

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

David Casado Marín^a, Pilar García Gómez^a and Ángel López Nicolás^{ab}

^a Department of Economics and Business, Research Centre for Economics and Health (CRES), Pompeu Fabra University (Barcelona).

^b Department of Economics, Technical University of Cartagena.

Abstract

Informal care is today the form of support most commonly used by those who need other people in order to carry out certain activities that are considered basic (eating, dressing, taking a shower, etc.), in Spain and in most other countries in the region. The possible labour opportunity costs incurred by these informal carers, the vast majority of whom are middle-aged women, have not as yet been properly quantified in Spain. It is, however, crucially important to know these quantities at a time when public authorities appear to be determined to extend the coverage offered up to now as regards long-term care.

In this context, we use the Spanish subsample of the *European Community Household Panel (1994-2001)* to estimate a dynamic ordered probit and so attempt 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 caregiving for more than a year has negative effects on labour force participation, the same cannot be said of those who “start caregiving” and “stop caregiving”.

JEL codes: J14, J2, I10

Keywords: informal care, female labour force participation, panel data models, ECHP, attrition bias

* The authors wish to thank the Institute for Fiscal Studies (*Instituto de Estudios Fiscales*) for financial support. David Casado also acknowledges the support of the Spanish Ministry of Education and Science through project SEC2003-045/ECO, while Ángel López and Pilar García do so with regard to projects SEJ2005-09104-C02-02 and SEJ2005-08783-C04-01.

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., 2006) 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, 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)¹. 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 (Edad & Vida, 2004).

Following the lead of other European countries, which have had universal public long-term care systems for some years, Spain has recently unveiled plans to develop a similar scheme known as the National Long-Term Care System (*Sistema Nacional de Dependencia* or SND) over the period 2007-2015². One of the main goals pursued through the SND, 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 SND seeks to make providing a certain amount of informal care compatible with having a paid job. This would not only avoid the negative consequences on an individual level associated with leaving the labour market (loss 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.

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 SND 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 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 the present 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 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 caregiving for more than a year has negative effects on labour force participation, the same cannot be said of those who "start caregiving" and "stop caregiving".

2. Background

From a methodological perspective, the main challenge involved in analysing the relationship between informal care and labour behaviour is the possible endogeneity of the former in a labour supply equation. 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 and 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 and 1996; Heitmueller, 2004; Crespo, 2006). As is well known, the success of this type of methodology depends on the availability of variables ("instruments") that are correlated with the regressor whose effect we are attempting to assess (informal care), but not with the dependent variable under investigation (labour force participation and/or hours worked). The instruments used in these studies 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 carers, despite using databases referring to different countries and different moments in time. Ettner (1995 and 1996) uses US data, as do Wolf and Soldo, yet obtains very different results: first, significant labour effects are detected when the woman providing informal care lives with the dependent person, basically in the form of a lower labour market participation; and second, although women caring for someone outside the household do not seem to participate less in the labour market, the number of hours worked is nevertheless lower than that of non-carers.

Ettner's results have been confirmed in part by more recent studies conducted using European data. Heitmueller (2004) uses instrumental variables to estimate, on the basis of the 2002 wave of the British Household Panel Survey, the effect on labour force participation of caregiving both inside and outside the household; his results show that only the first case results in a statistically significant decrease in the probability of being employed. Crespo (2006) uses data from the first available wave (2004) of the Survey of Health, Ageing and Retirement in Europe (SHARE) to estimate, by means of a bivariate probit, 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). Her results indicate that women who provide an "intense" level of care – i.e., live in the same household as the dependent, or give daily care elsewhere – have less probability of participating in the labour force both in the three southern European countries and in the three northern ones.

A third group of studies to be found in the literature is characterised, as we mentioned at the beginning of this section, by leaving the issue of simultaneity aside and concentrating on unobserved individual heterogeneity. In fact, although they use slightly different methodologies, the two studies that comprise this third group work from the same information source: the European Community Household Panel (ECHP). Thus, taking the first three waves of this survey (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. The results obtained, 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 fact of continuing to give care – not the fact of starting – that affects the number of hours worked, whereas in 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 to 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, on 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, Netherlands, Italy and Germany).

Finally, as we mentioned earlier, there is a fourth group of studies that have examined the relationship between informal care and labour behaviour by tackling both the simultaneity problem and the unobserved individual heterogeneity problem. 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 hours a year on the annual number of hours worked. The results obtained 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 both state dependence and individual heterogeneity. The model is estimated separately for two distinct samples of carers, which yield different results in each case: if we 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%).

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. 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, sex, 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 to 60 who participated in at least three 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. This subsample contains a total of 15,200 observations.

As can be seen in Table 1, caregiving was characterised according to two alternative classifications: first, given that several research studies have detected different effects on employment depending whether or not the dependent co-resides with the carer (Heitmueller, 2004), we divided the women carers in our sample into 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), the 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”.

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.

Table 1. Variables included in the analysis

Labour behaviour

nowork	1 if not working, 0 otherwise
fulltime	1 if working part-time, 0 otherwise
parttime	1 if working full-time, 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

Sociodemographic

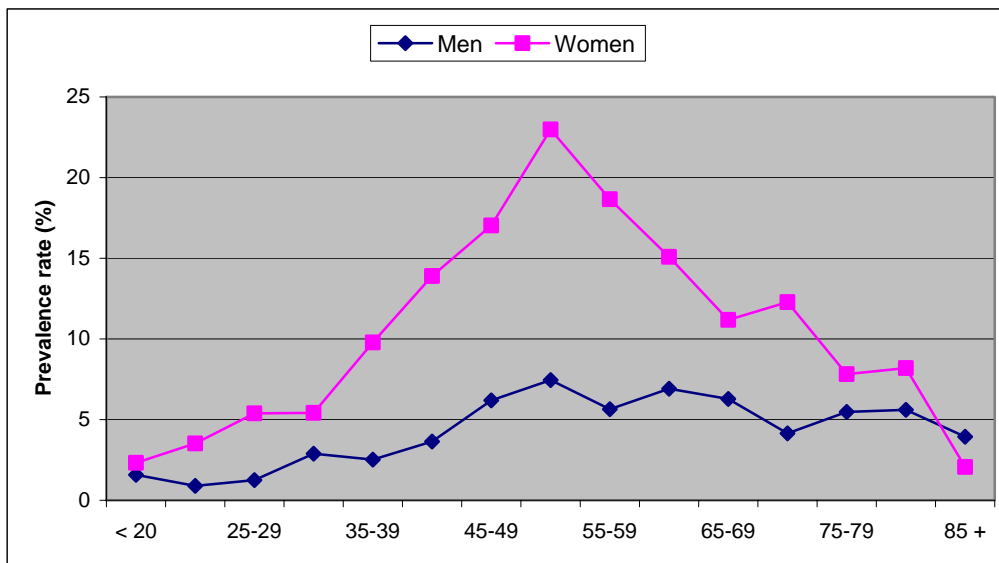
d3539	1 if aged 35 to 39, 0 otherwise
d4044	1 if aged 40 to 44, 0 otherwise
d4549	1 if aged 45 to 49, 0 otherwise
d5054	1 if aged 50 to 54, 0 otherwise
d5560	1 if aged 55 to 60, 0 otherwise
single	1 if single, 0 otherwise
widow	1 if widowed, 0 otherwise
sepdv	1 if separated or divorced, 0 otherwise
nch04	number of children in household aged 0 to 4
nch511	number of children in household aged 5 to 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
iscd57	1 if highest completed educational level is tertiary, 0 otherwise
iscd3	1 if highest completed educational level is secondary, 0 otherwise

Source: authors, based on the ECHP

3.2. Descriptive analysis

The relevance of focusing on the subsample of women aged 30 to 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 (Graph 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.

Graph 1. Percentage of informal carers: Spain, 1994

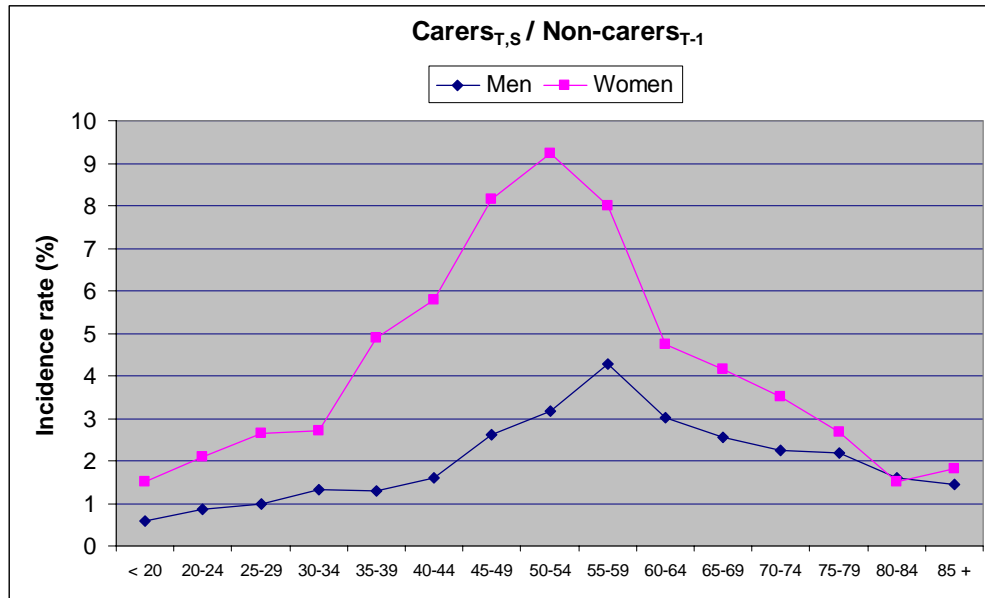


Source: authors, based on the ECHP

If we look at the dynamic incidence of the event “starting caregiving”, again notable differences can be seen both between men and women and in terms of age groups (Graph 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 caregiving role, it is generally the daughters and

daughters-in-law of these dependent people who provide the required help. However, when we come to 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.

Graph 2. Incidence rate of new carers (%): Spain, 1995-2001*



* Average incidence rates during the period.

Source: authors, based on the ECHP

Table 2 shows the descriptive statistics of the variables used in the analysis, calculated separately for women who care for and who do not care for a dependent adult over the various waves of the ECHP. The main features characterising the carers are as follows: labour force participation rate 13 percentage points lower; greater relative importance of part-time jobs among those who work; larger proportion of middle-aged cohorts; and clearly lower educational levels. 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: i) the differences between carers and non-carers as regards observable characteristics (age, marital status, educational level, etc.), ii) the existence of unobservable fixed factors (individual heterogeneity), iii) the state dependence that tends to characterise the labour behaviour of individuals over time, and iv) the attrition problems that tend to arise when working with panel data.

Table 2. Descriptive statistics of the variables included

	Non-carers		Carers	
	Mean	S.D.	Mean	S.D.
nowork	0.60	0.49	0.73	0.44
fulltime	0.34	0.47	0.23	0.42
parttime	0.05	0.23	0.04	0.20
carer_household	0.00	0.00	0.64	0.48
carer_elsewhere	0.00	0.00	0.36	0.48
starting caregiving	0.00	0.00	0.48	0.50
continuing caregiving	0.00	0.00	0.52	0.50
stopping caregiving	0.06	0.24	0.00	0.00
d3539	0.19	0.39	0.11	0.32
d4044	0.17	0.38	0.16	0.37
d4549	0.15	0.36	0.20	0.40
d5054	0.13	0.34	0.23	0.42
d5560	0.15	0.36	0.22	0.41
single	0.11	0.31	0.11	0.32
widow	0.04	0.19	0.05	0.22
sepdiv	0.05	0.22	0.04	0.19
nch04	0.18	0.45	0.17	0.44
nch511	0.34	0.63	0.35	0.64
sizehshd	3.78	1.35	4.22	1.44
sahvgood	0.16	0.36	0.11	0.31
sahgood	0.55	0.50	0.47	0.50
sahbad	0.07	0.26	0.10	0.30
sahvbad	0.01	0.12	0.01	0.11
lincome_ot	13.57	0.94	13.56	0.72
Northwest	0.15	0.35	0.17	0.38
Northeast	0.16	0.36	0.13	0.34
Madrid	0.10	0.30	0.08	0.27
Centre	0.14	0.35	0.17	0.38
South	0.18	0.39	0.19	0.40
Canaries	0.06	0.24	0.06	0.23
isced57	0.24	0.43	0.12	0.32
isced3	0.18	0.39	0.15	0.35
Number of obs.	24,469		3,151	

Source: authors, based on the ECHP

4. Methods

4.1. Econometric model

The econometric model we use to estimate the impact of informal care is an ordered probit.³ 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:

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

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 in microeconomic data, and ε_{it} depicts the purely random variation around the expected value of l_{it}^* (conditioned by 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

³ This is not the only possibility for modelling. For example, Bingley and Walker (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, 1991). Furthermore, its simple structure facilitates the introduction of delayed employment status variables capturing state dependence in the labour supply.

former along the lines of those proposed by Mundlak (1978) and Chamberlain (1984). That is,

$$\begin{aligned} \alpha_i &= \eta' \bar{X}_i + \lambda' l_{i0} + u_i \\ i &= 1 \dots N \end{aligned} \tag{2}$$

where \bar{X}_i contains the mean of vector X_i over T time periods for individual i , l_{i0} contains the values of the employment status variables in the initial period, and u_i is an independent random term of 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 (2002, p. 494) which consists in conditioning α_i to the initial employment status values⁴.

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

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

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:

⁴ See Contoyannis et al. (2004) for an application of this solution in the context of a model for self-perceived health status.

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

where Pr denotes probability.

There are two alternatives to consistently estimate the parameters of Equation (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:

$$\ln L = \sum_{i=1}^N \ln \int_{-\infty}^{\infty} \prod_{t=2}^T (\Pr_{itk} | C_{it}, X_{it}, l_{it-1}, \bar{X}_i, l_{i0}, u_i) du \quad (5)$$

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:

$$\ln L = \sum_{i=1}^N \sum_{t=2}^T \ln(\Pr_{itk} | C_{it}, X_{it}, l_{it-1}, \bar{X}_i, l_{i0}) \quad (6)$$

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 $\mathbf{l}_i = (l_{i2}, l_{i3}, \dots, l_{iT})$ 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 habitually in household panels, we come up against the phenomenon of attrition (Peracchi, 2002). Insofar as this attrition is related to the event we are interested in, the parameters – estimated by means of one or other of the two methods described 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 one of the following variables to the estimated model: i) a dummy variable indicating whether the individual has responded in the following wave; ii) a dummy variable indicating whether the individual is in the balanced panel; iii) a variable with the number of waves in which the individual has responded. The null hypothesis of non-existence of bias is rejected when any of these variables is significant, or when all three of them together (or any combination of two of them) are significant.

In any event, when the results for the attrition contrasts do not enable us to accept the null hypothesis of absence of attrition, it is possible to obtain consistent estimates even in the presence of attrition using the inverse probability weighting estimator, as suggested by Wooldridge (2002).

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 models are considered. The first includes the values in the first wave (1994) to estimate the response probabilities in the consecutive years. The second uses the values of the variables in $t-1$ to estimate the response probability in t ; in this case, given that 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 $\hat{p}_{it} = \hat{\Pi}_{i2}\hat{\Pi}_{i3}\dots\hat{\Pi}_{it}$, where $\hat{\Pi}_{it}$ represents the response probabilities estimated for each year (Wooldridge, 2002).

Lastly, we use the inverse of the predicted probabilities for each individual in each model ($1/\hat{p}_{it}$), or lacking that, the fitted probability in the second of the two models, 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:

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

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 two different models to assess the impact of caregiving on employment status. In both models the dependent variable is the ordered categorical variable $l_{it}=1, 2, 3$, corresponding to whether the woman declares herself to be in “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 – in the event that the woman is giving care in the current period – whether the place in which the care is given is relevant: caregiving at home, caregiving elsewhere, and non-caregiving. In Model 2, however, 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 give care in $t-1$ but does so in t), continue caregiving (gave care in $t-1$ and also does so in t), stop caregiving (gave 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). These four categories are parameterised with three dummy variables (one for each of the first three conditions). As mentioned earlier, both 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 proposed by Verbeek and Nijman (1992), in which the null hypothesis is the non-existence of this bias, are shown in Table 3 both for random effects specification and for pool specification. In these comparisons we can see that, when we use the dummy variable that captures whether the individual is in the sample the following year, the null hypothesis of non-existence of bias is rejected for both models under both specifications. However, with the rest of the variables, the null hypothesis is not rejected at conventional significance levels.

Table 3. Attrition test results*

	Ordered probit		Ordered probit (random effects)	
	Model 1	Model 2	Model 1	Model 2
Following year	1.66 (0.1980)	2.97 (0.0849)	4.09 (0.0431)	3.96 (0.0467)
All years	0.31 (0.5787)	0.01 (0.9167)	2.44 (0.1182)	2.44 (0.1185)
Number of years	0.14 (0.7074)	0.10 (0.7515)	0.51 (0.4733)	0.64 (0.4224)
Joint test	2.67 (0.4449)	4.27 (0.2335)	9.46 (0.0237)	9.31 (0.0255)

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

Note: Model 1 refers to the specification in which the treatment distinguishes between caregiving at home and caregiving elsewhere. Model 2 refers to the specification in which the treatment distinguishes between starting, continuing and stopping caregiving.

5.2. Model estimates

The rejection of the null hypothesis when the dummy variable “following year” is used suggests that it is necessary to use the IPW estimator. However, for the reasons mentioned in Section 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. Furthermore, in order to calibrate the robustness of the estimates, in Table 4 we present the results with two different ways of computing the weights \hat{p}_{it} . In the first of them we use information for the initial period (1994), whereas in the second we use information for the period t-1. Lastly, in the third column of Table 4 we present, for each model, unweighted estimates in order to be able to assess the potential impact of attrition bias.

Table 4 shows the results of the estimation of the various specified models. The first three columns of coefficients contain the results for the three models that distinguish between caregiving at home and caregiving elsewhere. Then the last three columns show the results of the three models estimated when we consider the dynamic behaviour of the carer; i.e., distinguishing between the effects in the first year, successive years and the year after caregiving stops.

For purposes of interpretation, although the scale of the ordered probit is arbitrary, the coefficients appearing in Table 4 enable us to know the direction of the effect of the different explanatory variables. The first outstanding feature is the importance of state dependence, as those 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 (IPW94 vs. IPW t-1) for both specifications considered.

Table 4. Dynamic ordered probit models for employment status

Variable	Treatment: caregiving at home or elsewhere			Treatment: starting, continuing or stopping caregiving		
	Ordered probit with IPW (1994 values)	Ordered probit with IPW (t-1 values)	Ordered probit without IPW	Ordered probit with IPW (1994 values)	Ordered probit with IPW (t-1 values)	Ordered probit without IPW
fulltime(1994)	0.8449	0.8885	0.8503	0.8494	0.8878	0.8529
parttime(1994)	0.3495	0.3896	0.3428	0.3527	0.3875	0.3444
fulltime(t-1)	2.0445	2.0755	2.0352	2.0386	2.0758	2.0290
parttime(t-1)	1.3033	1.2344	1.2229	1.3039	1.2429	1.2235
carer_household	-0.2331	-0.1583	-0.124			
carer_elsewhere	-0.2044	0.1223	0.0641			
starting caregiving				-0.0601	0.0224	0.0236
continuing caregiving				<i>-0.2081</i>	-0.1469	-0.1129
stopping caregiving				<i>0.1696</i>	-0.0298	0.0116
d3539	-0.0704	-0.0004	0.0144	-0.0691	0.0043	0.0195
d4044	-0.0205	0.1715	0.1497	-0.0217	0.1749	0.1535
d4549	0.0377	0.2213	0.1923	0.0296	0.224	0.188
d5054	-0.0163	0.1752	0.0889	-0.0303	0.1696	0.0767
d5560	-0.3515	0.0111	-0.1076	-0.3703	0.0003	-0.1204
single	0.7900	0.7195	0.6000	0.8185	0.7124	0.6201
widow	-0.3706	<i>-0.8120</i>	-0.3935	-0.3586	-0.8055	-0.3688
sepdiv	-0.114	0.0748	-0.082	-0.1092	0.0686	-0.0821
nch04	0.0934	0.0367	<i>0.0750</i>	0.0944	0.0455	<i>0.0791</i>
nch511	-0.0611	-0.0497	-0.0488	-0.0556	-0.0414	-0.042
sizehshd	-0.064	-0.0107	-0.0248	-0.064	-0.0131	-0.0268
sahvgood	0.1659	0.1425	0.1248	0.1589	0.1385	0.1177
sahgood	<i>0.1036</i>	-0.0095	0.0208	0.0986	-0.0109	0.0157
sahbad	0.2318	-0.0458	-0.0192	<i>0.2188</i>	-0.0467	-0.026
sahvbad	-0.1164	-0.1213	-0.0514	-0.1302	-0.1288	-0.0555
lincome_ot	0.0043	0.0254	0.0064	0.0042	0.0247	0.0066
Northwest	0.1377	0.0283	<i>0.0986</i>	0.1384	0.0301	<i>0.0924</i>
Northeast	-0.0185	-0.0913	-0.057	-0.021	-0.0945	-0.058
Madrid	-0.0586	-0.0852	-0.0858	-0.0609	-0.0887	-0.0876
Centre	-0.1922	-0.1868	-0.1957	-0.1948	-0.1909	-0.1975
South	-0.1749	-0.2314	-0.2114	-0.1749	-0.2290	-0.2128
Canaries	0.0039	-0.0352	-0.0043	0.0033	-0.0403	-0.0048
iscsed57	0.4111	0.3544	0.3643	0.4130	0.3566	0.3711
iscsed3	0.1684	0.1517	0.1800	0.1675	0.1486	0.1800
cut1	-0.1407	-0.176	-0.3158	-0.1313	-0.1694	-0.3259
cut2	0.194	0.1862	0.0251	0.2033	0.1913	0.0156
N	15200.00	12400.00	15200.00	15200.00	12500.00	15200.00
Log-likelihood	-6050.00	-4780.00	-6120.00	-6050.00	-4810.00	-6130.00

Note: **P-value <0.05**; *P-value <0.10*

5.3. Average treatment effect on the treated

As we mentioned earlier, the scale of the ordered probit model is arbitrary; therefore, 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 treatment effect on the treated (ATT); in other words, we have computed the extent to which the probability of being in each employment status changes for those individuals in each category in comparison with the same probability if they had not been carers.

Table 5 shows this average treatment effect on the treated for each of the models estimated. Graphs 3 to 5 depict both the average treatment effect on the treated and also the confidence intervals obtained by bootstrapping with 500 replications of samples with the same number of individuals resampled with replacement.

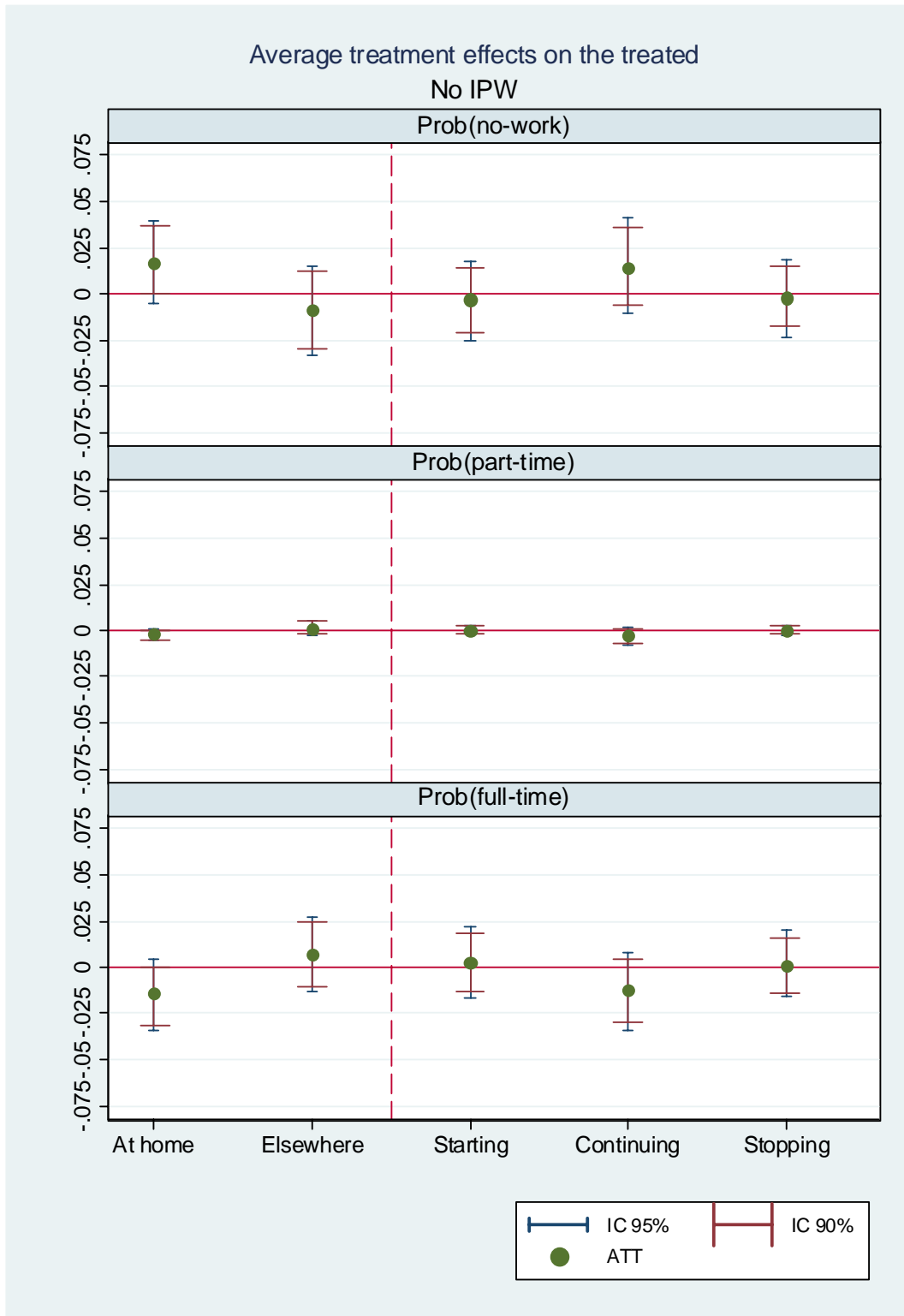
The results show, first of all, that the effects of informal care on employment status are restricted to the decision between working full-time and not working, since the estimated effects on the probability of working part-time are in all cases lower than 0.5%. Secondly, we find that caring for someone at home reduces the probability of working full-time, yet caregiving elsewhere does not appear to have any effect.

However, if we ask ourselves at what moment the possible change in employment status occurs, we find that it is as of the second year that the probability of working full-time diminishes. This is not surprising, in view of the number of hours dedicated to caregiving in each case. Lastly, the probability of being in one of the employment statuses considered is in most cases not significantly different between women who have stopped caregiving and women who have provided care neither in the current year nor in the previous year.

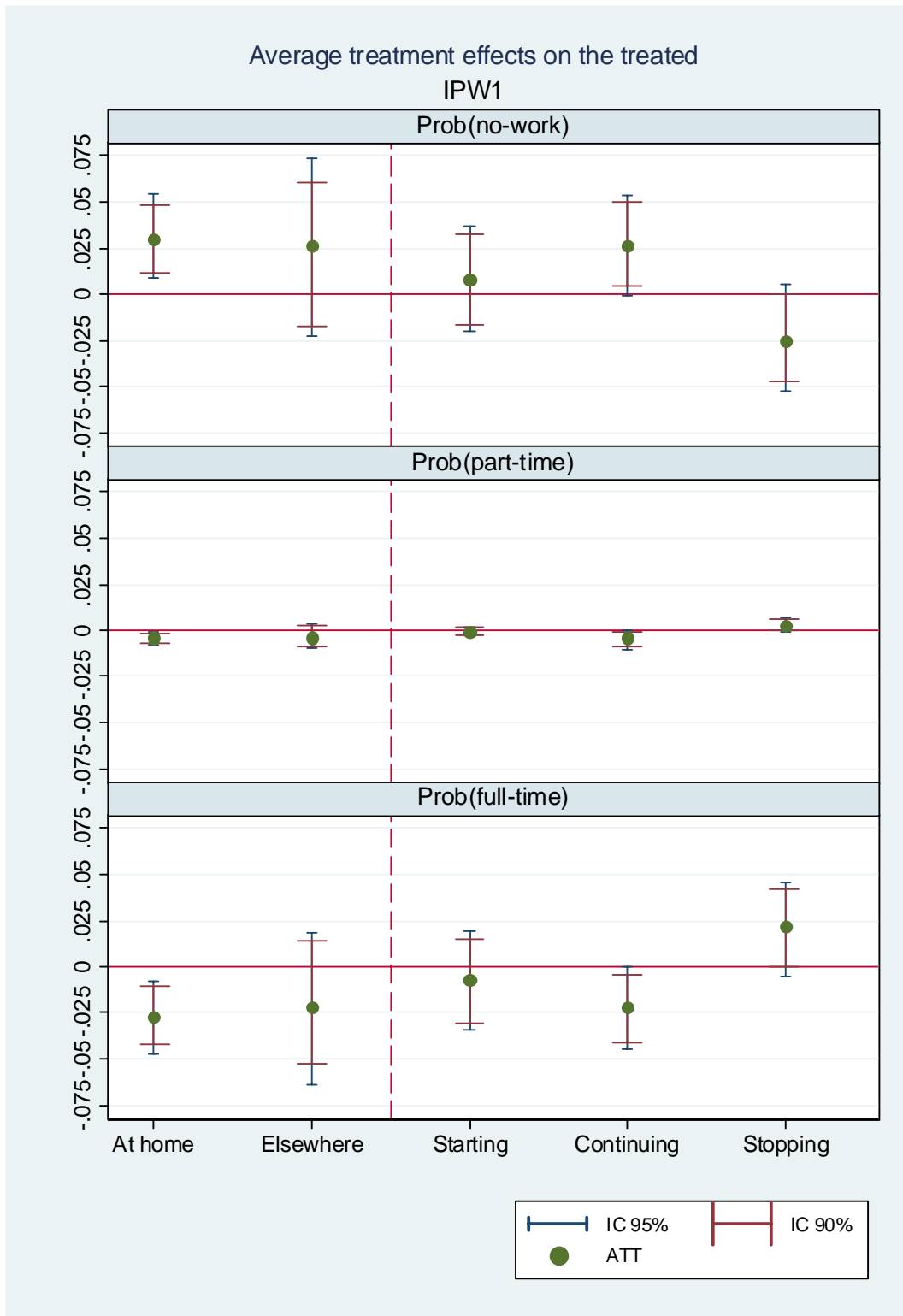
Table 5. Average treatment effect on the treated

	Prob(no work)			Prob(part-time)			Prob(full-time)		
	IPW	IPW-2	NoIPW	IPW	IPW-2	NoIPW	IPW	IPW-2	NoIPW
At home	0.03	0.020	0.017	-0.004	-0.003	-0.002	-0.027	-0.017	-0.014
Elsewhere	0.026	-0.016	-0.009	-0.004	0.003	0.001	-0.022	0.013	0.007
Starting caregiving	0.008	-0.003	-0.003	-0.001	0.0004	0.0004	-0.007	0.003	0.003
Continuing caregiving	0.026	0.018	0.014	-0.004	-0.004	-0.003	-0.022	-0.015	-0.012
Stopping caregiving	-0.025	0.004	-0.002	0.003	0.003	0.0002	0.022	-0.004	0.001

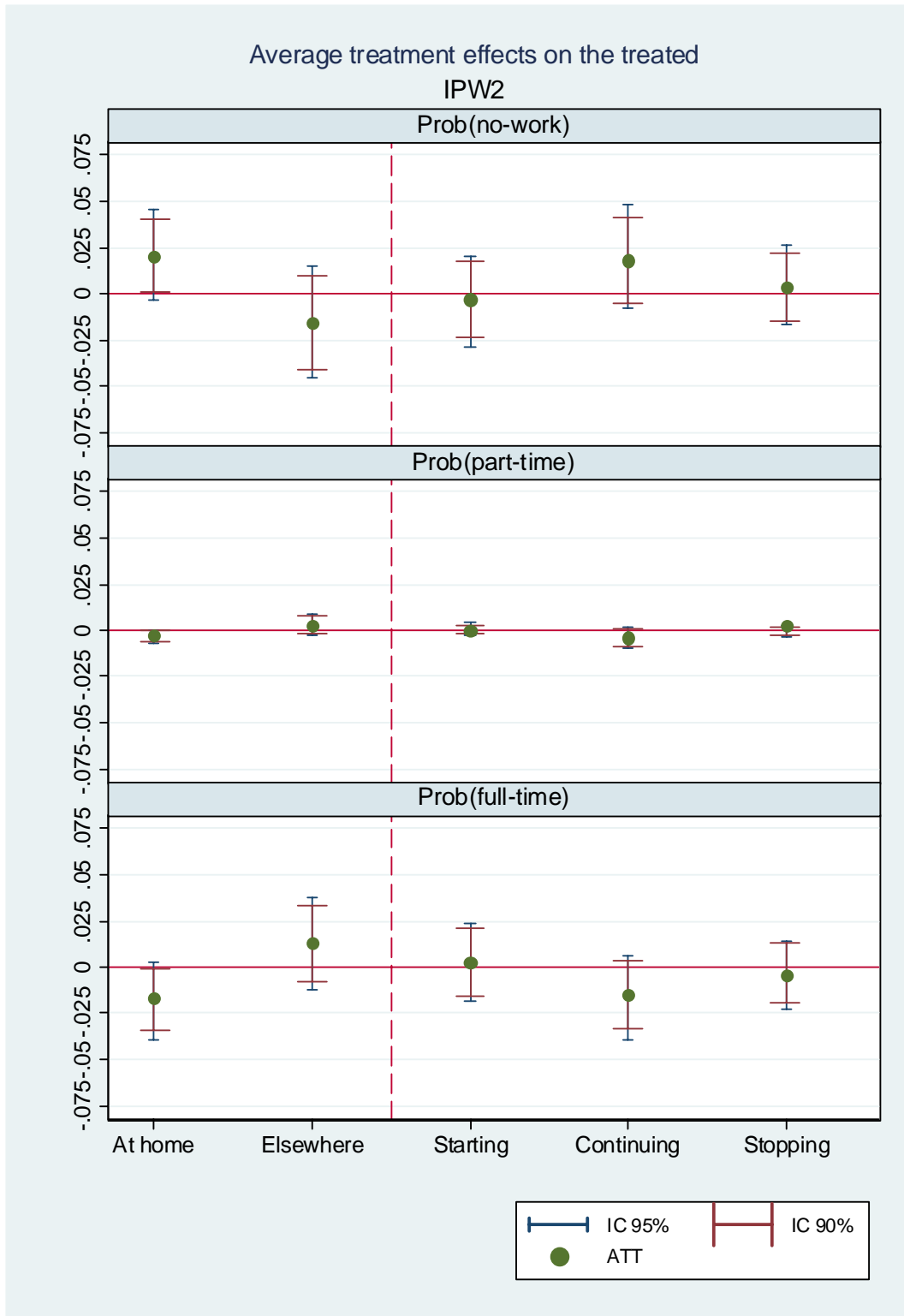
Graph 3. Average treatment effect on the treated: results without including weights in order to correct possible attrition bias



Graph 4. Average treatment effect on the treated: results including the weights that consider the characteristics of the individuals in the first period



Graph 5. Average treatment effect on the treated: results including the weights that consider the characteristics of the individuals in the previous period



6. Discussion and conclusions

The results obtained indicate that the labour effects of informal care occur basically among those women who care for someone at home, and among those who continue to provide care for more than one period. Those who finish an episode of informal care do not appear to have problems re-entering the labour market. As regards the type of job change made by women carers, they seem to be at the extensive margin (labour force participation), but not the intensive margin (hours worked), of employment decisions (Heckman, 1993). In this respect, although these results have been confirmed in other research on informal care (Ettner, 1996), 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 two types of implications for the design of public policies on long-term care. First, given that the labour effects appear to be concentrated in co-resident carers and those who provide care for long periods, the new SND benefits should be modulated according to the subjects' level of dependency; specifically, evidence exists for other countries to the effect that older people go to live with their adult children when dependency problems prevent them from continuing 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 result from chronic processes of a degenerative nature (Alzheimer's, cancer, etc.), in many cases continuity of care is also very probably capturing the seriousness of the dependency.

As the effects on labour behaviour seem to be reduced to an all-or-nothing choice (work full-time or stop working), the labour opportunity costs could probably be mitigated by developing more 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.

One important limitation of our analysis is that the results obtained cannot be interpreted, *stricto sensu*, in causal terms. Thus, despite having controlled for individual heterogeneity, part of the endogeneity problem probably persists, since we have not accounted for the simultaneity of the decisions “caring for someone” and “participating in the labour market”. In this respect, as a line of future research, we would suggest using the ECHP to replicate the analysis of the English case performed by Heitmueller and Michaud (2006) with data for more European countries. These authors use the British Household Panel Survey to conduct a dynamic analysis of the relationship between informal care and employment taking into account not only unobserved individual heterogeneity but also the simultaneousness of the two types of decisions.

References

- Ahn, N., J. Garcia and J. Herce (2005), *Health care expenditure and demographic uncertainty*, Documento de Trabajo 2005-07, FEDEA, Madrid.
- Bingley, P. and I. Walter (1997), "The labour supply, unemployment and participation of lone mothers in in-work transfer programmes." *The Economic Journal*, vol. 107, pp. 1375-1390.
- Cameron, A.C. and P.K. Trivedi (2005), *Microeconometrics. Methods and Applications*, Cambridge University Press, New York.
- Carmichael, F. and Charles, S. (1998), "The labour market costs of community care." *Journal of Health Economics*, 17: 747-765.
- Carmichael, F. and Charles, S. (2003), "The opportunity costs of informal care: does gender matter?" *Journal of Health Economics*, 22: 781-803.
- Casado, D. (2006), "La atención a la dependencia en España." *Gaceta Sanitaria*, 20 (Supl. 1): 135-142.
- Chamberlain, G. (1984), "Panel data" in Z. Griliches and M.D. Intriligator (eds.) *Handbook of Econometrics*, Vol. 1, Amsterdam: North Holland, pp. 1247-1318.
- Contoyannis, P., A.M. Jones and N. Rice (2004), "The dynamics of health in the British Household Panel Survey." *Journal of Applied Econometrics*, 19, pp. 473-503.
- Crespo, L. (2006), "Caring for parents and employment status of European mid-life women." Paper presented at the Second Workshop on Economics of the Family, held at the University of Zaragoza (9-10 November 2006).
- Edad & Vida (2004), *Estudio del modelo de atención a las personas mayores con dependencia en España*, Barcelona: Edad & Vida, 2004.
- Ermisch, J.F. and R.E. Wright (1993), "Wage offers and full-time and part-time employment by British women." *Journal of Human Resources*, vol. 28, pp. 111-133.
- Ettner, S.L. (1995), "The impact of parent care on female labor supply decisions." *Demography*, 32(1):63-79.
- Ettner, S.L. (1996), "The opportunity costs of elder care." *Journal of Human Resources*, 31(1):189-205.
- European Commission (2004), *Employment in Europe 2004: Recent trends and prospects*, Employment and Social Affairs, European Commission, Brussels.
- Heckman, J.J. (1981), "The incidental parameters problem and the problem of initial conditions in estimating a discrete time – Discrete data stochastic process" in C.F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications*, Cambridge MA: MIT Press.

- Heckman, J.J. (1993), "What has been learned about labor supply in the past twenty years?" *American Economic Review*, 83, pp. 116-121.
- Heitmueller, A. (2004), *The chicken or the egg? Endogeneity in labour market participation of informal carers in England*, IZA Discussion Paper No. 1366, IZA, Bonn.
- Heitmueller, A. and Michaud P.C. (2006), *Informal care and employment in England: Evidence from the British Household Panel Survey*, IZA Discussion Paper No. 2010, IZA, Bonn.
- Jimeno, J.F., J.A. Rojas and S. Puente (2006), "Modelling the impact of ageing on social security expenditures." *Documentos Ocasionales*, No. 0601, Banco de España.
- Johnson, R.W. and Lo Sasso, A.T. (2000), *The Trade-Off between Hours of Paid Employment and Time Assistance to Elderly Parents at Midlife*, report from the Urban Institute, Washington D.C.
- Mundlak, Y. (1978), "On the pooling of time series and cross-section data." *Econometrica*, Vol. 46, No. 1, pp. 69-85.
- OECD (2005), *Long-term Care for Older People*, OECD: Paris, 2005.
- Peracchi, F. (2002), "The European Community Household Panel: A review." *Empirical Economics*, 27: 63-90.
- Pezzin, L.E. and B. Schone (1999), "Intergenerational household formation, female labor supply and informal caregiving: A bargaining approach." *The Journal of Human Resources*, Vol. 34, No. 3, pp. 475-503.
- Spieß, C.K. and A.U. Schneider (2003), "Interactions between care-giving and paid work hours among European midlife women." *Ageing and Society*, Vol. 23, No. 1, pp. 41-68.
- Verbeek, M. and T. Nijman (1992), "Testing for selectivity bias in panel data models." *International Economic Review*, Vol. 33, No. 3, pp. 681-703.
- Vuittanen, T.K. (2005), *Informal Elderly Care and Female Labour Force Participation across Europe*, ENEPRI Research Report No. 13.
- Wolf, D.A. and Soldo, B. (1994), "Married women's allocation of time to employment and care of elderly parents." *Journal of Human Resources*, 29(3):1259-1276.
- Wooldridge, J.M. (2002), *Econometric Analysis of Cross Section and Panel Data*, Cambridge, MA: MIT Press.