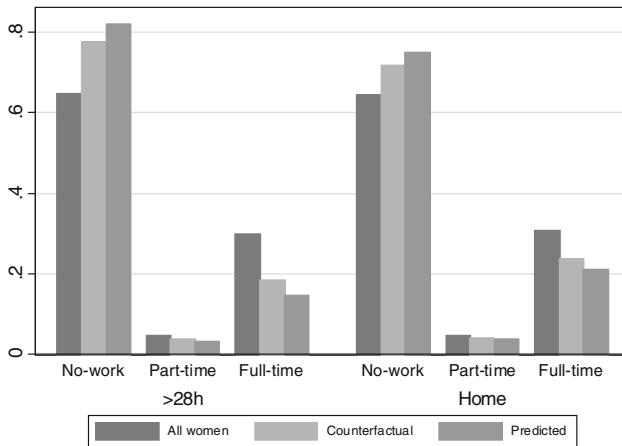
Fig. 3 Magnitude of the effects on employment for women who caregive more than 28 h/week and women who caregive at home. *Note* "All women" refers to the mean predicted probability of each employment status for all women aged 30–60 included in the sample of analysis; "Counterfactual" refers to the mean probability of each employment status for caregivers under the counterfactual scenario of no care provision; "Predicted" refers to the predicted probability of each employment for caregivers under the observed scenario of care provision (given the non linearity of the underlying model, these "predicted" probabilities differ slightly from the actual observed probabilities)

for those women who provide more than 28 h of care. The corresponding decreases among women who provide care at home are 9.1 and 11%. These estimates are shown in Fig. 3, where we can firstly note that the predicted probability of no employment for women who provide care at home or provide care for more than 28 h/week is greater than the corresponding predicted probability in the population of women aged 30–60. Indeed, for women who caregive more than 28 h/week (caregive at home) the counterfactual probability—i.e., the probability that would ensue if they were not caregivers—of working part-time is 0.040 (0.039) and that of working full-time is 0.185 (0.239), while their actual probabilities (in fact, the predicted probabilities estimated by the model under the observed scenario) are 0.033 (0.397) and 0.147 (0.212), respectively. As we mentioned earlier, these figures imply that providing informal care decreases the probability of being in part-time employment by 17.5% for women providing more than 28 h of care per week, and by 9.1% for women caregiving at home. Similarly, the probability of being in full-time employment is reduced by 20.5 and 11%, respectively.

5.4 Sensitivity analysis

In this section we carry out a test on the adequacy of our modelling assumption regarding the conditional exogeneity of the caregiver status. For this purpose we consider a model of two simultaneous equations with dynamics: a labour participation equation in which caregiving is one of the explanatory variables and a caregiving equation, allowing for dependence among their stochastic components.

Each individual i at each moment t decides whether he is going to provide care and participate in the labour market. Formally, we specify the following recursive bivariate dynamic model:

$$
\begin{aligned}
l_{it}^* &= \beta_l' X_{it} + \delta_l' C_{it} + \gamma_l l_{it-1} + v_{it}^l \\
C_{it}^* &= \beta_c' Z_{it} + \gamma_c' C_{it-1} + \delta_c' l_{it-1} + v_{it}^c \\
C_{it} &= I(C_{it}^* > 0), l_{it} = I(l_{it}^* > 0) \\
i &= 1, \ldots, N; \quad t = 1, \ldots, T
\end{aligned}
\tag{8}
$$

where C_{it} represents the decision to care and l_{it} the employment decision.

We specify composite error terms in the two equations, where u represents an individual fixed effect, possibly correlated with the explanatory variables, and ε is white noise.

$$
\begin{aligned}
v_{it}^l &= u_i^l + \varepsilon_{it}^l \\
v_{it}^c &= u_i^c + \varepsilon_{it}^c
\end{aligned}
\tag{9}
$$

In parallel to our strategy in the previous section, we model the individual fixed effect as suggested by Mundlak (1978); Chamberlain (1984) and Wooldridge (2005), i.e., we have modelled the individual fixed effects as shown in Eq. 10.

$$
\begin{aligned}
u_i^l &= \lambda_l' l_{io} + \phi_l' C_{io} + \eta_l' \overline{X}_i + \kappa_l' \overline{C}_i + \xi_i^l \\
u_i^c &= \lambda_c' l_{io} + \phi_c' C_{io} + \eta_c' \overline{Z}_i + \kappa_c' \overline{C}_i + \xi_i^c
\end{aligned}
\tag{10}
$$

where the random terms on the right hand side are white noise. We assume that the stochastic terms are jointly distributed following a bivariate normal distribution and therefore our model is a bivariate probit.

Note that this model explicitly specifies caregiving status as an endogenous variable. Note also that if the error terms in Eq. 8 are independent then each equation can be consistently estimated by a univariate probit (Maddala 1983). Our contention is that the modelling of the individual fixed effect according to Eq. 10 suffices to account for the endogeneity of caregiving status in the labour participation equation. If this is the case, we should find that, while a model that imposes $u_i^l = \xi_i^l$ and $u_i^c = \xi_i^c$ might still show dependence between v_{it}^l and v_{it}^c, lifting this restriction would drive this correlation to zero (ρ_{cl}). In fact, Maddala (1983) and Knapp and Seaks (1998) propose that a test of H_0: $\rho_{cl} = 0$ using a z-test and an LR test can act as an alternative to the Hausman test for exogeneity of the caregiving dummy variable in the labour participation equation.

Even if our model in the previous section is an ordered probit, the results for this test on the current model, where only the observability rule for the dependent variable changes with respect to the former model, are able to shed some light on the adequacy of our strategy. In particular, we find that the Mundlak representation of unobserved fixed heterogeneity renders the two error terms in Eq. 8 independent, so we are able to treat the caregiving dummy variable as exogenous in the labour equation. This lends support to our choice for the estimation of Eq. 3 in the previous section as a univariate ordered probit.

Table 9 Results for the test of exogeneity of the caregiving variable

	Maximum likelihood	
	ρ (SE)	Wald-test (prob $> \chi^2$)
Baseline	0.0770 (0.0607)	1.5974 (0.2063)
Baseline + Mundlak	−0.0490 (0.0879)	0.3092 (0.5782)

In order to carry out this sensitivity test, we will include in the employment equation the same covariates as in the previous models. Furthermore, we will consider a broad definition of caregiving (either at home and/or elsewhere and any number of hours per week). To achieve the non-parametric identification of the model, we include a dummy variable that captures whether there was an individual older than 65 in the household in the previous period in the caregiving equation. This exclusion restriction is grounded on the idea that the presence of individuals aged 65+ among the rest of the members of the household will affect the chances of labour participation only via the potential need for caregiving.[7]

We estimate the bivariate probit model shown in (8) by maximum likelihood. Table 9 shows the results of the endogeneity test, whereas the full set of results are shown in Table 10 in Appendix. We first estimate the bivariate probit assuming that η_l, η_c, κ_1 and κ_c are equal to zero, and thus that u_i^l and u_i^c are uncorrelated with the covariates. Secondly, we introduce the parameterisation of the unobserved fixed effect. Concerning our choice of exclusion restriction, it should be noted that our instrument (shorthanded to "older 65 $(t-1)$" in Table 10 in the Appendix) has a positive and significant effect on the probability of caregiving in both specifications. Thus, our instrument satisfies the requisites of excludability and relevance. The results regarding the endogeneity test itself in Table 9 suggest that the null hypothesis of no correlation between the random components of the error terms cannot be rejected for either model. This provides evidence regarding the consistency of the estimates obtained in the previous analysis.

These results are in line with those of Bolin et al. (2008) and Heitmueller (2007), who are unable to reject the exogeneity assumption.

6 Discussion and conclusions

Our analysis suggests that the labour effects of informal care affect mostly women who care for someone at home, and/or provide care for more than one period, and or provide more than 28 h/week of care. Therefore, unlike Viitanen (2005), we do detect labour supply opportunity costs associated to informal care. This underscores

---

[7] In order to check whether our instrument can be excluded from the labour equation, we have re-estimated the labour equation including all the explanatory variables plus the dummy variable that captures whether there was an individual older than 65 in the household lagged one period. We are not able to reject the null of no significance at conventional levels and we interpret this as evidence in favour of the orthogonality of our instrument to the error term in the labour equation. Results not shown but available from the authors upon request.

the importance of considering all potentially different types of informal care separately when analysing its effects on labour outcomes.

We also find that those who finish an episode of informal care do not appear to have problems re-entering the labour market. The transitions by women who provide informal care tend to be dominated by changes in the extensive margin (labour force participation), rather than the intensive margin (hours worked), of employment (Heckman 1993). Similar results have been found in other research on informal care (Ettner 1996), but these phenomena are very probably more acute in the Spanish case owing to the relative scarcity of part-time contracts in Spain (European Commission 2004).

Our results have implications for the design of public policies on long-term care. The labour effects appear to be concentrated in intensive carers (more than 28 h/week), co-resident carers and those who provide care for long periods. As we have argued, the high correlation among these conditions suggest that they all proxy the cared for subject's level of dependency. This also conforms with existing evidence for other countries showing that older people move in with their adult children when dependency problems prevent them from continuing in their own homes (Pezzin and Schone 1999). In consequence, the new SAAD benefits should be modulated according to the level of dependency.

Our results also prompt caveats about the likely effects of the new Long-Term Care System in Spain. On one hand, the gradual implementation of the system (individuals in worse health are being covered first) would in theory allow the effects on women's participation to diminish in the short run. However, the design of the different in-kind and monetary benefits suggests that some confronting effects could ensue. The provision of up to 20 weekly hours of formal care and the provision of subsidised day-care centres and residences should work in the direction of favouring women's labour participation. In contrast, the availability of monetary transfers (up to €500/month if severely disabled or €320/month if moderately disabled) for women acting as the main informal carer will decrease the opportunity costs of leaving the labour market. The net effect on the labour supply of co-resident carers is difficult to predict, but most likely it will depend on each woman's labour opportunity costs. Thus, women with low earnings will probably be induced to drop out of the labour force, while high earners are likely to be encouraged to remain at work.

Previous evidence suggests that working at firms that offer unpaid family leave exerts a positive influence on the chances of employment among caregivers (Pavalko and Henderson 2006). Accordingly, as the estimated effects of caregiving that we have obtained seem to be reduced to an all-or-nothing choice (work full-time or stop working), reforms aiming to diminish the opportunity costs of informal care should include flexible employment formulas, such as the reduction in working hours that already exists for maternity, duly incentivised economically so as to avoid perverse behaviour by employers.

A natural extension of our work, along the lines of that proposed by Viitanen (2005), would be to replicate the analysis for all the European countries for which the ECHP has data. Considering the great diversity that exists at European level as regards the flexibility of working hours (European Commission 2004) and the

coverage provided by long-term care systems (OECD 2005), this approach would reveal whether the results obtained for Spain are also forthcoming on examining other countries in the same region, and at the same time provide information as to what institutional factors lead to the largest reduction in the labour opportunity costs associated with informal care.

Appendix

See Table 10.

Table 10 Bivariate probit estimates

| | Excluding Mundlak specification of the unobserved heterogeneity | | | | Including Mundlak specification of the unobserved heterogeneity | | | |
| | Work | | Caregiving | | Work | | Caregiving | |
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE
Informal carer	−0.213*	0.104			0.069	0.178		
Informal carer ($t-1$)			1.296*	0.044			1.080*	0.049
Full time ($t-1$)	2.424*	0.042	−0.087*	0.040	1.902*	0.049	−0.187*	0.058
Part time ($t-1$)	1.975*	0.065	−0.166*	0.083	1.670*	0.071	−0.197*	0.091
Informal carer (1994)							0.460*	0.049
Full time (1994)					0.781*	0.050	0.142*	0.061
Part time (1994)					0.559*	0.073	0.035	0.089
Older 65 ($t-1$)			0.660*	0.041			0.390**	0.208
Older 65 (1994)							0.246	0.210
d3539	0.006	0.045	0.156*	0.064	0.039	0.073	−0.189**	0.098
d4044	0.020	0.048	0.371*	0.066	0.145	0.117	−0.156	0.143
d4549	−0.019	0.049	0.395*	0.065	0.182	0.157	−0.273	0.177
d5054	−0.118*	0.056	0.448*	0.068	0.091	0.190	−0.275	0.213
d5560	−0.325*	0.057	0.230*	0.073	−0.087	0.228	−0.327	0.246
Single	0.241*	0.052	−0.089	0.071	0.725*	0.272	−0.639**	0.343
Widow	0.085	0.092	0.100	0.082	−0.518	0.339	−0.307	0.271
sepdiv	0.117	0.075	−0.008	0.086	−0.036	0.173	−0.104	0.209
nch04	0.052**	0.032	−0.059	0.039	0.076**	0.046	−0.043	0.054
nch511	−0.028	0.022	0.012	0.025	−0.061**	0.037	0.014	0.039

Table 10 continued

	Excluding Mundlak specification of the unobserved heterogeneity				Including Mundlak specification of the unobserved heterogeneity			
	Work		Caregiving		Work		Caregiving	
	Coef.	SE	Coef.	SE	Coef.	SE	Coef.	SE
sizehshd	−0.015	0.013	0.043*	0.012	−0.039	0.039	0.282*	0.043
sahvgood	0.101*	0.049	−0.205*	0.057	0.084	0.061	−0.176*	0.070
sahgood	0.016	0.036	−0.165*	0.038	−0.020	0.045	−0.122*	0.048
sahbad	−0.221*	0.062	−0.130*	0.064	−0.042	0.079	−0.025	0.078
sahvbad	−0.364*	0.181	−0.452*	0.150	−0.057	0.212	−0.270	0.169
lincome_ot	−0.064*	0.018	0.017	0.019	−0.001	0.027	0.010	0.026
Northwest	0.036	0.049	−0.051	0.059	0.053	0.057	−0.017	0.063
Northeast	−0.065	0.049	−0.061	0.058	−0.065	0.055	−0.034	0.061
Madrid	−0.116*	0.057	0.002	0.066	−0.111**	0.066	0.014	0.071
Centre	−0.185*	0.053	0.107**	0.057	−0.176*	0.061	0.126*	0.061
South	−0.243*	0.050	0.059	0.056	−0.240*	0.057	0.088	0.060
Canaries	−0.115**	0.067	−0.072	0.083	−0.064	0.076	−0.061	0.088
isced57	0.434*	0.041	−0.146*	0.053	0.407*	0.048	−0.162*	0.059
isced3	0.113*	0.039	−0.037	0.046	0.114*	0.045	−0.049	0.050
dwave2	−0.202*	0.047	0.095**	0.055	−0.210*	0.063	−0.074	0.074
dwave3	−0.279*	0.049	0.109**	0.056	−0.278*	0.060	−0.033	0.068
dwave4	−0.117*	0.048	0.212*	0.055	−0.139*	0.053	0.103**	0.062
dwave5	−0.061	0.048	0.085	0.055	−0.084**	0.050	0.018	0.059
dwave6	−0.044	0.051	0.105**	0.061	−0.058	0.049	0.068	0.061
_cons	−0.193	0.266	−2.358*	0.290	0.795*	0.402	−2.532*	0.428
N	15,202				15,199			
Likelihood	−8721.5				−8428.9			

* P value <0.05

** P value <0.10

References

Ahn N, Garcia J, Herce J (2005) Health care expenditure and demographic uncertainty. Documento de Trabajo 2005–2007. FEDEA, Madrid

Bingley P, Walter I (1997) The labour supply, unemployment and participation of lone mothers in in-work transfer programmes. Econ J 107:1375–1390

Bolin K, Lindgren B, Lundborg P (2008) Your next of kin or your own career? Caring and working among the 50+ of Europe. J Health Econ 21:718–738

Cameron AC, Trivedi PK (2005) Microeconometrics, methods and applications. Cambridge University Press, New York

Carmichael F, Charles S (1998) The labour market costs of community care. J Health Econ 17:747–765

Carmichael F, Charles S (2003) The opportunity costs of informal care: does gender matter? J Health Econ 22:781–803

Casado D (2006) La atención a la dependencia en España. Gac Sanit 20(S1):135–142

Chamberlain G (1984) Panel data. In: Griliches Z, Intriligator MD (eds) Handbook of econometrics, vol 1. Amsterdam, North Holland, pp 1247–1318

Contoyannis P, Jones AM, Rice N (2004) The dynamics of health in the British Household Panel Survey. J Appl Econ 19:473–503

Crespo L (2007) Caring for parents and employment status of European mid-life women. Documento de trabajo CEMFI no. 6015, Madrid

Ermisch JF, Wright RE (1993) Wage offers and full-time and part-time employment by British women. J Hum Resour 28:111–133

Ettner SL (1995) The impact of parent care on female labor supply decisions. Demography 32:63–79

Ettner SL (1996) The opportunity costs of elder care. J Hum Resour 31:189–205

European Commission (2004) Employment in Europe 2004: recent trends and prospects. Employment and Social Affairs, European Commission, Brussels

Fundació Institut Català de l'Envelliment (2004) Estudio del modelo de atención a las personas mayores con dependencia en España. Edad & Vida, Barcelona

Heckman JJ (1981) The incidental parameters problem and the problem of initial conditions in estimating a discrete time—discrete data stochastic process. In: Manski CF, McFadden D (eds) Structural analysis of discrete data with econometric applications. MIT Press, Cambridge

Heckman JJ (1993) What has been learned about labor supply in the past twenty years? Am Econ Rev 83:116–121

Heitmueller A (2007) The chicken or the egg? Endogeneity in labour market participation of informal carers in England. J Health Econ 26(3):536–559

Heitmueller A, Michaud PC (2006) Informal care and employment in England: Evidence from the British Household Panel Survey. IZA discussion paper no. 2010, IZA, Bonn

Jimeno JF, Rojas JA, Puente S (2008) Modelling the impact of ageing on social security expenditures. Econ Model 25(2):201–224

Johnson RW, Lo Sasso AT (2000) The trade-off between hours of paid employment and time assistance to elderly parents at midlife. Report from the Urban Institute, Washington

Knapp LG, Seaks TG (1998) A Hausman test for dummy variable in probit. Appl Econ Lett 5:321–323

Maddala GS (1983) Limited-dependent and qualitative variables in econometrics. Cambridge University Press, Cambridge

Mundlak Y (1978) On the pooling of time series and cross-section data. Econometrica 46:69–85

OECD (2005) Long-term care for older people. OECD, Paris

Pavalko E, Henderson K (2006) Combining care work and paid work. Do workplace policies make a difference? Res Aging 28(3):359–374

Peracchi F (2002) The European community household panel: a review. Empir Econ 27:63–90

Pezzin LE, Schone B (1999) Intergenerational household formation, female labor supply and informal caregiving: a bargaining approach. J Hum Resour 34:475–503

Spiess CK, Schneider AU (2003) Interactions between care-giving and paid work hours among European midlife women. Ageing Soc 23:41–68

Verbeek M, Nijman T (1992) Testing for selectivity bias in panel data models. Int Econ Rev 33:681–703

Viitanen TK (2005) Informal elderly care and female labour force participation across Europe. ENEPRI Research report no. 13

Wolf DA, Soldo B (1994) Married women's allocation of time to employment and care of elderly parents. J Hum Resour 29(3):1259–1276

Wooldridge JM (2005) Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. J Appl Econ 20(1):39–54

Wooldridge JM (2007) Inverse probability weighted estimation for general missing data problems. J Appl Econ 141(2):1281–1301