I M P A L L A

International Master in Social Policy Analysis
by Luxembourg, Leuven, and Associate Institutes

KATHOLIEKE UNIVERSITEIT LEUVEN

FACULTY OF SOCIAL SCIENCES

DEPARTMENT OF SOCIOLOGY

What are the elderly caring characteristics influencing exit from work for European women?

A cohort and longitudinal approach

Supervisor: Prof. J. BILLIET
Rapporteur: Dr. P. HAUSMAN

MASTER THESIS
submitted by
Amandine MASUY

Academic year 2006-2007

EXECUTIVE SUMMARY

Population ageing is one of the major challenges the European countries have to face today. In the context of des-institutionalisation of health care and of responsibilisation of the community, the ageing 'burden' is re-turned to the family and especially to the women. At the same time, a high women employment rate is required to make the pension system sustainable. Therefore, more and more women in Europe will experience the situation of being carer and worker at the same time. But are these 2 roles compatible in the long term? What are the effects of caring for elderly on the working patterns?

The dissertation focuses on the relationship between **caring and the decision to stop work** among European working women. The research question is: ***Does caring for the elderly have an effect on the probability of quitting work for the European working women? What are the characteristics of caring that have an effect?***

The study sample is composed of **working women in the ECHP-1995**. The particularity of this study lies on its **longitudinal approach** and its **cohort comparison** perspective.

The results of the event history analysis on the all sample show that: 1) caring for elderly losses its positive effect on the probability of stopping work in multivariate analysis but when considering the characteristics of care, a significant effect of some predictors remain. The cohort approach shows that 1) the caring characteristics do not have an effect on the hazard rate of stopping work for the mid cohort (aged 31 to 49) and the old cohort (aged 50+) was the only one were characteristics of elderly care associated with a decrease in the odds of stopping work. 2) In the young and in the old cohorts, women providing light care are less likely to stop work than those not caring for elderly, while heavy care increase the probability of stopping work but only for the old cohort. 3) Neither women caring for elderly only, nor women caring both for children and elderly differ significantly from women not caring at all. This could be due to the small number of observations in those situations. Interestingly, caring for children only increases the propensity to give up work for young women but has the opposite effect for the old cohort. 4) Duration of care does not have an effect when the analysis is carried out on the 3 cohorts separately.

Policies aiming at keeping older workers longer at work should include measures helping the working and caring women close to retirement age to manage both work and care. More flexible schedules, temporary leave possibilities, respite care, and community care should be developed.

ACKNOWLEDGMENTS

I would like to thank,
Jaak Billiet for having allowing me to realise my project of dissertation on carers for elderly using longitudinal data and the methodological discussions we had,
Leen Vandecasteele and Hideko Matsuo for their methodological advice,

My father without who I would not have been able to prepare the data for the analyses,
Bruno Schoumaker[1] for making me benefit from his practical knowledge of the event history analysis and for the precious advice at the right time when handbooks were helpless,

Marie for very useful discussions at a crucial moment,
Jacques Hagenaars for methodological advice,
Judith Singer and John Willett for their very pragmatic handbook on longitudinal analysis and for having made syntax available on the net,

My mother for the discussions and re-reading,
Last but not least, Emi: I want to thank you for your unconditional support and your understanding: without you, I am not sure that I would have finished this Master.

[1] Professor at the Institute of Demography (UCL)

TABLE OF CONTENTS

LIST OF ANNEXES

LIST OF TABLES

LIST OF GRAPHS

SECTION I: INTRODUCTION

Population ageing & the caring-working dilemma for women in Europe

The process of **population ageing** is one of the main challenges the European countries have to face today. The fertility decrease together with the increase in life expectancy, have led to an irreversible and continuous increase in the absolute and relative number of very old and dependent people. According to the UN projections (2005) 10% (38 millions) of the EU-15 population will be aged 80 or more by 2050 (*annex 1*).

The question is: *How to meet the specific needs of this population in a sustainable way?* If residential and community care have been developed to various extents in the European countries, they cannot soak up the entire care demand of the elderly. *"The need to contain increasing costs in combination with the stated preferences of older people themselves to remain in their own environment for as long as possible, has led to what has been described as a* ***rediscovery of family care"*** (Johansson 2004)[2].

This question yields another one: *Is the today family able and willing to take up this responsibility?* Intergenerational solidarities are still very strong (Dooghe 1992). In most cases, the family, and especially women (Jenson & Jacobzone 2000), remain the main carers for the elderly[3]. The question could be re-formulated as: *How and to what extent can the caring role be added to the other social roles (as worker, child carer, etc.) women are fulfilling without jeopardize them*? Being better educated than the previous generations, the women of today are willing and pushed[4] to work. But, they are, at the same time, still considered as in charge of caring for children but also for the elderly. As the demand for elderly informal care is growing, the number of 'Multiple-role-women' is likely to increase as well[5].

Therefore, **knowing better to what extent women manage both care and work** would be helpful for drawing guidelines for **policies aiming at reaching, at the same time, a high level of women employment rate and sustainable informal care.**

---

[2] quoted by Mestheneos & Triantafillou 2005, p.7

[3] …before friends, neighbours and professionals (Van Tilburg 1997)

[4] One of the objectives of the Lisbon Summit (2000) was to increase the employment rate of European women. It is worth to notice that it is *"defined as key to improving the ratio between actives and retirees, as well as to broaden the contribution base for Social Security and general taxes"* (Sarasa 2005, p.5). Therefore increasing the employment of women is seen as part of the solution to population ageing (sustainability of the pension and health care systems).

[5] Demographic projections show that the elderly' spouse and daughters (in-law) should be the main sources of caring (Dooghe 1992).

This dissertation will focus on the effect of caring for elderly on the probability of quitting work. **The research question is twofold: *Does caring for the elderly[6] have an effect on the probability of quitting work for the European working women? What are the characteristics of caring that have an effect)?***

Most of the studies dealing with caring-working relationships focus on childcare (Spiess & Schneider 2002), among those investigating the relationships between work and elderly care, few concern European countries (Crespo 2006) and amongst them longitudinal methods of analysis are rarely used (Dentinger & Clarkberg 2002). Applying the event history analysis on the **ECHP 1995-2001** dataset to investigate the effect of **caring for elderly** (*hypothesis 1*) on the probability of quitting work of the 24592 surveyed European[7] working women, this dissertation is a tentative contribution to the knowledge of this phenomenon in a **European** and **longitudinal perspective**.

It is also interesting to note that the previous researches show conflicting results in this field: while some find a strong negative effect of caring on work, others do not find any significant relationship (Sarasa 2005). The second case could be partly explained by a too broad definition of caring. ***Multiple-role theory*** and the ***Lifecourse approach*** will be used to test

if elderly caring characteristics (e.g. intensity, duration) **have to be considered to find an effect of caring** *(hypothesis 2)* and **how these caring characteristics play over the active life stages**: resources and willingness for caring could change over time given the other roles woman have to fulfil *(hypothesis 3)*.

[6] *(Informal) care* refers to the daily help provided to an elderly relative or a friend (e.g. shopping, eating...).
[7] Except Luxembourg where the question on caring was lacking, the analysis concerns the EU-12 countries: Belgium, Denmark, France, Germany, Greece, Ireland, Italy, Spain, Portugal, The Netherlands, and UK.

SECTION II: LITERATURE REVIEW & RESEARCH HYPOTHESES

The literature review on the caring-working dilemma was constituted following a methodology explained in *annex 2*. In this section, the main results are presented according to the way the authors approach this theme and the theoretical backgrounds they used (1); some of their limitations and contributions will then be highlighted (2); and used to define the scope and the main hypotheses of this research (3). A comparative table of selected researches is displayed in *annex 3*

1. Caring for elderly-working as a 'dilemma' for women

A particularity of the caring-working relationship is that it can be studied in both directions: either the effect of work on care or the effect of care on work. Different populations can be considered: the caring women, the working women, or all the women[8] whatever their working status[9].

A. The decision of being at work among the carers *(working as dependent variable)***.** Studies on the carers' working status show that a selection process occurs among the carers who work: the **younger** (Barnes & Given 1995) and **better educated** are more likely to work, partly because the opportunity cost of caring is higher (Jenson & Jacobzone 2000). Another factor is the **intensity** of care provided to the elderly. Numerous studies highlight that carers providing intense care are less likely to work (Viitanen 2005) and this is observed in Northern countries and in Mediterranean ones as well (Crespo 2006). **Cohabiting with the dependent elderly** also decreases the likelihood of being at work (Heitmueller & Michaud 2006[10]). At another level, macro-variables such as **women employment rate** (*annex 4*), availability of formal care or labour policies, define the opportunities for seeking for or keeping a job for carers (Mestheneos & Triantafillou 2005)[11].

---

[8] A vast literature deals with gender differences in caring role but it will not be presented here because this research will focus on the effect of caring among women.

[9] The term '*work*' has been preferred to employment because work refers to the 'effective' status at the moment: either the woman works (whatever the kind of job she has – self-employed or not, full-time or not…) or she does not work (whatever the reason: unemployment, retirement, illness, etc). The '*working status*' is defined here as having 2 modalities: 'work' and 'not work'. 'Leaving the labour market', often used in the literature, does not fit the event analyzed in this paper: if the woman loses her job and becomes unemployed, there is no transition out of the labour market.

[10] quoted by Marín et al. 2006

[11] All these characteristics vary very much from country to country. For instance, the women employment rate of 2000 (*annex 3*) varies from more than 70% (Denmark), to 40% (Italy, Spain and Greece).It could be 'easier' in some countries to opt for a double career of worker-carer than it is in others.

B. The decision to care is influenced by working patterns[12] *(caring as dependent variable).* Another body of researches investigate the characteristics associated to the decision of becoming a carer. The findings on the influence of the working status on care are contradictory: Dautzenberg (2000, p.181) finds that "*paid work pulled women away from parent care in the sense that it reduced the likelihood that women become the carer of an elderly parent.*" Others recognize that caring is often assumed by individuals who are not working and that non-working carers provide more intensive care than others (Johnson & Lo Sasso 2000). However, they highlight that *"this is not because employed women are substantially less likely to assume caregiving responsibilities."* (Dentinger & Clarkberg 2002: 860). Other ones find no link between working status and the decision to become a carer (Spiess & Schneider 2002).

Other factors influence the probability of becoming a carer. For instance, Sarkisian & Gerstel (2004) show that adult children with a poor **health** are less likely to give help to their parents. As for the effect of education, it goes in both directions: because high **educated** women have usually higher wages, they can use it either for paying a full-formula of formal care (substitution) or for paying community care to complete their informal help in order to manage keeping their work while assuming part of caring responsibility too. (Sarasa 2005a). This author finds (2005b, p.18) that the **social class** and the associated cultural values also matter in the decision of working women to become a carer or not: *"farm workers are the most able to combine both loads probably because a mix of cultural values, work schedule flexibility and lack of substitutive services. On the opposite end, service class members are the least prone to assume heavy adult care giving duties"*. Finally amongst others, the **geographical distance** to a parent [...], the **cognitive and physical impairment of a parent**, but also **being unmarried** or **having many sibling** all affect the likelihood of becoming a carer (Dautzenberg 2000).

C. The allocation of time between caring and working *(working and caring as dependent variables).* Econometricians interested in the caring-working relationship, often consider it from the *allocation of time* theory (Becker 1965). The basic idea is that *"decisions on work hours and care hours are interrelated, because caregiving and employment compete for the caregiver's time resources. The model predicts that caregivers allocate their time in a way that an additional hour of time either in paid work, in leisure or caregiving generates the*

[12] *Working patterns* refer here to working-time.

same utility. [...] The trade-off depends on the relative marginal utility[13] *of paid work and care giving which, in turn, depends on relative wages."* (Spiess & Schneider 2002: 2). Systems of simultaneous equations are sometimes used for predicting both numbers of working-time and caring time.

Studies on the relation between changes in caring time and working time show that 1) **starting caring** adversely affects working-time [...]; 2) **increasing caring intensity** increases the probability of reducing working-time [while] the transition from non-work to work occurs most often when women **reduce caring** hours or stop caring altogether; 3) **stopping caring** is not associated with resuming usual working-time (Spiess & Schneider 2002). 4) Finally, Ettner (1995) shows that **co-residence** with a disabled parent leads to a large, significant reduction in working-time for women children, due primarily to withdrawal from the labour force.

Researchers studying simultaneously the determinants of caring and working highlight 2 ambiguities. The first is the **direction of the relation** (does caring influence working or is it rather the opposite?). The second concerns the effect of caring on working patterns. If, in most of the cases, caring is thought to have a **negative effect** (reduction in working-time or withdrawal from work), it could also have a positive effect: carers working more for having enough money to pay for formal care (income effect). Carmichael & Charles (1998)[14] point to *"two additional effects: [1] a demand for respite work: Participation in the labour market might be seen as a way of obtaining respite from the heavy caring work and the associated mental pressure; [2] a depressing effect on wages for those who do continue to undertake paid work."*

D. The influence of caring on quitting work *(stopping working as dependent variable).* Whether caring influences the probability of stopping work is also investigated. In the context of the Lisbon objective regarding women employment rate, understanding the mechanisms that lead working women to stop working is becoming more and more important. However, if the number of researches on that topic is increasing, the results are not converging. Stone & Short (1990)[15] find that some of the adult children who care for their parents are forced to rearrange their work schedules to accommodate for their caring responsibilities, but they do

[13] A sociological version of the marginal utility is found in the dichotomy 'home-centred'/'work-centred' lifestyles (Hackim 2000 quoted by Sarasa 2005). Women having a 'work-centred' lifestyle value more their careers and will try to avoid or to minimize caring. Women with a 'home-centred' lifestyle should more easily leave their job for caring (because they value more such home-related activity).
[14] quoted by Jenson & Jacobzone (2000, p.15).
[15] quoted by Johnson & Lo Sasso (2000, p.3).

not stop working[16]. Others authors find that, among working women, those caring are more likely than those not caring to leave work (Henz 2006; Dentinger & Clarkberg 2002).

When some specific caring characteristics are specified, the relationship with work seems more obvious. Pickard (2004) shows that when caring is provided for 20 hours a week or more **(intensive care)**, carers find it increasingly difficult to remain in paid work. Barnes & Given (1995), confirm that daughters who have left work are more involved in helping for the *Activities of Daily Living*[17] than the [still] working daughters. The consequences of caring depend also on the **personal characteristics of the carers** and on the **type** and the **quality of the carer's relationship with the dependent** (Henz 2004; Dentinger & Clarkberg 2002). Finally, the duration of caring (or rather the stability of the relation over time) plays a role. According to the *precarious ordering theory* developed by Wuest (1996), the caring relationship is a continuous process of '(re)-ordering': when a woman starts caring, the caring activity is quite disordered. Progressively, she will 'set boundaries' on the kind and the amount of care she is able to provide, she will 'negotiate' the division of caring responsibilities with the relatives and/or the professionals, and 're-pattern' her caring involvement in a way that allows managing 'competing and changing demands'. The process of precarious ordering is recurrent and is re-iterated each time that the demand for care changes. Therefore, the start of a caring spell can have bigger consequences on work. Once a combination of informal (and formal) care has been set up for responding to the elderly needs, the woman can more easily combine caring and working time. Henz (2006) shows that the probability of leaving work for a carer is higher at the **beginning of the caring spell,** than it is for a non carer, but that it decreases after a certain time.

Variables to control when studying the effect of caring on the probability of quitting work are also found in the literature. First of all, it has been proved that the propensity to leave paid work is far less frequent among **never-married** non-caring women than among women who have ever been married or who are cohabiting. The opposite is true for caring women: single women have a relatively higher rate of leaving work. (Henz 2006). Flexible work patterns allowing carers to better control their work environment, such as working from home or being **self-employed** can lower the positive effect of caring on quitting work (Twigg & Atkin 1994 in Pickard 2004). Working part-time or full-time also matters: Henz (2006) shows that carers working **part time** are more likely to leave work than full-time workers are.

[16] This study considers men *and* women carers. It could be a part of the absence of effect found.
[17] Activities of Daily Living (ADL) have been defined by Katz in 1969. It comprises help with washing, standing up, eating, going to the toilet and dressing). These activities are considered as the most demanding ones because they have to be performed every day without exception and following a fixed time schedule.

The author also highlight that the propensity to leave work is strongly associated with **social class**. Workers in routine occupations have the highest leaving rates.

Finally *"**age** is associated with different strategies of coping with caring demands at different life stages [for the older] a likely explanation is acquired pension rights. Once the respondent has obtained certain entitlements, leaving the labour market has less costly effects on the carer's future material standard of life. Younger carers, in contrast, reduce their commitment to the labour market without stopping work"* (Henz 2004). *"Women's age will also influence her perception of employment opportunities in several ways. Younger women are at the beginning of their careers, and face therefore potentially much more severe opportunity costs than do older women"* (Sarasa 2005, p.22). The determinant effect of age on the probability of quitting work leads researchers either to focus on a **specific age group** (e.g. Dentinger & Clarkberg 2002 study transition to retirement for the older workers) either to choose a **cohort analysis** (Henz 2006) referring to the ***Lifecourse theory*** that *"considers the timing of one's transition to the caregiving role, his or her competing roles in the context of the family, and the timing of life events such as retirement"* (Bengtson & Allen 1993) [18]. ***Role theory***[19] is another theoretical framework often used for explaining why women stop working. This theory is easily included in the framework of the Lifecourse approach. The idea is that people have different social roles to perform at the same time (spouse, parent, worker, child, carer for elderly, etc.). The constraints brought about by these various roles can produce pressure (*role strain*) and eventually lead the individual to give up one role to better fulfil the others. The association between caring and stop working can be interpreted as caring can produce role strain and constraints leading the woman to give up her worker role.

2. Contributions and limitations raised by previous researches

A. Geographical area. As seen, many researches have investigated the caring-working relationship. But most of these studies were done in the US (Crespo 2006) although the question of how women manage caring and working is, at least, as important in the ageing **Europe** of today.

The European researchers have started to address the question but most of the studies I have read so far either focus on one country (Henz 2006; Dautzenberg 2000) or on all Europe

[18] quoted by Kemp 2005, p.16

[19] The Role theory has been developed by Hill in the general context of family roles. The idea is that *"family roles are constructed by the expectations that other family members have to perform a particular role. Family crises [...] modify role expectations and, in turn, require behavioural change"*. (Kemp 2005, p.16)

but from an economical point of view (Sarasa 2005; Spiess & Schneider 2002) or also with the aim at explaining caring, (Marín et al. 2006; Crespo 2006), but **not aiming at explaining the caring effect on stopping working.**

B. Approach of the time. Few of the previous studies work on **longitudinal data,** while the caring-working relationship is, *per se*, an evolving relation over time (the effect of one on the other is not always immediate) and the choice the women have to made depends on her lifestage. As Henz (2006) and Dentinger & Clarkberg (2002) show, a longitudinal approach is needed when one wants to investigate the relation between caring and stopping working. The first author analyzes retrospective data and the second focuses on a specific cohort. **ECHP dataset allows using a longitudinal approach on panel data and considering different cohorts and their particularities.**

C. Sample selection. The results of the previous studies are not convergent, whichever approach of the relation is considered. Some authors find no relation, while others find a negative association between caring and working. Usually, bivariate associations are significantly negative. The problem arises in multivariate analyses. The divergence of the results can possibly be explained by sample selection - too small, not representative - (Spiess & Schneider 2002), by data limitations or econometric problems such as endogeneity or unobserved heterogeneity issues (Crespo 2006), by differences in estimation techniques or definition of caring (Johnson & Lo Sasso 2000). It has been shown that the effect of caring on stopping working is linked to the caring characteristics (intensity, duration, other competing roles, etc.). **Considering only the global effect of caring** (caring/no caring) **in multivariate analysis, could partly explain insignificant or poorly significant results.**

The 2 first approaches presented (working status of carers and workers who become carers) show that, choosing the working women as study population[20], implies not only a selectivity process among women[21] but also among carers. The first approach gives evidence that working carers are generally younger, better educated, providing lighter care, etc. They are not representative of the overall carer population. The second approach suggests that a woman will engage in a caring role only if she thinks that it is manageable with her other

[20] *Study population* (working women) has to be distinguished from *study sample* (the 24592 working women participating to ECHP in 1995).
[21] The selection issue will be discussed at the end of the *section III.*

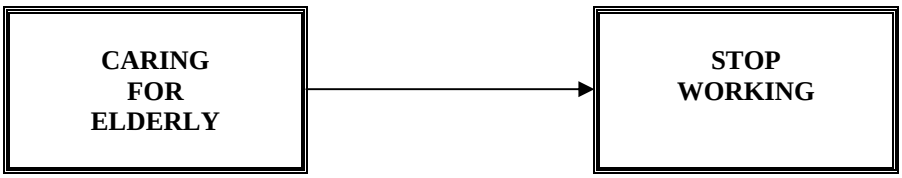
roles. Therefore the effect of caring can be lower because women who would have more difficulties in coping with caring and working will just not engage in both.[22]

D. Control variables. A last point to remember from the previous researches is the control variables. Studies aiming at predicting the probability of stopping working have found that work characteristics (flexibility, time-schedule, etc.), social class have to be controlled for. The age is determinant for understanding the circumstances in which the woman has to take the decision to stop or to continue working (multi-roles, opportunity cost of caring… are very dependent on her lifestage). Therefore, the Lifecourse approach is recommended. Studies on the determinants of caring show that education, wages, family situation and health status influence the probability of becoming a carer. They should therefore be considered also as control variables.[23]

3. Hypotheses formulation

Inspired by literature review and considering the information provided by the ECHP dataset, the hypotheses that will be tested in this paper are the following:

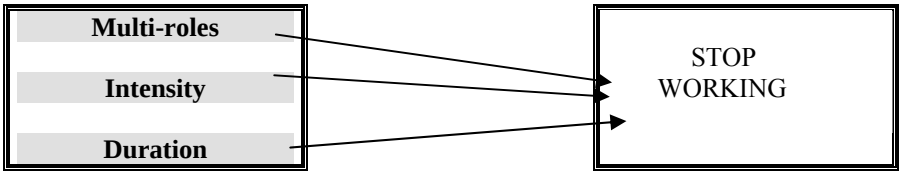
H1: Caring for elderly increases the probability of stopping work among European working women.



H2: The effect of caring on stopping working depends on the caring characteristics. Not all the working women involved in caring for elderly are as likely to stop working. Whether and to what extent caring influences the probability of quitting work depends on its characteristics (whether the care is light or heavy; whether the caring relation is a new one or long lasting; what are the competing roles the woman has to manage at the same time (*roles theory*).

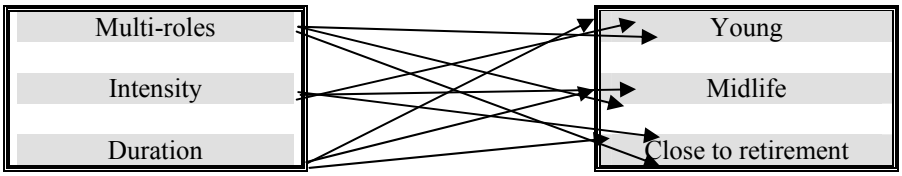
[22] This approach assumes that women can make a rational choice but it is not always possible. And even if a woman really weights the pros and the cons before starting care and reaches the conclusion that she can manage both working and caring, the constraints of her multiples roles may change and create roles strain that she did not had foreseen.

[23] The operationalization of the independent variables and the expected relation are presented in *section III.*



Control for
Socio-demographic, health & work characteristics

H3: The effect of the caring characteristics depends on the working woman's lifestage (Lifecourse approach). On the one hand, some kinds of care relations are more likely to appear at certain stages of the life (e.g. young women are less likely to provide care from a long time) on the other hand, some characteristics can have a stronger effect according to the period of the active life (e.g. having multi-roles roles to manage can have stronger effect for the women belonging to the old cohort who has become less attached to their work because they know that they will leave it soon.



Control for
Socio-demographic, health & work characteristics

SECTION III: TOOLS & LIMITATIONS

This section aims at presenting and justifying the choice of the dataset (1) and the statistical method (2) used for answering the research question, and at setting the limitations of this work (3). The section ends by displaying how the concepts and variables coming from the literature review are operationalized using the ECHP dataset, what is the expected direction of the relationship, and which are the potential reasons why no effect might be found (4).

1. Presentation and use of the ECHP Dataset

The analysis will be carried out on the European Community Household Panel (ECHP), a large-scale, longitudinal survey set up and funded by the European Union. The standardized questionnaire consists in different modules dealing with individual characteristics, household composition, income and expenditures, education, employment (working status and characteristics), etc. In the social relations module, **questions are asked on caring** activities at the time of interview (if the respondent is providing informal care, hours a week spent in care and if the care recipient is cohabiting or not). Therefore, having information on working and caring status as well as on socio-demographic characteristics, the research question can be addressed using the ECHP dataset.

Although the number of working women caring for elderly in the society is increasing, being a carer is not yet a widespread neither a fixed characteristic. Furthermore, the effect of caring on working can be not immediate. All that together leads to the conclusion that using panel data is necessary for addressing the research question: to follow the same individual over time increases the number of **transitions** of individuals having **a caring spell;** it also permits to consider the effect of **duration** of caring. For having the longest period of observation, it was decided to restrict the analysis to the 11[24] countries participating to the 8 waves of ECHP (from 1994 to 2001) and having included the questions on caring.

For Germany and the UK, the sample responding to the national survey (respectively SOEP and BHPS) will be analyzed instead of the ECHP survey that was not organized in these 2 countries for all the 8 waves. It was also decided to work on a **pooled** and

[24] Austria, Finland and Sweden joined the panel later on.

unbalanced[25] sample in order to have enough cases to make the analysis possible. The analyses were run over the **2nd to the 8th wave** because the question on intensity of care is not asked at the first wave. The study sample is made out of the European (11 selected countries) *panel*[26] women working in 1995 (2nd ECHP wave)[27]. A correspondence table between the8 ECHP waves and their use in the analysis is displayed in *annex 5.*

2. Statistical technique: Discrete-time Event history

The discrete-time event history has been chosen as the main method of analysis for 2 reasons. The first is that the nature of the study event requires a **longitudinal approach**: if one assumes that the decision of leaving work is not immediate, this transition in a specific period will be linked to the individual characteristics of the previous period[28]. The second reason is that being a carer for elderly (the main independent variable) is not necessarily a stable status. Event history analysis allows dealing with **time-varying predictors**[29].

A. Specific concepts and statistical tools. Some basic patterns need to be specified when using event history analysis. First, ***state(s)*** refer to the categories of the dependent variable (here: **working** or **not working**). Secondly, the study ***event*** is defined as the transition from one status to another: in this case, the **transition from working to not working**. Third, the ***beginning of the time*** refers to the moment (here: the time when the interview was made in 1995) when all the individuals in the sample are in the same initial state (here: working) and that no one has yet experimented the study event (here: stopping work)[30]. The first period

[25]Work on a balanced sample (here, only the women participating in the 8 waves) will dramatically decrease the sample size (and the number of transitions and caring spells as well). This would also lead to analyze a highly selected sample.
[26] By 'panel' women, we mean the women participating since the first wave (1994) of ECHP.
[27]Details about the sample and the selection bias are discussed in *section IV*.
[28]Given that the analyses are carried out on more than 2 waves, specific methods have to be used in order to identify the timing of the event and to use meaningful predictors to explain it. *"a simple logit approach […] has […] drawbacks that result from the fact that the choice of the period in which the event occurred or not is arbitrary*" (Vermunt 1997, p. 85).
[29]In a simple logit model, we should have chosen a fix value for caring (for instance, if the women is caring or not, in 1995). One can wonder the relevance of studying the relationship between caring in 1995 and quitting work status in 2000. If a woman has ever stopped caring for elderly in 1997 because she is depressed for any reason, and if she leaves her job in 2000 due to her psychological state, it could be very misleading to say that the fact that she stops work is related to becoming a carer in 1997. With event history, not only can the time of the event be located, but also, the time-varying nature of some predictors can be accounted for.
[30]The issue of the *left-censor* (people who already had experienced the event before the person-period under study) is discussed below (see 'limitations').

when it can occur is 1995-96[31]. Finally, the ***kind of study event*** should be specified. When defining the event as the **first exit** from work, it is made **unrepeatable**.

Because of the presence of *right-censored observations*[32], event history analysis needs to use specific statistics tools such as ***hazard*** and ***survival functions*** and ***median life*[33].** Providing that a **person period file** is created, the data analysis for discrete-time[34] event history analysis with a dichotomous dependent variable can be done using a simple binary logistic regression[35] (Allison 1995, p.254). The structure of the file allows dealing with time-varying variables and, interpreting the results in terms of hazard rates. The *annex 6* shows how the person-period file is created from the individual file.

B. Non independence of the observations. As the unit of observation is not anymore the individual but the *person-period*, a same woman can contribute to more than one observation. Actually, a woman will contribute to as many periods than the ones during which she is still at risk of leaving work *given she has not done so yet*. The problem of ***non-independency*[36]** of the observations belonging to a same woman is addressed by Singer & Willett (2003, p.384): *"the hazard function describes the conditional probability of event occurrence, where the conditioning depends upon the individual surviving until each particular time period and his value on the predictors in each time period. Therefore, it can be assumed that all records in the person-period data set are conditionally independent."*

C. Specification of predictors. Many specifications for the time-varying variables are possible[37]. It was decided to use 1-year-***lagged*** predictors: thus assuming that the woman's working status at the end of a given period is associated with the values of the predictors during the preceding wave/year. There are 2 main reasons for making this choice: (i) the predictors linked to employment will necessary be missing at the wave of the event (they are

[31] The events are recorded at the end of the period. Having experiencing an event in the period 1995-96 refers to the situation where a woman answered 1995 that she is working and in 1996 that she is not working.
[32]They are individuals for whom the event is not observed because they are still working when leaving the panel.
[33]These concepts are explained in *section IV*.
[34]Information on working status is recorded once a year, and not continuously.
[35]All the analysis were done using *SAS System* version 9.1.
[36] *"Because the same person is measured on several occasions, any unexplained person-specific time-invariant effect in the residuals will create a correlation across occasions. So, too, the outcome may have a different precision (and reliability) for individuals at different times, perhaps being more suitable at some occasions than at others. When this happens, the error variance may differ over time and the […] residuals will be heteroscedastic over occasion within person."* (Singer & Willett, 2003, p.37)
[37] For instance, we can consider the short term effect of caring for elderly (if the woman is caring at the current wave), the lagged effect (if the woman was caring 1, 2…years before), or a combination of it (if she was caring the previous year and still caring, etc).

state-dependent[38]); (ii) right-censor is the second reason: in the person period file, the last period a woman contributes to is either the period when the event occurs, or the one when she is 'censored' (see below). In the second case, there is information on the current value of the predictors[39].

3. Limitations of this study

Each research is subjected to potential biases. When it is not always possible to correct them, it is necessary to acknowledge them. Only the most relevant biases according to the research are presented and the possibility to correct them discussed[40]. Biases are presented by their source: the panel data, the ECHP questionnaire, and finally, the research theme.

A. Limitations due to panel data and the used method of analysis. *Attrition* is one of the main issues when analysing panel data. If among the study sample, those women leaving the panel before its end (here: 2001) differ from those who stay regarding the characteristics studied[41], then non-random attrition occurs and the results could be biased (Gallo et al. s.d.). A comparison between *attritors* and *non-attritors* is presented in *section III*. Because the aim is to predict the first exit (= the event) from work, women who leave the panel after having experienced the event are not problematic, because they are observed until they stop working. In other words, they are not right-censored. *Censoring* is another issue linked to panel data analysis. *Left-censors* (people that have already experienced the event before the observed period) are mainly a problem when one tries to predict the duration of stay within a given state. The problem in this case is that there is not information on the time the women started working. It is possible that women in the study sample did stop working for the first time in

[38] *"A predictor is state-dependent if its values at time t_j are affected by an individual's state […] at time t_j"* (Singer & Willett 2003, p.440). It is exactly the case here because the questions on employment characteristics (sector, time load, etc.) are only asked to the people who are still working at the time of the interview. Use of lagged predictors is the solution suggested by the authors in case of state-dependent predictors.

[39] The only censored observations for which current predictors values are available are those women who stay in the panel until the last wave without stopping working. They are then censored because the panel ends.

[40] In some cases, I have decided not to make the correction (because it was not possible on the available data or because it requires knowledge of specific programmes or methods that have not been taught in the IMPALLA programme). Time constraint does not allow me learning them. But, I will mention in which direction the results could be biased.

[41] For instance if the reason why the woman leaves the panel is linked to the event (she stopped working and is not at home at the moment of the survey because she is actively looking for a job - then the number of events in the study sample should be underestimated), or if the attritors have specific characteristics on important independent variables (e.g. caring for an elderly relative takes a lot of time and the women try to avoid other constraints, like spending time for answering a questionnaire) then we will lose cases that could give information on the main interest: the working pattern of the carer for elderly.

their life before the observed period[42]. Because the information available in the ECHP concerns only the period 1994-2001, it is not possible to distinguish the left-censor from the others. Solutions suggested by the literature (redefining the beginning of the time[43] or deleting the left-censors) are not applicable. Therefore, the solution chosen is **to (re)-define the study event as the first exit from work during the period 1995-2001. In this way, there are no left-censors**. As for the *right-censors*, event history analysis is a method allowing including them in the analysis. However, **the assumption of independence of right-censoring has to be made** (Singer & Willett 2003). A third concern, not specific to panel data but even more important to tackle in this kind of study is the *missing data*[44] treatment. Deleting a woman as soon as she has at least 1 missing value on one of the variables used in the analysis would dramatically reduce the sample, and it is very unlikely that missing data would occur at random. The solution applied **for the independent variables is to impute,** when available**,** the value of the preceding wave(s). This correction will allow carrying out the analyses on a larger and less 'selected' sample even if imputation implies a bias toward underestimation or delay of mutation on time-varying variables. The decision was also taken to not impute missing on the dependent variable to avoid misidentification of the timing of the event.

B. Limitations due to the design of ECHP and the content of the questionnaire. The first question to ask is *to what extent the dataset provides relevant information to answer the research question.* In the ECHP, respondents are asked if, without being paid, they are looking after children (1) dependent elderly (2) or both (3). If they answer positively, other detailed questions are asked dealing with the intensity of care and cohabitation with the care recipient. One can regrets the ***absence of information on other characteristics of caring*** which were proved to be important by the literature: the relationship between the elderly and the carer (Sarkisian & Gerstel 2004), the health condition of the elderly, the availability and use of community care (Marín et al. 2006) or working arrangements (Stier et al. 2001). Another weakness of ECHP for studying caring is the way the ***question on caring*** is formulated: only 'unpaid' caring is considered while in some of the ECHP countries,

[42] For instance, a woman who starts her first job in 1990 and stops working in 1994 but starts again working in 1995 would be in the study sample and left censored for the event stopping working for the really first time. For avoiding left censoring in this case, she should have been observed since 1990.
[43] In this case, a relevant starting point would be the time at the start of the first job for all the women (even if it corresponds to different times).
[44] Missing data refer here to what Rose (2000) calls *items non-response* (the person answers to the questionnaire but does not give an answer some questions) that he distinguishes to *unit non-response* (when the person did not answer the questionnaire at all).

allowances for carers are provided in order to compensate for the caring costs[45]. It could lead to an under-estimation of the real number of carers in those countries.[46]. Further, it is worth to notice that there is no clear, international, and unambiguous ***definition of what a carer is***. The way the carer is defined varies from survey to survey, what makes the comparison of results on the caring-working relationship difficult. In ECHP, no criteria of minimum intensity, of duration or kind of activities performed are set. Thus, whether the woman declares herself as carer or not depends on her own evaluation. The ***time-interval*** between 2 interviews is another important point to discuss. When using the discrete-time event history analysis, it is important that the time-interval (here: the year) is accurate enough to identify both the time when transition out of work occurs and what is the caring status at this moment. On the one hand, one can validly thinks that people do not enter and quit work each week and that a woman will not define herself as a carer if it happens that she exceptionally provides care for her mother during 1 week. When declared at interview, these 'statuses' refer probably to rather stable positions. However, short spells of inactivity may be not recorded as well as the effect of short caring spells on the work status at the moment of the interview might be overlooked[47] (Viitanen 2005), *annex 7*. This is also true for other independent time-varying variables. This limits the definition of the observed event defined here as: **not working at the moment of the interview** for women still working at the time of the previous interview.

C. Limitations due to the nature and the approach of the research theme. To study the event of quitting work using the event history analysis, it is needed to define a ***risk set*** composed by the individuals at risk of experiencing the event at a specific moment in time. The study sample is made out of all the women who are working in the ECHP 1995 and who are thus at risk of stopping working since that time. This choice may, nevertheless, involve a ***selection bias*** because the dependent variable is observed for a non-random sample (Spiess & Schneider 2002). **A comparison of the characteristics of the women included in the study sample and women not included** will be presented in the descriptive part to assess to which extent, and on which characteristics the selected women differ from the other women of the

[45] According to the Mestheneos & Triantafillou (2005) Report, such allowances exist in Belgium (Flanders), France, Ireland, Spain, Portugal, UK, The Netherlands and Denmark.
[46] Under-declarations can also arise when the carer does not recognise herself as such: caring is considered as natural and is more a characteristic of a relation than an independent activity.
[47] To illustrate this: a woman cares her aunt intensively between January and May 1998; she stops working in March and is still out of work in September when the 1998 interview takes place. She will not define herself as carer at this moment even if she was when she stopped working.

initial sample[48] (*annex 8*)**.** Another issue raised by the literature is the difficulty to establish the direction of the ***causality*** (Ettner 1995): *Does the working status determine if a woman will become a carer or not? Or is it the fact that she is a carer or not that determines her working status?* If this ambiguity on causality has to be noticed, I will not try to solve it in this paper because it is out of the scope of my research question[49]. A last limit due to the topic and its approach is that I have decided to work on pooled data and to carry the analyses on all the ***11 countries*** together (for justification of this choice, see *annex 9*). The purpose of the research is not to compare countries but rather to investigate a phenomenon – the caring-working dilemma the European women have to tackle. Because of the diversity of the European countries (in terms of cultural values, orientation and development of institutional settings and social policy targeting elderly and their carers), the results could better reflect the situation in some countries than in others. Nevertheless, elderly care and employment of women are on the increase in all the European countries during the 1990s. Therefore, the findings can be **interpreted as a sort of average situation in Europe for the period under study.**

These limitations led to a re-specification of the research question as follows: **Does the provision of unpaid care to the elderly affects the probability of quitting work during the period 1995-2001 for the European women at work? What are the characteristics of caring that really matter?**

[48] Heckman's correction and competing risks have been investigated as solutions for selectivity bias but is rejected because not applicable to the data or to the nature of the study event.

[49] Even if I will not investigate the causal effect of care on work, I did the test suggested by Taris (2000) to check the direction of the care-work relationship. The test shows that the magnitude of the effect of caring in 1995 on stopping working in 1996, controlling for not working in 1995, was bigger than the effect of not working in 1995 on caring in 1996, controlling for caring in 1995. The results confirm that it is more likely that caring explains stopping working than the opposite. Thus, it makes sense to investigate the influence of caring on working status among working women. *My aim is to discover the characteristics of caring that can predict the probability of quitting work and not to asses if caring is the very reason leading woman to stop working.*

4. Operationalisation and expected relationships

Concept	Operationalisation	Expected effect on stopping working	Reason why the variable could not show the expected effect
		Explanatory variables	
Cohort	*- age at the 1995 wave grouped in 3 cohorts:* *COH1*='1' up to 30 yrs: the '**young**'; *COH2*='1' from 31 to 49: the '**mid**'; *COH3*='1' for 50 and over: the '**old**'.	It can be assumed that the youngest cohort and the old cohort will be more likely to stop working, independently of their caring status, for different reasons (the first cohort because of child bearing and the last cohort because of being closer to retirement age)	The cut points for defining the age groups can influence the results. The choices were based on stages in the active life, assuming that usually, until 30, the woman is at the beginning of her career and is still not stabilised, that she is better installed in her midlife and that from 50, in most of the countries, retirement is felt as a possibility.
Old care	*- Present daily activities include, without pay, looking after children or other persons who need special help because of old age,…?* =>*OLDCARE*=1 if the respondent is caring for elderly (and/or child)	**Mid**: *positive effect*	Lack of precision about the kind of activities, risk of under-declaration of people who care without identify themselves as carer (see limitations). Also differences in women employment rate may affect the result. If only 1 country with low employment rate has been considered, a stronger relationship caring-working might have been found.)
Multi-roles[50]	*- Present daily activities include, without pay, looking after children or other persons who need special help because of old age?* => The 3 modalities are distinguished: child care only, elderly care only, child and elderly care[51]	**Young** : '*ONLY CHILD*' *positive* **Mid** : '*CHILD & ELDERLY*' *positive* **Old** : '*CHILD & ELDERLY*' *positive*	Too few effective caring at the same time for child and elderly (only 175 cases out of the 6917 waves with an event)
Intensity of elderly care	*- Number of weekly hours spent looking after elderly* *-Co-residence with elderly.* =>*HEAVY*=1 either if the respondent provides care for more than 2 h/a day or if the elderly lives at home. If the respondent is carer but does not fit any of these conditions, *LIGHT*=1	**Young, mid, old**: *positive*	The use of co-residence as a proxy of intensity could affect the relation because if co-residence is often linked to heavy care, it can also make the caring more manageable (no long journey between carer and elderly living place, less stress, support from other household members)

[50] Childcare might be a better indicator than the number of children. If a woman has children aged 16 or more, she usually does not have to spend hours to care for them daily (she is not providing childcare). As for the women of the old cohort, they may take care of their grandchildren and this can be time consuming (they are providing childcare).

[51] Note that *OLDCARE* groups the 2 last modalities

Concept	Operationalisation	Expected effect on stopping working	Reason why the variable could not show the expected effect
Duration of elderly care	*Present daily activities include, without pay, looking after children or other persons who need special help because of old age,…?* => if the respondent was caring the year before but not the previous year, *FIRST*=1; if the respondent was caring the year before[52] and also the previous one, *SEC*=1	**Mid, old**: *positive for the old and very positive for the mid* (according to the precarious ordering theory (Wuest 1997) the effect of caring can be higher at the beginning of the spell when caring responsibility and intensity is instable.	Because of left-censor on the variable (people who did start caring before the panel), the real duration of care spell cannot be estimated. The way *the variable* is built allows only distinguishing 1 year-spell from longer ones. The effect of care on work could appear with a delay (the accumulation of stress and fatigue can lead to stop working) According to Henz (2006), duration of the caring activity could indicate a selection effect.
Control variables			
Health	*- Health in general?* => *BADHEALTH* =1 if the respondent evaluates her health as 'fair' or 'worse'	**Old**: *positive* (it could lower the effect of care when controlling for it)	If the health status is a consequence of caring, the 2 effects could be difficult to distinguish Furthermore, those having a bad health will be less likely to care.
Education	*- Higher level of education completed* => if secondary, *SECON*=1 =>if superior, *SUP*=1	**Young, mid, old**: *negative* (higher opportunity cost: loss due to quitting the employment status are more important)	The considered levels of education could have different meanings over the generations and across countries
Social class	*Occupation in current job* => if manual, technicians, etc, *BLCOL*=1	**Young, mid, old:** *positive*	Social class can be highly correlated with education and 1 of the 2 variables could soak up all the effect.
Sector	*Status of employment* => if self-employed, *SELF*=1 => if public sector, *PUBLIC*=1[53]	**Young, mid, old:** public: *negative* (advantages and fringe benefit can help to manage work and other constraints. Self-employed: either *positive* if self-employed is demanding and time consuming, or *negative* if self-employed allows more flexible schedules.	The diversity of self-employment does not allow drawing one-way hypothesis on the direction of the relation.
Support	*-Number of household members aged 16+* *-Co-residence with elderly.* => If the number of household members aged 16+ (reduced by one if there is a cohabiting elderly) >= 2, *SUPPORT*=1[54]	**Young, mid, old:** *negative*	This indicator refers only to the potentially supportive cohabitants. There is no information on their characteristics, neither on their ability or willingness to provide effective support.

[52] The caring predictor is 1-year lagged. It means that the value of *CAREDUR* in person-period 1997-98 will be '2' if the women was caring in 1995-96 and 1996-97, '1'if she is caring in 1996-97 but not in 1995-96 and '0' if she is not caring in 1995-96.

[53] By building dummies in this way, we aim at distinguishing self-employment from private sector.

Concept	Operationalisation	Expected effect on stopping working	Reason why the variable could not show the expected effect
Part-time	*- Main job part-time/full-time? (Auto-evaluated).* => if the respondent claims to work part-time, *PARTIME*=1	**Young, mid, old:** *negative*: it can be assumed that those working part-time can more easily combine working-time with other constraints (it is not/less needed to stop working totally)	▪The question does not specify the number of hours. Part-time here means less than national full-time number of hours. The availability of part-time work, its valuation, and the linked rights vary from one country to another. ▪A positive effect is also possible, according to Henz (2006) by assuming that part-time workers are less attached to be at work.
Income from work as a % of household net income	*-Personal net income from work* *- Household net income*[55] => if the income coming from the work of the woman counts for at least half of the household net income, *PCTINC*=1	**Young, mid, old:** *negative* (opportunity cost too high).	The imputation for the household net income may not reflect the real situation.

[54] This indicator intends to capture the potential help the women can benefit from in her household for to either caring or completing her other roles. The correction (minus 1 if there is a dependent elderly cohabitant) avoids counting the elderly to care as a possible support. This indicator seems to us better than the fact to be married or not because a women can also be a carer for her husband (is this case, he cannot be considered as a support).

[55] I did not choose the wage of the women as economic predictor but rather the part her income represents in total household income. It seemed to me a better indicator of the opportunity cost of stopping work.

SECTION IV: EXPLORATORY ANALYSES

Exploratory analyses of the study sample are carried out for 2 reasons. The first is to check if the potential problems identified in section III materialize and if so, to evaluate how they could influence the results (1). The second reason is to introduce the explanatory analysis (with life table, bivariate association, etc.) (2).

1. Selection, attrition & missing data issues

The study sample made out of **24592** women working in 1995 were drawn from the initial sample of the **66195** women participate in 1994 to the first ECHP wave (1994). They produce **100483** person-periods to analyse.[56]

A. Selection. The difference in numbers between the initial sample and the study sample comes from the selection process. The next table (*table 1*) shows the reasons why women have not been selected and the number dropped for each reason (*annex 8*).

Table 1: Processing the study sample and removal reasons

Total women participate at the 1st wave (1994) of the ECHP (***initial sample***)	**66195**
Reason for removal	
Not observed in 1995	6705
Did not answer the working status question in 1995	77
Not working in 1995	34821
Study sample	
'Panel' women working in 1995	24592

Source: ECHP 1994-2001

As compliance to panel surveys can vary across countries, a first question deals with the link between the criteria of sample selection and the countries. *Table 9* (*annex 10*) shows that 20% of the initial sample of Ireland[57] is lost due to drop out of women not observed in 1995, and most of the women who did not answer the working status question in 1995 belong to the German sample (64 out of a total of 77). Otherwise, the main losses are due to the removal of the women having declared not to work at the 1995

[56] 6984 women belonging to the cohort aged up to 30 produces a total of 26414 person-periods; the 12919 women aged 31-49 57248 person-periods and the 4689 women aged 50 + 16821 person-periods

[57] Ireland is one of the countries for which attrition of the total sample is the highest (EPUNET 2004).

wave (this concerns more than half of the initial sample for Portugal, Spain, Italy, Greece, and France). The proportion of the initial sample remaining in the study sample varies from 23% (Spain) to 50% (Germany, UK, and Denmark). This proportion is closely linked to the **women employment rate** in the respective countries. When comparing the ECHP percentage of women working in 1995 across the considered countries[58] to the women employment rate published by Eurostat for 1995 (*table 10, annex 10*), one observes that the ECHP figures are slightly lower than those produced by Eurostat. It can be partly explained by the fact that no age interval has been imposed in the ECHP calculation.[59]

The study sample differs from the removed sub-samples (the non-working and not observed women in the 1995 wave) according to their respective **socio-demographic characteristics** (*table 11, annex 10*). The study sample is younger (median age 38.6 against 52 for the non-working and 45 for the not observed women) and more often higher educated (23.7% versus 6.8%, and 14.9%. Less than 25% of the study sample evaluates their health as fair or worse, while this is the case of 47.4% of the non-working women and 38.7% of the not observed in 1995. These differences are not surprising: it is known that the higher educated with a good health are more likely to work (the 'healthy-worker effect' is a well known selection bias in epidemiology). As for the family situation (*SUPPORT)* and the caring status *(OLDCARE)*, the sub-populations do not differ so much.

Remembering that the literature reveals a self-selection mechanism among the working carers (age, education, intensity of care), the presence of differences in the ECHP between working and non-working carers in 1995 is investigated: 9% of non-working women are carers this year against 6% of the working women. The non-working carers are older (median age: 56) than the working carers (median age: 45) (*table 12, annex 10*). In conformity with the literature, it can be noted that the working carers are more educated (49% have, at least, completed secondary school against 28% for not-

[58] The total is the sum of the women working in 1995 (study sample) and the women answering that they did not work at the 1995 wave.

[59] Percentage of women working in the sub-sample 'not observed in 1995' (attritor) has also been computed to see if these attritor women have a specific working status. Compared to the percentage of working women at ECHP 1995 the differences are not so big and vary across countries: attritors work more often in Ireland, Italy, and Spain and less often in the other countries.

working carers). They also differ in the intensity of the provided care: 6.3% of the not-working carers provide care for more than 2 hours a day, while this is the case of only 3.2% of the working carers.

In 1995, the working women carers are younger than the non-working women carers in each of the considered countries. Also, younger women, whenever working or not are less often carers than the older ones: 2% of the youngest age group (up to 30 years old), 6% in 31-49years group and 10% among the 50+). The carers in the study sample of working women are thus not representative of the general population of women carers, but this could be explained if the women are considered along their successive lifestages: the oldest cohort (aged 50+) is less likely to be still at work because they are closer to retirement age. Once retired, they should have more time for caring. At the same time, this cohort is more likely to have either a dependent parent or even a spouse to care for. There is a selection among carers who are working, and a selection among the working women who decide to care. Engaging both in work and caring depends on the stage in one's life (age, family and/or work career…) when the decision has to be taken, the multiple roles the woman has to manage and possibly the alternatives or support she can benefit from (formal care, other potential carers within the household, etc).

However, this selectivity does not prevent from investigating the effect of caring on working among the 'selected' working women population. To take the age selection into account, the relation will be investigated across broad cohorts defined by their age (the 'young' aged up to 30; the 'mid' or middle aged 31-49; and the 'old' aged 50+) at the 1995 ECHP wave.

B. Attrition. Attrition has been computed on the study sample (*table 13, annex 11*). 1404 out of the **24592** working women of 1995 did not answer the questionnaire the following year. Therefore, the sample size in 1996 is reduced by 5.7%[60]. Attrition occurs also the following years, but the number of women leaving the panel was almost stable over time. A total of **15342** women are observed until the end of the panel (62.39% of the study

[60] Permanent and temporary attrition are not distinguished. As soon as a woman does not answer the questionnaire once, all her data on the following years are considered as missing: for the missing data on the dependent variable, we cannot know if she had stopped working during her year(s) of absence.

sample)[61]. The question is: *whether or not the attritors differ from the non-attritors regarding the event and the independent variables that will be used in the analysis?* As for the **event**, (*table 14, annex 11*) it shows that the percentage of women stopping working among the non-attritors (34%) is nearly twice that of the attritors (18%). But the attritors are not observed over the whole period and, they would have been as likely to experience the event (random attrition) as the non-attritors if they would have been observed over as many years as the non-attritors. The same table shows that the 2 sub-groups are quite similar on the independent variables If one cannot state that there are no attrition bias at all, the descriptive results let us think that they are not so important.

C. Missing data. There are small percentages of non-responses for the variables used in the analysis. Nearly 15% of the women there are missing values for, at least, one wave on the question dealing with part-time, but this is an exception: for the other variables, the proportion of individuals without any missing value is close to 99% (*annex 12*). Imputation has been done for independent variables[62].

2. Preparatory analyses

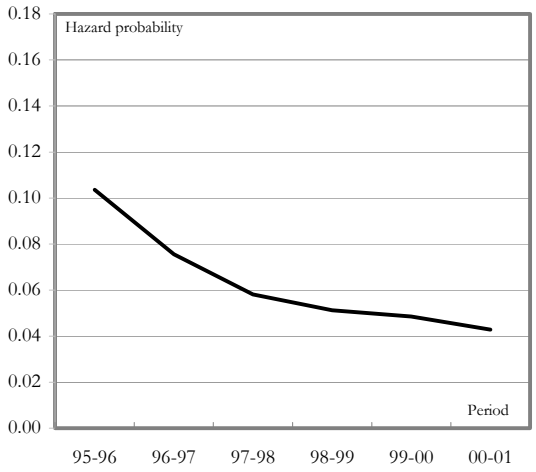
Before starting the event history multivariate analysis, the function of the time has to be specified and association between predictors and the event (stopping work) has to be checked.

A. Function of the time. Event history analysis allows to study events for which the probability of occurrence varies over time. In this case, the time measure is the year of the ECHP waves. I do not think that the probability of stopping work is higher in some of the study years than in the others: the hazard function could be flat (independent of the time). Nonetheless, because only the first exit during the period under study is considered, one can imagine that the hazard rate will decrease over time[63]. The first way

[61] Non-attritors are more frequent in the oldest cohort (**72%** compared mid: **69%,** young: **59%**.
[62] See *part 3 section III* in which the limitations of the study are discussed.
[63] In the reality, women can exit work more than once between 1995 and 2001 (5695 of the 6917 women who stopped working, re-start working and 37% of them (2089) quit again their work afterwards. If we would have chosen repeated event history, the function of the time would have been different. We did not make this choice because the research question deals with the effect of caring on stopping working and not with instable working lives having frequent transitions in and out work. The consequence is that, because

to check the time function is to look at the *life table* (*table 16, annex 13*) or the hazard function below (*graph 1*). It confirms that the hazard rate is quite high at the beginning (10% of the women working at the beginning of the time-interval 1995-96 will have stopped at the end of this period) and decreases afterwards. After a certain threshold, the decrease slows down. The time function seems to be **quadratic**. When comparing the results of logistic regressions testing different time-specifications (*table 16, annex 13*)[64], the quadratic function of time is preferred. Therefore, *TIME* and *TIMESQ* will be included in the regression analyses. The time is centered (*TIME*=person-period - 1) to make the intercept of the multivariate models meaningful[65].

Graph 1: Hazard function for the all study sample



Source: ECHP 1994-2001

B. Association between cohort and other predictors. To verify that the predictors chosen on the basis of theory are effectively associated with the dependent variable, separate logistic regressions were carried out for each predictor (bivariate analysis, *annex 14*).

---

these 'multiple event' women are withdrawn after experiencing their first event during the panel, the remaining risk set is composed of women who are less likely to experiment the event).

[64] Following Singer and Willett (2003), logistic regression on stopping working have to be done for the constant, linear, quadratic, cubic, forth power and general (all the period dummies). Comparing the AIC statistics and considering the parsimony criteria, the quadratic model seems to be a good choice.

[65] The intercept is the predicted hazard rate when all the predictors equal zero. By centering the time as period -1, the intercept gives the predicted hazard rate for the period 1 (1995-96) where *TIME*=0 for those individuals belonging to the reference categories for all the predictors.

Except from *SUPPORT*[66], all the predictors are significantly associated with the probability of stopping work. Women belonging to the old cohort are more than 3 times as likely to stop working the next year as those of the mid cohort (odds=3.38). The odds of stopping work the next year for high educated women is 40% (odds=0.39) to those having only primary diploma (reference category). Working part-time rather than full-time increases the odds of stopping work the next year by 2 (odds=2.02) while working in the public sector rather than in the private sector decreases the probability of stopping working the year after (odds=0.45). Comparing the AIC statistics[67], the cohort is the variable that explains the most in bivariate analysis. As for the effect of caring for elderly, it is significant in bivariate analysis (as shown in the literature) but its magnitude is not so high: becoming a carer increases the odds by 17% (odds =1.17). Cross-tables between the cohort and the other independent variables show that the distribution of the control variables varies very much from one cohort to another[68]. For instance, for 46% of the person-periods of the old cohort, women evaluate negatively their health status during the last year compared to 18% for the youngest and 28% for the middle-aged. In 30% of the person-periods the old cohort was self-employment the year before (compared to 14.5% for the middle aged and 8% for the youngest cohort). Finally, working part-time is also more frequent for the old cohort (30.7% against 24% for the middle aged and 18% of the youngest) (*annex 15*).

These 2 elements together led us to choose a 'cohort specific' analysis[69]. Life tables were computed for each cohort separately (*tables 20-22, annex 15*). The hazard rate (*graph 2*) shows that the 3 cohorts have the same shape: starting with a high hazard rate and decreasing slowly afterwards. The old cohort has nearly a flat hazard function and also the highest hazard for each period. One explanation is that, its members are

[66] Although the effect of *SUPPORT* is not significant, it is kept for the forthcoming analyses for two reasons: the first is that it is theoretically interesting and the second is that we cannot be sure that this variable will not have an effect when controlling for the other predictors.

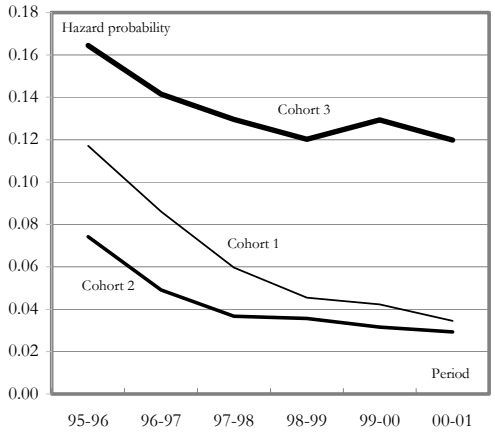
[67] Education and income AIC statistics cannot be compared to the others because the sample size is not the same due to the observations deleted for missing values. AIC statistics (or Akaike's Information Criterion) is a criterion for selecting among models. The AIC is a number associated with each model. The smaller the value, the best the model fits.

[68] Usually cross-tabulations are not used in data with censored observations. However, if we consider each observation as independent (see above), we can use it as an exploratory tool. The proportions are computed over the number of person-periods for the cohort and the characteristics refer to the observed value of the predictors at the previous wave (1-year lagged time-varying predictors).

[69] If the cohort is included as a predictor among others, it could soak up the effect of the other predictors.

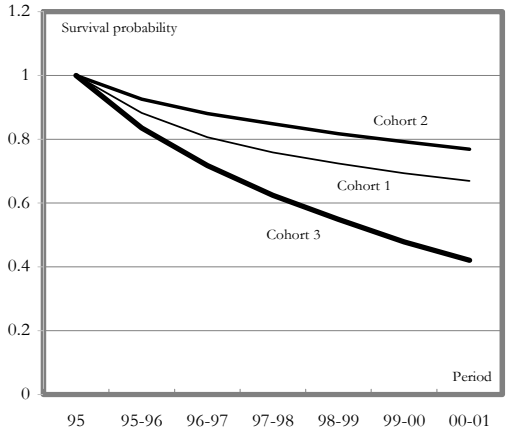
close to retirement age (stopping working is more frequent in this cohort[70] and those who have not stopped yet are as likely to stop afterwards[71]).The youngest cohort has the steepest curve: the hazard rate decreases fourfold, between 1995-96, (0.12) and 2000-01, (0.04). It can be partly explained by a higher proportion of instable careers (multiple events) and a selection among the survivors (those who have found a stable job and are trying to keep it) (*graph 3*).

Graph 2: hazard rate by cohort



Source: ECHP 1994-2001

Graph 3: Survival function by cohorts



Source: ECHP 1994-2001

[70] **Half** of the cohort experience the event (2350 over 4689) compared to **28%** of the youngest cohort (1945/6984) and **20%** of the mid one (2622 over 12919)
[71] They could be younger or wanting to delay retirement.

SECTION V: EVENT HISTORY MULTIVARIATE ANALYSIS

The main results of the analyses carried out of the person-period file are presented in this section. To begin with, a description of the characteristics of the person-period with an event and the person-periods without event is presented (1). Afterwards, the 2 first hypotheses are investigated: whether there is an effect of caring or not, or of its characteristics when considering the all sample (2). Finally, the third hypothesis will lead to a deeper analysis comparing the effect of the characteristics within the cohorts (what are the characteristics that matter for each specific cohort?) and between cohorts (does the effect of one specific characteristic differ from one cohort to another?) (3).

1. Descriptive analysis[72]

Considering the different person-periods produced by a same woman as independent, the control variables can be distributed over all the person-periods. Comparing the distributions over the person-periods where the event occurs to the others (without event) can give some hints on how the characteristics of the women having stopping work the next year differ from the others (*annex 16*).

When the total study sample is considered, the activity sector (*PUBLIC* and *SELFWORK*), the working time (*PARTIME)*, the percentage of the household income coming from the woman's work (*PCTINC)*, the educational level (*SUP*) and the health status (*BADHEALTH)* show different patterns for the 'event' and the 'non event' person-periods: the event (exit work the next year) person-periods show a higher percentage of women declaring having a bad health, being self-employed and part-time working, and a lower percentage of women whose work contribute at least half of the household income and who work in the public sector. However, the magnitude of the difference varies across cohorts.

In the **oldest cohort,** the differences between event and non-event person-periods are smaller. When having a bad health and working part-time or in the public sector is more frequent among those who will exit work the next year, for the other characteristics

[72] These descriptive results have to be considered as a kind of bivariate analysis (without controlling for other variables.

the 2 subgroups are quite similar. This can be explained by the fact that leaving work might be a **'normal' event** for this cohort reaching retirement age. Bad health can anticipate retirement. In some countries, it is also possible to take a 'progressive' retirement by decreasing working-time. This is a possible explanation for the observed differences in percentage of working part-time: to work part-time might precede stopping working. The public sector often offers working arrangements or temporary leaves (e.g. maternity, sickness, caring, etc.).

PARTIME and *PUBLIC* play also as discriminators for the 2 other cohorts (the '**young**' and the '**mid**'). The effect of part-time is consistent with the literature (Henz 2006) showing that women working part-time are less 'attached' to their job. The public sector offers also in several countries specific possibilities to take working arrangements or leaves. Having a bad health does not play for the **young** cohort. Other factors preventing leaving work for the 2 younger cohorts are to contribute for half of the household income and to be higher educated. This analysis highlights the importance of adopting a cohort perspective.

Notes: Because the number of carers is small, the effect of a predictor is considered as significant when alpha<0.1. Due to lack of room, the results of the analyses are put in annex 17 *(only the bivariate, the control, and the best models are presented in the main text). The methodology of model building is presented in* annex 18*. For each hypothesis, descriptive analyses linking predictors and caring characteristics were done for helping interpretations (*annex 19*).*

2. Caring and caring characteristics effects

A. Caring for the elderly influences the probability of leaving work (H1) (*table 24, annex 17*). As shown by the literature, caring for the elderly has a significant positive effect on the probability of stopping work the next year in bivariate analysis (odds=**1.17**). In the first multivariate model tested – 'full-control' (with all the independent variables) – (*model C*), *OLDCARE* turns to be not significant while all the other variables are[73].

[73] For all the control variables, the association is of the expected sign: feeling to have a bad health and having a blue collar work are associated with a higher probability of stopping work, while being more

Among the predictors, belonging to the old cohort has the biggest positive effect: the odds of stopping work for women aged 50+ in 1995 is 2.59 the odds of those aged 31-49. One hypothesis formulated was that the caring characteristics effects vary from cohort to cohort. If this is true, it could partly explain that including the cohort as control soaks up the effect of *OLDCARE*. Different models were fitted to obtain a multivariate model were caring is significant, the cohort predictors were the first withdrawn. When *PCTINC*, *PARTIME* and the cohort predictors are simultaneously excluded (*model D*), caring for the elderly has a significant but very small positive effect on the probability of stopping work the next year (odds=**1.09**)[74]. But as soon as one of these predictors is included, *OLDCARE* become insignificant. **Hypothesis 1 is thus infirmed**: caring for elderly does not have an effect on the propensity to stop work when controlling for other variables.

B. The effect of caring on quitting work depends[75] on the caring characteristics (H2) (*table 25, annex 17*)**.**The same model, with all the independent variables, is run separately for the 3 considered caring characteristics.

As for the **intensity of care**, it is interesting to see that *LIGHT* has a negative effect: working women providing light care are less likely to stop working the next year than those who are not caring for elderly[76]. A possible explanation is that these women are more attached to their job either because they need money to pay for formal care to complement the care they provide (income effect), either because working provides them with a breath from caring. Cross-tabulations show that these women are more often higher educated and having a white collar job than the others. The opposite is true for working women providing heavy care. These women have also a worse health status

---

educated, contributing for at least half of the household income, living with a 'supportive' adult (non significant) and working in the public sector are associated with a lower hazard rate of stopping work. For part-time work and self-employment, both directions were defendable. The empirical results show a positive relationship in both cases: self-employed and part-time workers seem to be less attached to their work. The direction of the relationship for the significant control variables is stable over all the following models.

[74] When comparing with the nested control *model E* – without *OLDCARE*), one can see that including it does not improve the fit of the model (the AIC are similar: *model D* = 46846, model E= 46848)

[75] 'Depends on' does not refer to interactions here. It means that, according to the characteristics of care considered in the model, the effect can be different. Because of small numbers, it is not possible to include in a same model all the caring characteristics into a single one.

[76] The effect is almost the same/similar (odds=0.83) in multivariate (*model D*) as in bivariate (*model A*, odds=0.88) analysis.

(38% against 28%). As expected, *HEAVY* increases the probability of stopping work the next year (gross[77] odds =1.50) compared to the non-caring women. However, this effect diminishes a lot and losses its significance in the full control model.

In bivariate analysis, both caring for the first time (*FIRST)* and providing care for at least 2 years consecutively (*SEC*) is associated with a higher hazard rate of stopping work the next year compared to those who do not care for the elderly. As expected, the effect of the first year caring is bigger (odds=1.21 against 1.15) but not so much. Both predictors lose their significance in the full control model.

The last characteristic tested is multiple-roles. As expected, *OLD&CHILD* has a positive effect. Caring at the same time for elderly and children is associated with a higher hazard rate of stopping work the next year than for those women not caring at all. The effect is even bigger in the full control model (odds=1.15). For the 2 others groups, the results are surprising. *ONLY CHILD* is associated with a lower hazard rate in the bivariate model while the expected positive effect is found in the full control model. The opposite occurs with *ONLY OLD*: the expected positive effect is found in the bivariate analysis but not in the multivariate (negative and insignificant). These changes can be explained by the fact the multi-roles vary very much from one cohort to another.

Hypothesis 2 is confirmed: when considering specific caring characteristics rather than the simple fact to care or not for elderly, the results are different. Intensity and multiples roles seem to be 'resistant' to the control variables while duration of care and the original variable *OLDCARE* are not.

3. Cohort approach (H3)

A. Cohort specific models. As it was the case in the all sample, duration predictors do not affect the hazard rate of stopping work. This runs counter the literature. However, it could be explained be the inaccuracy of the real duration of caring (ECHP does not allow identifying the start of the caring spell). It is also interesting to note that caring intensity and multi-roles play a role only in the youngest and in the old cohort. The working status of the mid cohort seems not to be affected by caring activities.

[77] Gross odds refer to the bivariate model.

*1) The **'young'** cohort* (*table 26, annex 17*)**.** Caring for elderly is not frequent in the youngest cohort (it is found that women provides heavy care in only 67 person-periods and women caring at the same time for children and elderly is found in 61 person-periods). It is therefore not surprising that these predictors do not have significant effects either in bivariate or in multivariate analysis. Actually, only caring for a child and providing light care are caring characteristics that have an effect on the probability of stopping work (both in bivariate and multivariate analysis).

As it was for the all sample, the effect of *LIGHT* is negative (gross odds=0.61). In the 'best' model for intensity (*model C*), without *SUPPORT* and *BLCOL* (insignificant in the 'full control' *model B*), but controlling for the other variables included in the model, the odds of stopping work the next year for young working women providing light care is 56% the odds of those not providing care to elderly. This negative effect of *LIGHT* is bigger than it was in bivariate analysis. In the 'best' model for multi-roles (*model E*), without *SELFWORK* and *BLCOL* (insignificant in the 'full control' *model – D*), controlling for other variables included in the model, the positive effect of *CHILDONLY*[78] is also slightly bigger (odds=1.44) than it was in bivariate analysis (odds=1.38). The significant effect of childcare is easily explained by the fact that leaving work because of child birth is very likely to occur at this lifestage. The duration of care provided to the elderly does not affect the probability of stopping work the next year. For instance, the odds of stopping work decreases by 27%[79] when moving from the first period (1995-96) to the second one. The important control variables (with larger magnitude and significant effects) are similar in both intensity and multi-roles models. All of them go in the expected direction. One can notice the high positive effect of *PARTIME*: *all things kept equal,* among the younger working women, the odds of stopping work for the part-time workers are almost twice (odds are about[80] 2.04) those of the full-time workers. Other characteristics particularly important in this cohort are: 1) being higher educated: it decreases the odds by 40% compared to those having a lower education level; 2) having a secondary school level has also a negative effect, but it is

[78] the reference category here is 'not caring at all'
[79] Odds are quite similar (around 0.716 *TIME* *1.02 *TIMESQ* =0.73) in all the best models.
[80] The odds are nearly equal in both models, the specification of caring and the other control variables do not really affect the odds.

smaller. Working in the public sector reduces the odds of stopping work by around 30% in both models[81]. Women contributing to at least 50% of the household net income from their job are about 35% less likely to give up work. The *SUPPORT* predictor has a particular behaviour: it becomes significant when multi-roles predictors only are included. Part of the explanation could be found in the fact that working women caring for children only are more likely to have a supportive adult (e.g. the father – *table 31, annex 19*). The 'best' fitted model for this cohort (with the smallest AIC) is the one including multi-roles predictors (*model F*)[82].

*2) The **'mid'** cohort* (*table 27, annex 17*). The exploratory analysis shows that this cohort is the one who has the highest proportion of survivors at the end of the panel (77%, *graph 3*). The multivariate models with the caring statistics demonstrate now that, controlling for the socio-demographic and economic variables, neither intensity nor multi-roles or duration have a significant effect (*models B, C, D*). In bivariate analysis (*model A*) each of these variables has, at least, one significant predictor: *HEAVY* is associated with an increase of 40% in the odds of stopping work compared to the 'mid' cohort women not providing care for the elderly. Compared to women not caring at all, women caring only for children are slightly more likely to stop work the next year (odds=1.18), and women caring both for children and the elderly are even more likely (odds=1.49). Finally, the hazard rate of stopping work the next year is higher for women being at the beginning of the care spell (*FIRST)* than for women not caring for the elderly. However, all those caring effects reduce in magnitude and turn to be insignificant in multivariate analysis. On the contrary, all the control variables are significant. For the mid cohort, only the socio-demographic and economic characteristics matter for predicting the exit from work the next year. Therefore, the 'best' *model (E)*, with the best fit, includes only these predictors[83]. It should be noticed that the 'mid' cohort is the only one where *BLCOL* keeps a significant effect in multivariate analysis. The odds of

[81] The exact odds are not mentioned because they slightly vary from one best model to another. However the size of the effect is similar in all of them.
[82] Because all the models refer to the sample, one can use the AIC statistic to compare their fit even if they are not nested (Singer & Willett 2003).
[83] Their effect goes in the expected direction.

stopping work decreases by 31%[84] when moving from the first person-period (1995-96) to the second one. The predictors having an important effect in the 'young' cohort have also an important effect for the 'mid' cohort: positive effect of *PARTIME,* negative effect of *PUBLIC, PTCINC, SECON,* and *SUPER*. But in this cohort, *BADHEALTH* has also a quite important effect while it was not the case for the young cohort. In the 'mid' cohort, declaring to have a bad health condition is associated with an increase of 46% in the odds of stopping work the next year, *all things kept equal*.

*3) The **'old'** cohort,* (*table 28, annex 17*)*.* As said before, the particularity of the 'old' cohort is that stopping work is, in some cases a 'normal' and expected event. However, controlling for the other variables, the time has not a negligible effect. For instance, the odds of stopping work decreases by 22%[85] when moving from the first person-period (1995-96) to the second one. By investigating the effect of caring, one can wonder if caring is a factor explaining the timing of (early) retirement[86] for this cohort. As in the young cohort, some predictors of intensity and multi-roles are still significant in the multivariate analysis and *BLCOL* is excluded from the 'best' models because of insignificant effect *once controlling for the other variables.* But the caring predictors that really matter and the changes occurring when going from the bivariate to the 'best' model, as well as the insignificant control variables differ. To begin with, *OLD&CHILD* is significant and positive in bivariate but not anymore in multivariate analysis. The reverse is true for *CHILDONLY:* its effect is negative and it gains its significance in multivariate. In the 'best' model, the odds of stopping work the next year for the oldest working women caring only for children is 87% the odds of those not caring at all. One possible explanation of this unexpected negative effect is that women of this cohort who are at the end of their career and are also caring for children (most probably their grandchildren) are the most 'active' ones. Also, caring for grandchildren is less constraining and probably better schedulable over time than caring for her own children. As for intensity, what is interesting here is that both, *LIGHT* and *HEAVY,* keep a significant effect in multivariate analysis and in the expected direction. *All things kept*

[84] Odds equals (0.66 *TIME* *1.05 *TIMESQ*=0.69) in the best model.
[85] Odds around (0.86 *TIME* *1.02*TIMESQ*=0.88) in all the best models.
[86] The available data do not allow testing if the woman stops working earlier when she is carer. However, the fact that caring as a significant effect indirectly means that.

equal, compared to the women who do not care for elderly, those providing light care are less likely to stop work the next year (odds=0.76) while those providing heavy care are more likely to do that (odds=1.18). The control variables having an important effect are the same as in the mid cohort. The 'best' model for this cohort (with smallest AIC) is the one with intensity caring characteristics.

***B. Cohorts comparison*[87].** The cohort comparisons concern only the intensity and multi-roles effects in the youngest and the older cohorts. The idea is to see if the caring characteristics that were significant in the cohort specific models are significantly different from one cohort to another. None of the characteristics were significant in the mid cohort and duration was not significant for any cohort. They are therefore not included in the comparison. The youngest cohort is the reference category.

1) Multi-roles, (*table 29, annex 17*)*.* As expected from the cohort specific model, *CHILD ONLY* is the sole predictor of multi-roles that is still significant in multivariate analysis when the sample includes both the youngest and the oldest cohorts. The 'best' model (model C) shows that, *all kept constant, BADHEALTH* (odds=1.26), *PARTIME* (odds=1.589) and, to a smaller extent, *SELFWORK* (odds=1.09) are associated with an increase in the hazard rate of quitting work the next year compared to their respective reference categories. On the other hand, *PCTINC* (odds=0.77), *PUBLIC* (odds=0.75), *SECON* (odds=0.83), *HIGH* (odds=0.64) and *SUPPORT* (odds=0.82) compared to their respective reference categories are associated with a decrease in the hazard rate.

The cohort continues to play an important role for predicting the hazard rate of stopping work. *Controlling for other variables in the model,* among the women who are not providing care to children only, the estimate odds of stopping work the next year for the old cohort is near 2 times as big as for the youngest cohort (odds=1.97). In the young cohort, women caring for children only have an odds of stopping work the next year 44% higher than women not caring at all (odds=1.43). The *CHILDONLY* effect is significantly different in the old cohort. As shown in the cohort specific model, caring only for children is associated with a lower hazard rate of stopping work in the oldest cohort

[87] The magnitude of the caring characteristics effect in the different cohorts could not have been compared using the cohort specific models (intercepts are different). For doing a real comparison, cohorts have to be included in the same model (same intercept) and cohort dummies as well as interaction between cohorts and caring characteristics should be included (*tables 29-30, annex 17*).

(odds=1.43*0.58=0.83). The importance of considering the cohort while investigating caring for children only is confirmed by the fact that the model with interactions have the lowest AIC statistics.

2) Intensity, (*table 30, annex 17*). Only *LIGHT* is significant in the 2 cohort specific models. In both cases, the effect on the hazard rate of stopping work the next year is negative. The question is: *whether or not this effect is significantly different (in term of direction and/or magnitude) in the 2 cohorts?* The 'full-control' (*model B*) shows that the effect of light care does not differ significantly in both cohorts (the interaction term is not significant). The best model excluded the interactions terms and the insignificant *BLCOL* control variable[88].

Hypothesis 3 is confirmed, whether or not caring characteristics have an effect, but also the direction of the effect varies from cohort to cohort.

[88] All the predictors are significant and their effect goes in the expected direction.

Table 2: Best models for the 3 hypothesis

Best models	Hypo. 1 (see table 24)		Hypothesis 2 (see table 25)						Hypothesis 3[89]			
			Multi-roles		Intensity		Duration		Multi-roles (see table 29)		Intensity (see table 30)	
Predictors	**Odds**	**Pr>Chi**	**Odds**	**Pr>Chi**	**Odds**	**Pr>Chi**	**Odds**	**Pr>Chi**	**Odds**	**Pr>Chi**	**Odds**	**Pr>Chi**
Intercept		<.0001		<.0001		<.0001		<.0001		<.0001		<.0001
TIME	0.708	<.0001	0.729	<.0001	0.729	<.0001	0.729	<.0001	0.775	<.0001	0.781	<.0001
TIMESQ	1.035	<.0001	1.033	<.0001	1.033	<.0001	1.033	<.0001	1.022	0.0037	1.022	0.0032
BADHEALTH	1.472	<.0001	1.347	<.0001	1.339	<.0001	1.342	<.0001	1.262	<.0001	1.259	<.0001
SECOND	0.711	<.0001	0.767	<.0001	0.773	<.0001	0.772	<.0001	0.826	<.0001	0.825	<.0001
SUPER	0.52	<.0001	0.611	<.0001	0.617	<.0001	0.615	<.0001	0.638	<.0001	0.641	<.0001
SUPPORT	0.886	0.0017	0.791	<.0001	0.81	<.0001	0.803	<.0001	0.825	0.0007	0.854	0.005
BLCOL	1.089	0.0013	1.062	0.0264	1.063	0.0246	1.065	0.0205				
SELFWORK	1.249	<.0001	1.116	0.0017	1.119	0.0013	1.125	0.0007	1.089	0.0635	1.103	0.0333
PUBLIC	0.566	<.0001	0.613	<.0001	0.616	<.0001	0.616	<.0001	0.753	<.0001	0.767	<.0001
PARTIME			1.592	<.0001	1.626	<.0001	1.622	<.0001	1.589	<.0001	1.654	<.0001
PCTINC			0.679	<.0001	0.679	<.0001	0.676	<.0001	0.768	<.0001	0.772	<.0001
COH1			1.632	<.0001	1.604	<.0001	1.605	<.0001	1.972	<.0001	1.681	<.0001
COH3			2.711	<.0001	2.601	<.0001	2.599	<.0001				
OLDCARE	1.092	0.0753										
HEAVY					1.108	0.1275					1.13	0.1645
LIGHT					0.832	**0.0189**					0.718	0.0023
CHILDONLY			1.119	**0.0002**					1.434	<.0001		
CHILD&OLD			1.153	**0.0947**					1.114	0.4527		
OLDONLY			0.944	0.3676					0.892	0.1456		
FIRST							1.067	0.3731				
SEC							0.902	0.1402				
CHILDONLY*COH3									0.580	<.0001		
Sample size[90]	98582		97345		97345		97345		41173		41173	
AIC intercept	49002.00		47473.00		47472.51		47472.51		25866.57		25866.57	
AIC covariate	46846.00		44105.00		44111.15		44116.37		24590.81		24629.81	

ECHP 1994-2001

[89] The results displayed refer to the cohort comparison models (for young and old cohorts).

[90] The sample size refers to the number of observations (person-periods) used for the analysis (after exclusion of the observations with missing values on at least 1 of the predictors).

SECTION VI: DISCUSSION

This research has been presented as a 'tentative' contribution.
The effect of caring on stopping working is not easy to tackle. The number of carers found in the study sample is not high. Furthermore the ECHP dataset is not specially designed for tackling the caring relationship. Previous researches displayed mixed findings due to selection or other bias and it is not possible to avoid all of them.

1. Findings

Notwithstanding, some interesting results were found. Especially, **the cohort approach has** proved to be relevant for investigating the caring characteristics effects. First of all, the caring characteristics do not have an effect on the hazard rate of stopping work for the mid cohort. In the 2 other cohorts (young and old), women providing light care are less likely to stop working than those who do not care for elderly, while heavy care has a positive effect but only for the old cohort. For the multi-roles effect, only childcare have still an effect in the multivariate analysis. Interestingly, caring for children only increases the propensity to give up work for young women but has the opposite effect for the old cohort[91]. **Longitudinal approach** highlighted that, at the beginning (without controlling), the hazard rate of (first) exit from work during the period 1995-2001 differ from one cohort to another. For the old cohort, next to retirement age, the hazard function is flatter than in the 2 younger cohorts.

2. Small size effect

However, the magnitude of the caring characteristics effect is very small. Caring for elderly is rarely the main reason for stopping work. Structural variables like current economic situation or 'lifespan events' like child birth, retirement, etc. could explain more. There are many ways to help working women to care for elderly without having to give up work. Using formal care as complement, reducing working-time (Sarasa 2005, p.

[91] The effect of caring for children only is not directly linked to the caring for elderly effect. It would be too easy to conclude that caring both for children and elderly does not have an effect on working patterns. Literature on roles strain shows the opposite. A reason why the results do not show a significant effect could be the small number of cases in the situation and especially within the youngest cohort.

25), changing job, decreasing leisure time (Barnes & Given 1995, p.375) are some of the alternative strategies used to care when keeping work[92].

The effect of the control variables is also quite small. From the literature review, most of the important control variables were included, at least in the first model. But one important is missing: the country. While working on the all sample of the 11 European countries was justified for methodological reasons, it could also explain the weakness of the effects found: some characteristics, like part-time refers to different country specific situation and the effect can vary in magnitude and even in direction from one country to another. Working women from countries with an overall lower women employment rate may be more 'selected' in term of level of education, attachment to work, career orientation. Finally, the diversity of 'women-friendly' working and caring policies will also affect the overall caring-working relation.

3. Dataset and statistical methods

For sure, ECHP is not oriented towards ageing and intergenerational issues. The question used to identify carers is not precise enough and questions on important aspects of care are simply missing (e.g. the health status of the care recipient, the relationship between the carer and the care recipient, the use of formal care)[93]. Specific National surveys on the carers' population exist but they do not allow comparing carers to not carers. Recently, a European survey on the 50+ population as been launched (SHARE) but most of this population is close to retirement, which excludes a cohort analysis. At this day, only the first wave is available for the research and longitudinal analysis is not yet possible. Therefore, despite the limitations of the dataset, to address the research question using the ECHP was worthwhile: question on caring is asked and longitudinal and cohort approaches were possible.

[92] In a qualitative study, Arksey (2002) interviewed working carers on how they manage both caring and working. Various strategies were recorded: rearranged work schedules, self-employment, reducing working-time, using lunch-time for caring (related) activities, reducing travelling time between care and work places and living place (by cohabiting with the care recipient, for instance),etc.

[93] The national surveys for UK and Germany are slightly different from the ECHP core questionnaire. The question on intensity of care is not asked in Germany. Therefore, the co-residence with elderly has been used as a proxy for intensity in this country.

4. Sample and outcome focus

This paper focuses on a specific population of carers (working women) and, on a specific work outcome (stopping working).

The population of carers is more diverse: men are increasingly becoming and more carers (especially when there is neither daughter nor wife available for caring). Also, given the increase in life expectancy without disability and the 'verticalisation' of the family structure (more co-existing generations but fewer members per generation), the retired spouses are expected to represent a big part of the future carers. Further research should investigate the particularities of these groups of carers and the consequences of caring on their other roles.

As said before, studying the consequences of caring on the working pattern of women is especially relevant in the European context of today. The Ageing challenge asks at the same time for a high women employment rate and for more informal carers to supply shortcomings or unavailability of formal care. Besides giving up work, caring can have many other consequences on working: lethargy, tiredness, lack of concentration, stress (Pickard 2004, p.7), loss of earning in the short, and in the long term (Barnes & Given 1995, p.375). These other outcomes deserve the researcher's attention too. A longitudinal approach would allow tackling the long term effects. Finally, to accurately measure the duration of care, longitudinal survey should also collect data at the start of the caring relation.

CONCLUSION

Are there research questions that cannot be answered?

The previous results suggest that 1) providing light care does not have a negative effect on work. Working women belonging to the social network of the elderly can provide them with some help without jeopardizing their working status. But politicians have not to interpret the negative effect found in the analyses as if providing light care was a protective factor from giving up work and that if more people in the community accept to contribute to elderly care, formal care could be deleted. According to the literature, a more likely explanation is that, usually, an elderly has few carers (often only one), and to keep working allows this carer to live a more balanced life and to have money to pay complementary formal cares. This implies that formal care services should be available. Otherwise, light care quickly turns in heavy care, which increases the probability of stopping work. 2) The old cohort is the most likely to be affected by caring for elderly. This cohort shows the highest proportion of carers (10%) is found. This cohort is also close to retire and the literature shows that it makes these women less 'attached' to keep working the remaining years . One hot concern in the European Union is to keep older worker at work. Providing heavy care could lead to early retirement. Therefore, older worker activation policies should take the caring dimension into account.

A multi-dimensional approach should be adopted for tackling the problem. Labour policies like care leave and flexible schedules could help women to easier allocate their time between caring and working. Social policies like carer allowances could prevent carer from the high costs linked to their caring activities (journeys, medicines, formal care, etc). Care policies like respite care[94], community care, etc. could alleviate the caring burden and prevent the carers from stress and tiredness at workplace. Alongside these guidelines, policies specific to each national situation should be implemented.

The previous analyses have provided some answer the research question. However, the research pathway has produced at least as many questions as answers: the

[94] *Respite care services* are places where nurses and social workers provide elderly with temporary care. It allows carers having a breath or fulfilling her/his other commitments.

starting point was a topic of interest (caring) and a will of approaching it from a longitudinal perspective. Small number of cases, inaccuracy in the available data asked, selection, etc. lead me to ask one question: does the availability of adequate data only determine research? Or have you first to define your research question and then try to find or raise the correspondent data? I think that the solution is in the middle: question without any data on it cannot be answered but, at the same time, investigating only those topics on which very good data are available is creating another selection bias and leading to miss the understanding of most of the social phenomenon. The story and pathway of this study are presented in *annex 20*.

BIBLIOGRAPHY

Allison, P., 1999. *Logistic Regression Using the SAS System: Theory and application.* USA: SAS Institute Inc.

Allison, P., 1995. *Survival Analysis Using the SAS System: A Practical Guide.* USA: SAS Institute Inc.

Allison, P., 1991. *Event History Analysis: Regression for Longitudinal Event Data.* London: Sage University Paper. Series: Quantitative Applications in the Social Sciences, no 46.

Allen, S., 1994. Gender Differences in Spousal Caregiving and Unmet Need for Care. *Journal of Gerontology*, *Social Sciences.* 49(4), pp. 187-195.

Allen, S., Foster, A., Berg, K., 2001. Receiving Help at Home: The Interplay of Human and Technological Assistance. *Journal of Gerontology: Social Sciences* 56(6), pp. 374-382.

Arksey, H., 2002. Combining informal care and work: supporting carers in the workplace. *Health and Social Care in the Community*, 10, pp. 151-161.

Barnes, C., Given, C., 1995. Parent Caregivers: A Comparison of Employed and Not Employed Daughters. *Social Work,* 40(3), pp. 375-381.

Beauchemin, C., Schoumaker, B., 2005. Migration to Cities in Burkina Faso: Does the Level of Development in Sending Areas Matter?. *World Development,* 33(7), pp. 1129-1152.

Becker, G., 1965. A Theory of Allocation of Time. *Economic Journa*l, 75, pp. 493-517.

Blossfeld, H. P., Rohwer, G., 2002. *Techniques of Event History Modelling: New Approaches to Causal Analysis.* (2nd ed.). New Jersey: Lawrence Erlbaum associates, Inc.

Borg, C., Hallberg, I., 2006. Life satisfaction among informal caregivers in comparison with non-caregivers. *The Authors. Journal compilation,* pp. 427-438.

Chappell, N., Blandford, A., 1991. Informal and Formal Care: Exploring the Complementarities. *Ageing and Society.* 11, pp. 299-317.

Crespo, L., 2006. Parental Caregiving and Employment Status of European Midlife Women [Preliminary Version] ftp://ftp.cemfi.es/pdf/papers/pew/Laura%20Crespo.pdf [5/17/07]

Crets, S., 1996. Determinants of the Use of Ambulant Social Care by the Elderly. *Social Sciences and Medicine*, 43(2), pp. 1709-1720.

Daily, M., Lewis, J., 2000. The concept of social care and the analysis of contemporary welfare States. *British Journal of Sociology* 51(2), pp. 28-198.

Dautzenberg M., Diederiks., J., Philipsen, H., Stevens F., Tan, F., Vernooij-Dassen, M., 2000. The competing demands of paid work and parent care. *Research on Aging,* 22, pp. 165-187.

Dentinger, E., Clarkberg, M., 2002. Informal Caregiving and Retirement Timing Among Men and Women: Gender and Caregiving Relationships in Late Midlife. *Journal of Family Issues,* 23, pp. 857-879.

Dooghe, G., 1992. Informal caregivers of elderly people: An European review. *Ageing and Society*, 12, pp. 369-379.

EC, 2002. *The life of women & men in Europe. A statistical portrait* (Data 1980–2000). http://ec.europa.eu/employment_social/employment_analysis/gender/life_women_and_men.pdf [2/11/07]

Economic Policy Committee 2001. *Budgetary challenges posed by ageing populations: the impact on public spending on pensions, health and long term care for the elderly and possible indicators of the long term sustainability of public finances* EPC/ECFIN/655/01 EN final.

Eurostat, 2003. *ECHP UDB manual: European Community Household Panel Longitudinal Users'Database. Waves 1 to 8 Survey years 1994 to 2001.*

Eurostat, 2007. *Women employment rate (16-64 years) by country (1995-2006).* http://www.eustat.es/elem/ele0002500/tbl0002546_i.htlm [5/28/07]

Egidi, V., Spizzichino, D., 2006. Perceived health and mortality: A multidimensional analysis of ECHP, Italian data. *Genus* LXII(34), pp. 135-153.

Esping-Andersen, G., 1990. *The Three Worlds of the Welfare Capitalism.* New Jersey: Princeton University Press.

Esping-Andersen, G., 2002. A New Gender Contract. In: G. Esping-Andersen et al. (eds.) *Why do we need a New Welfare State.* Oxford University Press, pp. 69-95 .

Ettner, S., 1995. The Impact of Parent Care on Women Labour Supply Decisions. *Demography*, 32(1), pp. 63-80.

Eurofound (2005), *Working time developments – 2004.* http://eurofound.europa.eu/eiro/2005/03/update/tn0503104u.html [3/17/07]

Fisher, K., Fouarge, D., Muffels, R., Verma, V., 1996. *Examining flexible labour in Europe the first three waves of the ECHP.* [Working paper]. University of Tilburg.

Gallo, F., Mastrovita, S., Siciliani, I., s.d. *The Nature of Sample Attrition in the ECHP*. (Istat: Italian National Institute of Statistics. ECHP's NDU)

Heckman, J., 1979. Sample Selection as a Specification Error. *Econometrica* 45, pp. 153-61.

Hemerick, A., 2002. The Self-Transformation of the European Social Model(s). In G. Esping-Andersen et al. (eds.). *Why do we need a New Welfare State.* Oxford: Oxford University Press. pp. 173-213.

Henz, U., 2006. Informal Caregiving at Working Age: Effects of Job Characteristics and Family Configuration. *Journal of Marriage and Family,* 68, pp. 411-429.

Henz, U., 2004. The effects of informal care on paid work participation in Great Britain: a Lifecourse perspective. *Ageing and Society*, 24(6), pp. 851-880

Hunt, C., 2003. Concepts in Caregiver Research. *Journal of Nursing Scholarship,* 35(1), pp. 27-32.

Jaumotte, F., 2003. *Women labour force participation: past trends and main determinants in OECD countries*, OECD. Economics Department Working Paper No. 376, OECD, Paris.

Jenson, J., Jacobzone, S., 2000. Care Allowance for the frail Elderly and their Impact on Women Care Givers. *Labour Market and Social Policy Occasional Papers* n° 41. http://www.cprn.com/documents/24968_en.pdf [5/17/07]

Johnson, R., Lo Sasso, A., 2000. The trade-off between hours of paid employment and time assistance to elderly parents at midlife (Working Paper). Washington, DC: Urban Institute. In line http://www.urban.org/uploadedPDF/elderly_parents.pdf [5/5/07]

Kemp, A., 2005. *Evaluating the Effectiveness of National Policies Supporting Family Caregivers of Old Adults.* Thesis submitted to the Faculty of the Virginia Polytechnic Institute and State University in partial fulfilment of the requirements for the degree of

Master of Science in Human Development. Approved: Michelle L. Stevenson, Chair, Rosemary Blieszner, Tammy L. Henderson. Blacksburg, Virginia.

Kalton, G., Citro, C., 2000. Panel surveys: Adding a fourth dimension. In: D., Rose (ed.). *Researching social and economic change: The uses of household panel studies.* London: Routledge, pp. 36-57.

Kroger, T., 2001. *Comparative research on Social Care. The state of the art*. SOCCARE Project Report 1. European Commission, Brussels.

Kroger, T., 1999. *New kinds of families, new kinds of social care.* SOCCARE Project: Families, Work, and Social Care in Europe. Final report.

Leroy, X., Neirynck, I., 1992. Solidarités *sociales et santé des aînés : Les aînés vivant à domicile (Tome 1).* Bruxelles : Rapport de la recherche demandée par le Collège Réuni de la Commission Commune de Bruxelles-Capitale à l'initiative des Ministres J. L. Thys et J. Chabert.

Marín, D., Gómez, P., Nicolás, Á ., 2006. *Labour and income effects of caregiving across Europe: an evaluation using matching techniques* [Very preliminary]

Mestheneos, E., Triantafillou, J., 2005. *PanEuropean Background Report PEUBARE.* http://www.uke.unihamburg.de/extern/eurofamcare/publikationen.php [3/17/07]

Meuleman B., Reeskens T. & Billiet J., 2007. How many countries are needed for multilevel SEM? A simulation study. *Sixth International Amsterdam Conference on Multilevel Analysis* (16-17 April 2007), Amsterdam.

Miller, B. and Mc Fall, S., 1991. Stability and Change in the Informal Task Support Network of Frail Older Persons. *The Gerontologist,* 31(6), pp. 735-745.

Moen, P., Robison, J., Fields, V., 1994. Women's work and caregiving roles: A life Course approach. *Journal of Gerontology Social Sciences,* 49(4), pp. 176-186.

OECD, 2004. *Employment Outlook* http://www.oecd.org/dataoecd/42/55/32494755.pdf [2/19/ 07]

OECD, 2005. *Society at a Glance: OECD Social Indicators-2006 edition*. OECD Publishing.

Orloff, A., 2002. Women's Employment and Welfare Regimes Globalization, Export Orientation and Social Policy in Europe and North America. *Social Policy and Development Programme,* Paper No 12 (United Nations Research Institute for Social Development).

Pacolet, J. et al., 2000. *Social Protection for Dependency in Old Age : A study of the fifteen EU Member States and Norway.* Aldershot: Ashgate Publishing.

Pickard, L., 2004. *Caring for Older People and employment: A review on the literature prepared for the Audit Commission[UK].* http://www.auditcommission.gov.uk/Products/NATIONALREPORT/4BDD4E6361C64430A08A0174D7A6880/LitReview01final.pdf [17/5/ 08]

Rose, D., 2000. Household panel studies: An Overview. *In* D., Rose, ed. *Researching social and economic change: The uses of household panel studies.* London: Routledge.

Rozario, P., Morrow-Howell, N., Hinterlong, J., 2004. Role Enhancement or Role Strain: Assessing the Impact of Multiple Productive Roles on Older Caregiver Well-Being. *Research on Aging,* 26(4), pp. 413-428.

Sainsbury, D. (eds.) 1999, *Gender and Welfare State Regimes.* Oxford: Oxford University Press.

Sarasa, S., 2005. Women's Employment and the Adult Caring Burden. *DemoSoc Working Paper* 20057, Barcelona.

Sarasa, S., 2005b. How do Midlife Women Manage Adult Care and Labour in Different Welfare Regimes? *ESPAnet Conference 2005,* September 22-24, University of Fribourg, Switzerland.

Sarkisian, N., Gerstel, N., 2004. Explaining the Gender Gap in Help to Parents: The Importance of Employment. *Journal of Marriage and Family*, 66, pp. 431-451.

Singer, J., Willett, J., 2003. *Applied Longitudinal Data Analysis: Modelling Change and Event Occurrence.* Oxford: Oxford University Press.

Spiess, K., Schneider, U., 2002. Midlife caregiving and employment: an analysis of adjustments in work hours and informal care for women employees in Europe. *Policy implications for Europe* ENEPRI Working paper no. 9.

Stier, H., Lewin-Epstein, N. et al., 2001. Welfare Regimes, Family Supportive Policies, and Women's Employment along the Lifecourse. *American Journal of Sociology,* 106(6), pp. 160-173.

Taris, T., 2000. *A Primer in Longitudinal Data Analysis.* London: Sage publications Inc.

Tennstedt, S. et al., 1993. Determining the Pattern of Community care : Is Coresidence More Important Than Caregiver Relationship?. *Journal of Gerontology: Social Sciences.* 48(2), pp.74-83.

UCLA Academic Technology Services, Stat Computing, SAS, textbook Examples: Applied Longitudinal Data http://www.ats.ucla.edu/stat/sas/examples/alda/chapter10/aldasasch10.htm[5/17/07] http://www.ats.ucla.edu/stat/sas/examples/alda/chapter11/chapter11.htm[5/17/07] http://www.ats.ucla.edu/stat/sas/examples/alda/chapter12/chapter12.htm [5/17/07]

United Nations 2005. *World population Prospects. The 2004 revision*, Vol.1: Comprehensive tables, New York, United Nations.

Van Tilburg, T., 1998. Losing and Gaining in Old age: Changes in Personal Network Size and Social Support in a Four Year Longitudinal Study. *Journal of Gerontology: Social Sciences.* 53(6), pp. 313-323.

Vermunt, J., 1997. *LogLinear Models Event Histories.* Advanced Quantitative Techniques in the Social Sciences Series, 8, London, Sage Publications.

Viitanen, T., 2005. *Informal Elderly Care and Women Labour Force Participation across Europe.* ENEPRI Research Report n°13 http://www.enepri.org/Reviser/enepritarja.pdf [5/17/07]

Wuest, J., 1997. Illuminating environmental influences on women's caring. *Journal of Advanced Nursing,* 26, pp. 49-58.

Annex 1:

Ageing population

Table 3: Population aged 80+ (%)

Country	1995	2050
Belgium	3.8	10.8
Denmark	3.9	8.4
France	4.3	10.9
Germany	4.1	12.2
Greece	3.2	9.5
Ireland	2.5	7.6
Italy	4.1	15.2
Netherlands	3.1	10.1
Portugal	3.0	9.9
Spain	3.2	12.3
United Kingdom	3.9	8.8

Source: United Nations 2005. *World population Prospects. The 2004 revision*, Vol.1: Comprehensive tables, New York, United Nations

Annex 2:

Literature review methodology

Papers and documents used for the literature review were identified using:

1) Google with various combinations of the following keywords: "caregiver/carer", "elderly/old", "employment/work", "women" "Europe".

2) ERIC (EBSCO) [a free database on the net] for papers having "caregiver/carer" in the title and "employment/work" and "elderly/old" somewhere in the text.

3) Starting with these first readings, the keywords list have been lengthened with synonyms or related terms. The new list has been used for a second research in the **BLACKWELL** and **J-STOR** numerical databases.

4) The **bibliography** of the papers read was also used to identify important readings to make, but all of them could not be found.

5) Finally, for avoiding missing less frequently quoted papers (nor in the bibliographies neither on the net), 2 journals specialized in aging issues were systematically reviewed from 1990 onwards: ***Aging & Society*** and ***Research on Ageing.***

Exclusion criteria: papers dealing non developed countries[1], with mentally ill elderly (too specific case), with formal support for carers and too ancient papers, although exceptions were made for reference papers.

Efforts were made to systematize the research of the literature. If non-availability and time pressure did not allow reading all the identified papers, I had the feeling to have reached a saturation point when reading the last papers as they did not give much more information than the previous ones.

[1] I did not restrict the literature review to the European situation because a great deal researches made on the caring-working relationship come from US.

Annex 3:
A selection of research results on the caring-working relationship

Table 4: Research on the caring-working relationship

Authors	Description	Research question	Control	Findings
Crespo (2006)	**Period** : 2004 **Country:** 2 groups of countries: Southern European countries (Spain, Italy and Greece) and Northern European countries (Sweden, Denmark and Netherlands) **Method:** bivariate probit analysis **Survey:** SHARE **Sample:** women approaching retirement (i.e. aged 50 to 60) in Europe	Whether there actually exists a significant opportunity cost of undertaking intensive parental caring activities in terms of employment status for European mid-life women. Analysis of the potential relationship between labour market participation decisions and caring for elderly parents.		- Women who provide intensive care, either because they are live-in carers or because they have to look after someone outside the household on a daily basis, have a lower probability of being in employment, both in southern and in Northern European countries.
Dentinger & Clarkberg (2002)	**Period** : 1994-1995 **Country:** USA **Method:** Discrete time Event history analysis **Survey:** Cornell Retirement and Well-being Study **Sample:** 763 pension-eligible men and women	The ways in which informal caring influences the transition to retirement and how this relationship is shaped by gender.		- Certain types of caring shape the timing of retirement and is moderated by gender (e.g. wives caring for their husbands have retirement odds 5 times greater than women who are not carers, whereas husbands caring for their wives are substantially slower to retire.

Source: Self-elaboration from readings

Table 4: Research on the caring-working relationship (following)

Authors	Description	Research question	Control	Findings
Heitmueller & Michaud (2006) [in Marin et al, 2006]	**Period**: 1991-2003 **Country:** UK **Method:** bivariate probit adjusting both for state dependence and unobserved heterogeneity **Survey:** 13 waves of the British Household Panel **Sample:** n.a.	Effect of providing care on labour supply for all carers regardless of whether they look after someone living in the same household or living elsewhere.	Tackling simultaneity and unobserved individual heterogeneity.	- Labour supply is not statistically different than that of non carers. - The probability of being employed is lower for women (-6%) and men (-4.7%) when they look after a live-in dependent.
Henz (2004)	**Period**: 1994-1995 **Country:** GB **Method:** Probit models in a dynamic Lifecourse approach **Survey:** Family and Working Live Survey **Sample:** representative sample of men and women aged 15 to 69.	▪ What is the timing of informal caring to a sick, disable, or elderly person? ▪ What are the effects of caring on employment?	Age, family configuration, employment characteristics	- Most carers look after only one dependent person during theirs lives, only 1/3 look after a second dependent before the age of 65 years. - Among carers, 1/3 were never employed before caring, 1/3 said that caring has no effect on their work arrangements, 1/3 reports most commonly that they stopped working.
Johnson & Lo Sasso (2000)	**Period** : 1993-1994 **Country:** US **Method:** a simultaneous equation model **Survey:** 2 waves from the Health and Retirement Study –HRS **Sample:** men and women aged 53 to 65 with at least one living parent	To estimate the impact of looking after an elderly parent for more than 100 hours in a given year on the number of working hours within the same year.	Tackling simultaneity and unobserved individual heterogeneity	- The labour supply for both men and women in such situation is 23% and 28% lower than the labour supply of middle age men and women (53-63 years) who do not provide care.

Source: Self-elaboration from readings

Table 4: Research on the caring-working relationship (following)

Authors	Description	Research question	Control	Findings
Marín,Gómez & Nicolás (2006)	**Period** : 1994-2001 **Country:** 12 EU countries **Method:** Match women who have become carers with "control" women who are deemed to be comparable in all relevant characteristics and compute a non-parametric measure of the effect of becoming a carer on the outcomes mentioned above. **Survey:** ECHP **Sample:** women aged 30 to 60	[What is]the effect of caring on labour outcomes such as employment, full time employment (conditional on employment), and income for different European countries.	Age and gender, educational attainment, health status, household size, number of children in the household, marital status, job characteristics (only when we analyze working women), the logarithm of equivalent household income at the start of the sequence and country and wave interactions.	- For women who are working before becoming a carer there is no statistically significant change in the chances of being employed. - For women who were not working prior to becoming a carer, there is a statistically significant decrease in the chances of entering work. - Negative and significant effect on labour income, which tends to be offset by a parallel increase in social transfers, except in the case of women with low levels of education in the Southern countries.
Pavalko & Artis [Spiess & Schneider (2002) and in Henz, (2004)]	**Period** :1984 and 1987 waves **Country:** US **Method:** n.a. **Survey:** National Longitudinal Survey (NLS) **Sample:** Women aged 50 to 64 caring for husbands, ill or disabled children, parents, or grandchildren.	- Relationship between employment and caring for an ill or disabled friend or relative for 3 years. - Changes in working-time associated with the end of caring.	Demographic controls, caring status (in the employment estimation), employment variables (status, wages, satisfaction), and work history variables.	- Working women and women with a higher work status are as likely to start caring as other in 1987. - Job tenure displays a positive, though marginal, effect on the odds of caring, which supposedly reflects better chances of workplace flexibility. - Impact of caring on working-time changes depends on whether caring is being taken up or terminated: starting caring adversely affects working-time, stopping care provision is not associated with working-time.

Source: Self-elaboration from readings

Table 4: Research on the caring-working relationship (following)

Authors	Description	Research question	Control	Findings
Sarkisian & Gerstel (2004)	**Period** : na **Country:** UK **Method:** zero-inflated negative binomial regression. **Survey:** National Survey of Families and Households **Sample:** 7350 men and women	To what extent can the gender gap be explained by structural variation, especially the different rates of employment and kinds of job that women and men tend to hold?	age, race, education, and physical health; marital status and number of minor children; and proximity to parents, parents' health-related need, parents' financial need, parents' gender, parents' marital status, number of siblings and siblings-in-law family.	- Employment status and job characteristics, especially wages and self-employment, are important factors in explaining the gender gap in the help given to parents, and that these operate similarly for women and men.
Spiess & Schneider (2002)	**Period** : 1994-2001 **Countries:** 12 EU countries **Method:** maximum-likelihood probit estimation and ordinary-least square regressions. "Difference-in-difference model" **Survey:** ECHP **Sample:** women aged 45 to 59	association between changes in weekly working hours and changes in weekly caring hours	*Individual characteristics & circumstances at the time of the first interview in 1994* (age, education, nationality, health, employment status, family status & household type). *Macrolevel variables:* ratio of the population 65+ to midlife women, the female unemployment rate, *variables capturing the influence of long-term care policy*, country-group dummies.	- Starting or increasing informal caring is normally associated with a reduction in the number of working hours per week. - There does not appear a positive effect on working hours for women terminating a caring spell or reducing caring hours. - The (negative) link between start of caring and a working time reduction seems to be particularly strong in Northern European countries where formalised home care is more frequent and where re-entry into the labour market may be easier.

Source: Self-elaboration from readings

Table 4: Research on the caring-working relationship (end)

Authors	Description	Research question	Control	Findings
Spiess & Schneider (2002) [in Marin et al, 2006]	**Period** : 1992-1994 **Countries:** 12 EU countries **Method:** difference-in-differences strategy on 2 groups of European countries: Mediterranean (in this group they also include Ireland), and non-Mediterranean countries **Survey:** ECHP **Sample:** women aged 45–59 years who participated in the labour force in at least one of the 2 years studied.	Changes in the number of working hours are caused by the 3 following events: to start caring, to continue caring and to stop caring.		-In Mediterranean countries the number of working hours is affected only when women continue giving care, while for other countries the effects are found when they start caring.
Viitanen (2005)	**Period** : 1994-2001 **Countries:** 13 European countries **Method:** dynamic probit models **Survey:** ECHP **Sample:** women aged 20 to 59	Effect of informal caring on labour behaviour	Control for unobserved individual heterogeneity, state dependence and attrition bias.	- Country specific and show that the probability of being in employment is affected by being an informal carer only in the case of Germany. - When she separates the analysis for different subgroups of the population, she finds statistically significant effects for middle aged women (Belgium, Finland, and Germany) and single women (Greece, Netherlands, Italy, and Germany) in different countries.

Source: Self-elaboration from readings

Annex 4:

Women employment rate

Table 5: Women employment rate (16-64 years)

Country	1995	2000
Belgium	45.0	51.5
Denmark	**66.7**	**71.6**
France	52.1	55.2
Germany	55.3	58.1
Greece	38.1	41.7
Ireland	41.6	53.9
Italy	35.4	39.6
Netherlands	**53.8**	**63.5**
Portugal	**54.4**	**60.5**
Spain	31.7	41.3
United Kingdom	**61.7**	**64.7**

Source: Eurostat 2007

Annex 5:

ECHP waves and their use in the analysis

Table 6: Use of ECHP waves

ECHP Waves	**Use for the analyses**
1994	- To impute missing data on following waves - To compare the characteristics of the study sample to the part of the women who participate to the beginning of the panel in 1994 and not during the next year (1995).
1995	- Selection of the study sample and comparison with women excluded from the analysis due to their working status in 1995. - Beginning of the time (period 0) - Lagged predictors for period 1 - Information on the part of the study sample that is censored in 1996
1996-2001	- Periods 1-6 in the person period file during which an event can occur = periods introduced in the person period file.

Source: Self-elaboration

Annex 6:

From the individual file to the person-period file

individual file

id	censor	w_event	N_period	Age 95(ti)	Care_old95	care_old96	care_old97	care_old98	care_old99	care_old00	care_old01
ID1	*0*	3	6	20	1	1	1	0	0	0	0
ID2	3	0	1	44	0	0	.	.	.	.	.
ID3	8	0	6	34	0	0	0	0	1	1	1

…

Person-period file

INDIVIDUAL	Period	Event	censor	Age 1995(ti)	Caring old(tvl)
1	1	**1**	0	20	1
2	1	0	**1**	44	0
3	1	0	0	34	0
3	2	0	0	34	0
3	3	0	0	34	0
3	4	0	0	34	0
3	5	0	0	34	1
3	6	0	**1**	34	1

id: individual number (fictive here)

censor: first time when the women is not observed (right-censor)

n_period: number of periods at risk

'.' Means that there is missing data

(ti): time-invarying predictor

(tvl): time-varying predictors lagged by one period

w_event: wave was the event is reported. In the case of ID1 it is reported at wave 3 (1996) and occurred between waves 2 and 3, which is the person-period

The creation of a person period file from an individual file consists in switching from a file with the individual as observation unit to a file with the person period as observation unit: each woman has a number of observations corresponding to the number of periods for which she is still at risk of experiencing the event. This implies that each of the 24592 women of the study sample contribute at least to one observation. All of them were working at the moment of the interview in 1995. Therefore they are all at risk of quitting work during the period 1995-96 (until the time of the next interview). Women who either experiment the event during this first person period either do not answer in 1996 (censored) contributes to only 1 person period. The maximum *person-period*[2] a woman can have is 6 (if she stays in the panel and is still working until the 2000). The person-period file contains different elements: (1) the *ID* for identifying the individual; (2) the *time indicator* (the period) which identifies the specific moment when

[2] The term of person period has been chosen for making the distinction between the observations in the person period file (when the woman is at risk) and the more general term of period (time-interval).

information was recorded ; (3) the *outcome variable* [event]; (4) predictors variables with either *time-invariant predictors* which have identical values across each of the person's multiple records [here: the cohorts built on age in 1995 1: less than 31; 2: 31 to 49; 3: 50+] or *time-varying predictors* which can vary over time [in our case: all the predictors, except the cohort, are time-varying.] (Singer & Willett 2003)

The predictor values of 1995 are used for predicting the value of the outcome at the end of the first person period (1995-96). The predictor values of 2000 are used for the 7th and last person period. Thus, choosing '1-year-lagged' specification means that we will not make use of the predictors of the last period (2000-01) and that we need to have predictors for the wave (wave 1995) before the first person period (1995-1996).

Annex 7:

Illustration of cases misidentified due to time-interval bias

Table 7: Illustration of misidentified cases

1995	Jan –Aug 96	Sept-Dec 96	If the interview occur in November 1996
Working	*Caring* No working	No caring No working	Misspecification of caring
Working	*No Caring* No working	Caring No working	
Working	Caring *No working*	Caring Working	Event not identified
Working	No Caring *No working*	No Caring Working	
Working	Caring *No working*	No caring Working	Event not identified and misspecification of caring
Working	No Caring *No working*	Caring Working	
In all those cases, the woman is at risk of stopping working in 1996 because she was working until the end of 1995. When the event occurs before the moment of the next interview, and if a change occurs in the characteristic between the moment of the event and the interview or if she works again, misidentification can occur.			

Source: Self-elaboration

Annex 8:

Processing the study sample and selection issue

Two solutions for correcting the selection bias (competing risks and Heckman's correction) were investigated and then rejected because they are not applicable to the data or to the nature of the event. **Competing risks** should solve the problem of selectivity by keeping all the women as study sample (either working or not). However, the event would have to be redefined as an increase or a decrease in working-time during the period under study. The problem is that this method requires that all the women are at risk of experiencing the 2 events [working more hours and working less hours] (Singer and Willett 2003). But, a woman who is not working *at the beginning of the considered time* cannot experiment the event of working less. So, the method is not applicable in this case. **Heckman's two-step selection procedure (1979)** allows correcting the selectivity bias only if an instrumental variable explaining the selection process can be identified. The next table depicts the selection steps and shows the different reasons of exclusion from the study sample.

Table 8: Selection process of the study sample

<table>
<tr><td colspan="4">All the women in wave 1
↓
Was the woman participating in wave 2(1995)? [3]</td></tr>
<tr><td rowspan="2">No: attrition</td><td colspan="3">Yes
↓
What did she answer to the question on working status (PE003):
ILO main activity status at the time of interview?</td></tr>
<tr><td>No answer</td><td>Answered to be inactive
(unemployed, economically inactive, discouraged worker)</td><td>Answered to be active
↓
Study sample[4]</td></tr>
</table>

Source: Self-elaboration

[3] I have decided to work only on the women participating to the panel since the beginning in 1994 but only to consider the waves 1995-2001 for the analyses. The 2 main reasons of this choice are: first, there is an important question on caring that is not asked at the first wave in 1994. The other reason is that, restricting the sample to the women in the panel since the beginning allows using the information of the first wave for imputation.

[4] If a woman does not answer the working status question before having stopped working or having left the panel, she will be considered as right censored the first wave where she does not answer. For instance, a sample woman who claims working in 1995, 1996 and 1997 and who does not answer the question in 1998 will be censored in 1998. Missing values on working status is not a problem if it occurs after the event.

It would be very difficult to find one substantive variable that could be explanative of being excluded for any of the 3 reasons mentioned above, rather than being included in the study sample. Also, applying the correction with a wrong variable could even worsen the situation rather than solving the selection bias problem. I have decided not to use Heckman's correction in this essay. This implies that the scope of the findings is limited to investigate whether the ECHP women working in 1995 are more or less likely to exit the labour market at least once until 2001.

Annex 9:

The decision to analyse the 11 countries together

At the really beginning of this project, I have planned to compare 4 countries: Denmark, The Netherlands, Italy and Ireland (1 for each Welfare Regime, according to Esping-Andersen's typology (1990).

After the lecture on the event history analysis and the seminar which taught us the big model technique for comparing countries using country dummies, I have decided to restrict the analysis to 1 country[5]: the Netherlands was chosen, because of its sample size (bigger than the Irish and the Danish) and of the quality of its data (less missing than for Italy). I was also interested in a country with a 'hybrid' model (sometimes linked to the Nordic Welfare Regime). But, when I did start with the exploratory analyses, the number of observations of being a carer at wave n and having the event at wave n+1 was too small for the kind of method I have planned to use.

Thus, I decided to work on the 11 countries participating to the ECHP panel in 1994 and asking the question on caring (my main independent variable). Among these countries, only 1 (Denmark) represents the Nordic Welfare Regime type, while the 4 Mediterranean countries are present. Because of this disequilibrium, I though that a Welfare Regime approach would not be appropriated. I rather have planned to introduce variables taking into account the diversity of the national institutional settings, like the development of community care or the availability of residential care. But, finally, they were not included because it seems difficult to define comparable and relevant macro characteristics for 3 reasons: (i) the definition of the services for elderly varies from country to country; (ii) papers and reports that try to compare the national settings are contradictory and the date of introduction of those services, their coverage, and the part of beneficiary contribution are not always specified; (iii) finally, availability of services in one country does not imply that the elderly or their carers will make use of them[6].

---

[5] For making a real comparison of the effect of the characteristics in the different countries, I should have put country dummies and their interaction with my variables of interest. This would have increased substantially the number of predictors in the analysis. I was afraid of having too few cases for some of the terms included in the model.

[6] The multi-level analysis has been rejected due to small number of countries participating to the panel. According to Meuleman et al. (2007, p.21), multi-level requires a minimum of 40 countries to be carried out validly.

Thus the analyses concern the situation of a sort of 'average' European working women. While this choice implies that we lose in understanding the diversity of the different countries, we can consider that the working-caring dilemma is similarly present in all the study countries and that it still makes sense to deal with the situation of the European women.

Annex 10:

The study sample and the removed sub-samples: a comparison of their socio-demographic characteristics

For the 'not observed in 1995' sub-sample, we used the characteristics of the 1994 wave. For the others sub-samples, the 1995 characteristics are considered, the working status recorded this year served as a basis to select the study sample.

Table 9: Selection and countries sample

Countries	**Total**	**Not observed w2/1995**		**Working status not answered**		**Not working**		**Study sample**	
	N	**N**	**% lost**	**N**	**% lost**	**N**	**% lost**	**N**	**% kept**
All	66195	6705	10.13	77	0.12	34821	52.60	24592	37.15
Denmark	3053	420	13.76	1	0.03	1098	35.96	1534	**50.25**
Netherlands	4953	431	8.70	1	0.02	2312	46.68	2209	44.60
Belgium	3507	321	9.15	1	0.03	1739	49.59	1446	41.23
France	7493	839	11.20	9	0.12	3792	**50.61**	2853	38.08
Ireland	4982	998	**20.03**	0	0.00	2477	49.72	1507	30.25
Italy	9070	543	5.99	0	0.00	5746	**63.35**	2781	30.66
Greece	6590	658	9.98	0	0.00	4019	**60.99**	1913	29.03
Spain	9276	1332	14.36	0	0.00	5789	**62.41**	2155	23.23
Portugal	6064	327	5.39	1	0.02	3210	**52.94**	2526	41.66
Germany	6347	414	6.52	**64**	1.01	2638	41.56	3231	**50.91**
UK	4860	422	8.68	0	0.00	2001	41.17	2437	**50.14**

Source: ECHP 1994-2001

Table 10: % working women at ECHP 1995 and the Eurostat employment rate 1995

	ECHP 1995		ECHP 1994	Eurostat figures 1995
Country	**N**	**% working**	**% working 1994 not observed 1995**	**Female employment rate (16-64 years)**
Denmark	2632	58.3	42.7	66.7
Netherlands	4521	48.9	42.0	53.8
Belgium	3185	45.4	40.6	45.0
France	6645	42.9	34.0	52.1
Ireland	3984	37.8	**44.2**	41.6
Italy	8527	32.6	**33.7**	35.4
Greece	5932	32.2	28.4	38.1
Spain	7944	27.1	**28.7**	31.7
Portugal	5736	44.0	36.6	54.4
Germany	5869	55.1	49.8	55.3
UK	4438	54.9	50.9	61.7

Source: ECHP 1994-2001, EUSTAT 2006

Table 11: Socio-demographic characteristics[7]

	Not working 1995		Study sample		Not observed 1st wave	
Median age	51.98		***38.65***		45.00	
	N	**%**[8]	**N**	**%**	**N**	**%**
Bad health	16511	**47.40**	6560	***26.70***	2593	**38.70**
Support	28315	81.30	21784	88.60	5593	83.40
Primary education	22627	**68.93**	9489	***39.81***	3566	**56.03**
Secondary education	7966	24.27	8701	36.51	1853	29.12
Further education	2233	***6.80***	5643	**23.68**	945	***14.85***
Old Care	3048	8.75	1506	6.12	501	7.49
N						

Source: ECHP 1994-2001

[7] The 77 women who were removed because they did not answer are not considered.

[8]The percentages are computed on the valid observations after imputation for the study sample. For old care, there are 2 missing values among 'not working', 4 for 'study sample' and 17 for 'first wave'. For education there are 1995 missing values among 'not working' (5.7%), 759 for 'study sample' (3.1%) and 341 for 'first wave' (5.1%).

Table 12: Working and not working carers in 1995

Carers in 1995		
% cohort 1- young	3%	
% cohort2 – mid	8%	
% cohort3 – old	10%	
	Working	**Not working**
N carers in the 3 cohorts	**1506**	**3048**
Median age	45	56
% carers in the 3 cohorts	6	9
% secondary education + (valid)	0.49	0.28
% caring 2h+ /day	3.2	6.3
Total cohort 1	6984	6816
N carers cohort 1	163	226
% in cohort 1	**2**	**3**
Total cohort 2	12919	8260
N carers cohort 2	837	864
% in cohort 2	**6**	**10**
Total cohort 3	4689	19745
N carers cohort 3	506	1958
% in cohort 3	**10**	**10**

Source: ECHP 1994-2001

Annex 11:

Attrition table and characteristics

Table 13: Attrition

Attrition table				
Wave	**Study sample**	**Wave attrition**	**Cumulative attrition**	**% Staying**
1995	24592	0	0	100
1996	23188	1404	1404	94.29
1997	21583	1605	3009	87.76
1998	20157	1426	4435	81.97
1999	18877	1280	5715	76.76
2000	17555	1322	7037	71.39
2001	15342	2213	9250	62.39

Source: ECHP 1994-2001

Table 14: Comparison attritors/non-attritors[9]

	Attritor (n=9250)		**Non-attritor** (n=15342)	
	N	**%**	**n**	**%**
Stop work (1996-2001)	5270	**34.35**	1647	**17.81**
Predictor values in 1995				
BADHEALTH	2249	24.31	4311	28.10
SUPPORT	8152	88.13	13632	88.85
SECOND	3543	39.83	5158	34.53
SUPER	2115	23.78	3528	23.62
OLDCARE	493	5.33	1013	6.60
LIGHT	231	2.50	486	3.17
HEAVY	262	2.83	527	3.44
FIRST	246	2.66	464	3.02
SEC	247	2.67	549	3.58
CHILD ONLY	3035	32.81	5597	36.48
OLD CARE ONLY	289	3.12	605	3.94
CHILD &OLD	204	2.21	408	2.66
PART-TIME WORK	2227	24.30	4019	26.48
BLUE COLLAR	3393	38.39	6329	43.43
SELF-EMPLOYED	1431	15.47	2586	16.86
PUBLIC SECTOR	2623	28.36	4709	30.69

Source: ECHP 1994-2001

[9] Comparisons refer to 1995, the only year where information is available for the all study sample. The proportions are computed on the valid observations separately for attritors and non-attritors.

Annex 12:

Description of the independent variables used in the analyses

Table 15: Description of the independent variables

Variables	ECHP question	% individuals without missing[10]	Dummies for the regression analysis	Specification
Time	*Wave*	100	-	-
Cohort	*Age at the time of the survey in 1994*	100	*COH1*='1' if age 1995 up to 30 *COH2*='1' if age 31-49 *COH3*='1' if age 50+ (Ref. cat. in the cohort comparison models is the *COH1*)	Time-invarying (1995 value)
Old care	*Do your present daily activities include, without pay, looking after children or other persons who need help because of old age?*	99.44	*OLDCARE*=1 if the respondent is caring for elderly (and child) (Ref. cat.: no care for elderly)	Time-varying lagged
Multi-roles	*Do your present daily activities include, without pay, looking after children or other persons who need special help because of old age?*	99.44	*ONLY CHILD=1* if the respondent is only caring for children *ONLY OLD=1* if the respondent is only caring for elderly *CHILD & OLD=1* if the respondent is caring for children and elderly (Ref. cat.: no care)	Time-varying lagged
Intensity of elderly care	*- Number of weekly hours spent looking after elderly* *-Co-residence with elderly.*	99.23 99.89	*HEAVY*=1 either if the respondent provides care for more than 2 hours a day or if the elderly lives at home. *LIGHT*=1 if the respondent is carer but does not fit any of these conditions (Ref. cat.: no care for elderly)	Time-varying lagged

Source: Self-elaboration

[10] Only the code '-9' (non response) until the last person-period is considered as missing. '-8' (not concerned) is not.

Table 15: Description of the independent variables (end)

Variables	ECHP question	% individuals without missing	Dummies for the regression analysis	Specification
Duration of elderly care	*Do your present daily activities include, without pay, looking after children or other persons who need special help because of old age?*	99.44	*FIRST*=1 if the respondent was caring the year before but not the previous year, *SEC*=1 if the respondent was caring the 2 preceding years (Ref. cat.: no care for elderly)	Time-varying lagged
Health	*- Health in general?*	99.28	=>*BADHEALTH*=1 the respondent evaluates her health as 'fair' or worse (Ref. cat.: (very) good)	Time-varying lagged
Education	*- Higher level of education completed*	99.09	*SECON*=1 if higher level completed is secondary school, *SUP*=1 if higher level completed is higher education (Ref. cat.: primary education)	Time-varying lagged
Social class	*Occupation in current job*	90.79	*BLCOL*=1 if manual, technicians, etc. (Ref. cat.: white collar)	Time-varying Lagged
Sector	*Status of employment*	94.80	*SELF*=1, if self-employed *PUBLIC*=1 if public sector (Ref. cat.: private sector)	Time-varying Lagged
Support	*-Number of household members aged 16+* *-Co-residence with elderly.*	100 100	*SUPPORT*=1 if number of household member aged 16+ (reduced by one if the is an elderly cohabiting) is, >= 2 (Ref. cat.: no support)	Time-varying Lagged
Part-time	*- Main job part-time/full-time? (Auto-evaluated).*	85.56	*PARTIME*=1 if the respondent works part-time (Ref. cat.: full-time)	Time-varying Lagged
Income from work as a % of household net income	*-Personal net income from work* *- Household net income*[11]	100 99.19	*PCTINC*=1 if the income from work of the woman count for at least half of the household net income. (Ref. cat.: < 50% household income)	Time-varying Lagged

Source: Self-elaboration

[11] I did not choose the wage of the women as economic predictor but rather the part her income represents in total household income. It seemed to me a better indicator of the opportunity cost of stopping work.

Annex 13:

Specification of time function

Table 16: Life table (for the all study sample)

ECHP wave	Time interval	Risk set	Events	Censored observations	Hazard function	Survival function
2	*1995*	24592	0	0	0	1
3	1995-96	24592	2548	1790	0.104	0.896
4	1996-97	20254	1530	1723	0.076	0.829
5	1997-98	17001	988	1368	0.058	0.781
6	1998-99	14645	751	1126	0.051	0.740
7	1999-00	12768	619	926	0.048	0.705
8	2000-01	11223	481	10742	0.043	0.674

Source: ECHP 1994-2001

To build the life table we need to know the number of events (***3rd column***) and of censored observations (***4th column***) for each person-period. The risk set at the beginning of the time-interval n (***2nd column***) equals the risk set of the previous period (n-1) minus the number of events and censored observations that did occur during this n-1 period. For instance, the risk set at the beginning of the time-interval 1997-1998 = 20254 – 1530 -1723 = 17001.

The hazard rate at the end of the period n (***5th column)*** equals the number of events occurring during the time-interval divided by the risk set of this time-interval. It is the conditional probability that a woman working in 1995 exits the labour market during this period given that she is still in the panel and has not yet experienced the event.

The value of the survival function (***6th column)*** at the period n is defined as the probability of being still 'at risk' of experiencing the event at the end of this period. Because there are censored observations, the survival function cannot be computed directly (Singer and Willett 2003, p.336). Under the assumption that censoring is independent, it can be computed from the hazard rate of the previous and the concerned periods. Therefore, the probability of not having yet stopped work at the end of the period 1996-97 is obtained by computing the probability of not experiencing this event (to survive it) neither in 1995-96 nor in 1996-97= (1-0,1036)*(1-0,0757).

Note that there are 2 ways of being censored at the end of the last period (2001): either the woman was still working in 2000 but did not participate to the 2001 wave (attrition) or she

answered that she was still working in 2001. In both cases, we do not observe the event before the end of the data collection.

The life table displays the event histories of the 24592 women working in 1995: at the beginning, all of them are working. One year after, 2548 exit work and 1790 had left the panel. Therefore, 20254 women enter the next year. Both event occurrence and censoring become less frequent over the time. At the beginning of the 6th (and the last) year only 11223 working women remain observable in the panel (46% of the initial sample).

The next table displays the results of the logistic regressions carried out for choosing the time specification. The general model includes all the period dummies. *ONE* is a constant. For choosing the best specification, one has to consider the AIC (the smaller the better), the parsimony (the difference in AIC between cubic and quadratic is too small for preferring the cubic specification) and the theory.

Table 17: Model fit statistics for the different time-specifications

Model	Covariates Included	AIC	_2_Log_L	Difference with the preceding model	Difference with the general model
General	period1 period2 period3 period4 period5 period6	49623.141	49611.141	.	.
Constant	*ONE*	50368.406	50366.406	.	755.264
Linear	*ONE TIME*	49670.905	49666.905	**699.500**	55.764
Quadratic	*ONE TIME TIME^2*	49621.350	49615.350	**51.555**	4.208
Cubic	*ONE TIME TIME^2 TIME^3*	49620.835	49612.835	2.515	1.694
Fourth	*ONE TIME TIME^2 TIME^3 TIME^4*	49621.168	49611.168	1.668	0.026

Source: ECHP 1994-2001

Annex 14:

Bivariate models for caring and control variables

Table 18: Bivariate models

Predictors	DF	Estimate	Standard Error	Wald Chi-Square	Pr>Chi sq	diff AIC	odds	IC 95%
OLD CARE								
Intercept	1	-2.6158	0.0129	40847.3017	<.0001			
OLDCARE	1	0.1633	0.0479	11.6287	0.0006		1.177	(1.072 -1.293)
	intercept only		50367.835	**with covariates**	50358.659	9.176		
COHORT								
Intercept	1	-3.0366	0.02	23069.5487	<.0001			
COH1	1	0.5044	0.0309	266.5201	<.0001		1.656	(1.559-1.759)
COH3	1	1.2188	0.0299	1661.1025	<.0001		3.383	(3.191-3.587)
AIC	**intercept only**		50368.406	**with covariates**	48794.216	1574.190		
BAD HEALTH								
Intercept	1	-2.7532	0.0157	30776.1627	<.0001			
BADHEALTH	1	0.4578	0.0259	312.8643	<.0001		1.581	(1.502-1.663)
AIC	**intercept only**		50368.406	**with covariates**	50069.138	299.268		
EDUCATION								
Intercept	1	-2.2674	0.0174	17049.5797	<.0001			
SECON	1	-0.4698	0.0286	269.1606	<.0001		0.625	(0.591-0.661)
SUPER	1	-0.9514	0.0367	671.088	<.0001		0.386	(0.359-0.415)
AIC	**intercept only**		49002.279	**with covariates**	48189.781	812.498		
SUPPORT								
Intercept	1	-2.5796	0.035	5444.574	<.0001			
SUPPORT	1	-0.0286	0.0374	0.5859	0.444		0.972	(0.903-1.046)
AIC	**intercept only**		50368.406	**with covariates**	50369.834	-1.428		
BLUE COLLAR								
Intercept	1	-2.712	0.0168	25924.9189	<.0001			
BLCOL	1	0.2521	0.025	101.2921	<.0001		1.287	(1.225-1.351)
AIC	**intercept only**		50368.406	**with covariates**	50269.973	98.433		
PART_TIME								
Intercept	1	-2.8092	0.0156	32631.8468	<.0001			
PARTIME	1	0.7036	0.0262	722.9486	<.0001		2.021	(1.920-2.127)
AIC	**intercept only**		50368.406	**with covariates**	49692.51	675.896		
SECTOR								
Intercept	1	-2.4737	0.0164	22681.3683	<.0001			
SELFWORK	1	0.3216	0.031	107.2639	<.0001		1.379	(1.298-1.466)
PUBLIC	1	-0.7387	0.0328	507.0365	<.0001		0.478	(0.448-0.509)
AIC	**intercept only**		50368.406	**with covariates**	49493.664	874.742		
% INCOME FROM WORK								
Intercept	1	-2.4984	0.0142	31041.7072	<.0001			
PCTINC	1	-0.5649	0.0321	310.6184	<.0001		0.568	(0.534-0.605)
AIC	**intercept only**		48786.800	**with covariates**	48445.99	340.81		

Source: ECHP 1994-2001

Annex 15:

Cohort differences

Table 19: Distribution of the control variables in the 3 cohorts[12]

	Cohort 1 (N=26414)		Cohort 2 (N=57248)		Cohort 3 (N=16821)	
Predictors	**N**	**%**	**N**	**%**	**N**	**%**
BADHEALTH	4682	**17.73**	16008	**27.96**	7700	**45.78**
SECON	10925	**43.00**	18946	**33.49**	3902	**23.49**
SUPER	6532	25.71	16397	28.99	2887	17.38
SUPPORT	23201	87.84	50971	89.04	13819	82.15
BLCOL	9676	36.63	22574	39.43	7868	46.77
PARTIME	4679	**17.71**	13500	**23.58**	5157	**30.66**
SELFWORK	2059	**7.80**	8327	**14.55**	5056	**30.06**
PUBLIC	6595	24.97	21671	37.85	5066	30.12
PCTINC	6819	26.09	16515	29.05	5043	31.07

Source: ECHP 1994-2001

Table 20: Life table: young cohort (up to 30 years)

Time	Risk set	N of Events	N of Censored	Hazard	Survival
1995	6984	0	0	**0**	1
1995-96	6984	818	690	**0.117**	0.883
1996-97	5476	471	626	**0.086**	0.807
1997-98	4379	261	444	**0.060**	0.759
1998-99	3674	167	362	**0.045**	0.724
1999-00	3145	133	256	**0.042**	0.694
2000-01	2756	95	2661	**0.034**	**0.670**

Source: ECHP 1994-2001

Table 21: Life table: mid cohort (31-49)

Time	Risk set	N of Events	N of Censored	Hazard	Survival
1995	12919	0	0	**0**	1
1995-96	12919	959	799	**0.074**	0.926
1996-97	11161	547	831	**0.049**	0.880
1997-98	9783	359	732	**0.037**	0.848
1998-99	8692	310	629	**0.036**	0.818
1999-00	7753	244	569	**0.031**	0.792
2000-01	6940	203	6737	**0.029**	**0.769**

Source: ECHP 1994-2001

[12] The effectives are in number of p-periods. The percentages are computed on the valid observations.

Table 22: Life table: old cohort (50+)

Time	Risk set	N of Events	N of Censored	Hazard	Survival
1995	4689	0	0	**0**	1
1995-96	4689	771	301	**0.164**	0.836
1996-97	3617	512	266	**0.142**	0.717
1997-98	2839	368	192	**0.130**	0.624
1998-99	2279	274	135	**0.120**	0.549
1999-00	1870	242	101	**0.129**	0.478
2000-01	1527	183	1344	**0.120**	**0.421**

Source: ECHP 1994-2001

Annex 16:

Comparison of the person-periods characteristics with and without event

Table 23: Person-periods characteristics

	Cohort 1				**Cohort 2**				**Cohort 3**				**All the cohorts**			
	Event				**Event**				**Event**				**Event**			
	0		**1**		**0**		**1**		**0**		**1**		**0**		**1**	
	N	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
BADHEALTH	4293	17.5	389	20	15054	**27.6**	954	**36.4**	6445	44.5	1255	53.4	25792	**27.6**	2598	**37.56**
PCTINC	6537	**26.7**	282	**14.5**	16117	**29.5**	398	**15.2**	4456	30.8	587	25	27110	**29.3**	1267	**19.06**
SUPPORT	21441	87.6	1760	90.5	48592	89	2379	90.7	11921	82.4	1898	80.8	81954	87.59	6037	87.28
BLCOL	8930	36.5	746	38.4	21364	39.1	1210	46.1	6666	46.1	1202	51.1	36960	39.5	3158	45.66
SELFWORK	1867	7.6	192	9.9	7763	**14.2**	564	**21.5**	4204	29.1	852	36.3	13834	**14.79**	1608	**23.25**
PUBLIC	6282	**25.7**	313	**16.1**	21213	38.8	458	17.5	4547	31.4	519	22.1	32042	**34.25**	1290	**18.65**
SECOND	10114	41.3	811	41.7	18163	33.2	783	29.9	3442	23.8	460	19.6	31719	34.52	2054	30.62
SUPER	6250	**25.5**	282	**14.5**	15969	29.2	428	16.3	2604	18	283	12	24823	**27.02**	993	**14.81**
PARTIME	4034	**16.5**	645	**33.2**	12529	22.9	971	37	4240	29.3	917	39	20803	**22.23**	2533	**36.62**

Source: ECHP 1994-2001

Annex 17:

Hazard rate models for stopping work

Table 24: Hazard rate models for stopping work (effect of caring)

Models	Bivariate (A)		Control model (B)		Full control H1 (C)		Best model (D)		Best model: control(E)	
Predictors	**Odds**	**Pr>Chi**	**Odds**	**Pr>Chi**	**Odds**	**Predictor>Chi**	**Odds**	**Pr>Chi**	**Odds**	**Pr>Chi**
Intercept				<.0001		<.0001		<.0001		<.0001
TIME	0.688	<.0001	**0.729**	<.0001	0.729	<.0001	0.708	<.0001	0.708	<.0001
TIMESQ	1.040	<.0001	1.033	<.0001	1.033	<.0001	1.035	<.0001	1.034	<.0001
BADHEALTH	1.581	<.0001	1.342	<.0001	1.342	<.0001	1.472	<.0001	1.473	<.0001
SECOND	0.625	<.0001	0.772	<.0001	0.772	<.0001	0.711	<.0001	0.709	<.0001
SUPER	0.386	<.0001	0.615	<.0001	0.615	<.0001	0.520	<.0001	0.519	<.0001
SUPPORT	0.972	0.444	**0.804**	<.0001	0.802	<.0001	0.886	0.0017	0.882	0.0011
BLCOL	1.287	<.0001	1.065	0.0199	1.065	0.0202	1.089	0.0013	1.088	0.0013
SELFWORK	1.379	<.0001	1.125	0.0008	1.125	0.0007	1.249	<.0001	1.252	<.0001
PUBLIC	0.478	<.0001	**0.616**	<.0001	0.616	<.0001	0.566	<.0001	0.567	<.0001
PARTIME	2.021	<.0001	**1.621**	<.0001	1.622	<.0001				
PCTINC	0.568	<.0001	**0.676**	<.0001	0.676	<.0001				
COH1	1.656	<.0001	**1.607**	<.0001	1.605	<.0001				
COH3	3.383	<.0001	**2.595**	<.0001	2.597	<.0001				
OLDCARE	1.177	0.0006			***0.975***	0.6221	**1.092**	0.0753		
Sample size[13]			97345		97345		98582		98582	
AIC intercept			47472.51		47472.50		49002.00		49002.30	
AIC covariate			44115.52		44117.30		46846.00		46847.50	

Source: ECHP 1994-2001

[13] The sample size refers to the number of observations (Person period) used for the analysis (after exclusion of the observations with missing values on at least 1 of the predictors).

Table 25: Hazard rate models for stopping work (effect of caring characteristics)

Models	**Bivariate (A)**		**Control model (B)**		**Full control multi-roles(C)**		**Full control intensity (D)**		**Full control duration (E)**	
Predictors	**Odds**	**Pr>Chi**	**Odds**	**Pr>Chi**	**Odds**	**Pr>Chi**	**Odds**	**Pr>Chi**	**Odds**	**Pr>Chi**
Intercept				<.0001		<.0001		<.0001		<.0001
TIME	0.688	<.0001	**0.729**	<.0001	0.729	<.0001	0.729	<.0001	0.729	<.0001
TIMESQ	1.04	<.0001	1.033	<.0001	1.033	<.0001	1.033	<.0001	1.033	<.0001
BADHEALTH	1.581	<.0001	1.342	<.0001	1.347	<.0001	1.339	<.0001	1.342	<.0001
SECOND	0.625	<.0001	0.772	<.0001	0.767	<.0001	0.773	<.0001	0.772	<.0001
SUPER	0.386	<.0001	0.615	<.0001	0.611	<.0001	0.617	<.0001	0.615	<.0001
SUPPORT	0.972	0.444	**0.804**	<.0001	0.791	<.0001	0.810	<.0001	0.803	<.0001
BLCOL	1.287	<.0001	1.065	0.0199	1.062	0.0264	1.063	0.0246	1.065	0.0205
SELFWORK	1.379	<.0001	1.125	0.0008	1.116	0.0017	1.119	0.0013	1.125	0.0007
PUBLIC	0.478	<.0001	**0.616**	<.0001	0.613	<.0001	0.616	<.0001	0.616	<.0001
PARTIME	2.021	<.0001	**1.621**	<.0001	1.592	<.0001	1.626	<.0001	1.622	<.0001
PCTINC	0.568	<.0001	**0.676**	<.0001	0.679	<.0001	0.679	<.0001	0.676	<.0001
COH1	1.656	<.0001	**1.607**	<.0001	1.632	<.0001	1.604	<.0001	1.605	<.0001
COH3	3.383	<.0001	**2.595**	<.0001	2.711	<.0001	2.601	<.0001	2.599	<.0001
HEAVY	1.496	<.0001					1.108	0.1275		
LIGHT	0.879	0.0846					0.832	**0.0189**		
CHILDONLY	0.875	<.0001			1.119	**0.0002**				
CHILD&OLD	1.046	0.5752			1.153	**0.0947**				
OLDONLY	1.165	0.0097			0.944	0.3676				
FIRST	1.208	0.0059							1.067	0.3731
SEC	1.152	0.0285							0.902	0.1402
sample size			97345		97345		97345		97345	
AIC intercept			47472.50		47473.00		47472.51		47472.51	
AIC covariate			44115.50		44105.00		44111.15		44116.37	

Source: ECHP 1994-2001

Table 26: Hazard rate models for stopping work (effect of intensity, duration and multi-roles for young cohort)

Models	Bivariate(A)		Intensity						Multi-roles						Duration	
			Full control(B)		Best model(C)		Best control(D)		Full control(E)		Best model (F)		Best control(G)		Full control (H)	
Predictors	Odds	Pr>Chi	Odds	Pr>Chi	Odds	Pr>Chi	Odds	Pr>Chi	Odds	Pr>Chi	Odds	Pr>Chi	Odds	Pr>Chi	Odds	Pr>Chi
Intercept				<.0001		<.0001		<.0001		<.0001		<.0001		<.0001		<.0001
TIME	0.654	<.0001	0.715	<.0001	0.716	<.0001	0.716	<.0001	0.701	<.0001	0.701	<.0001	0.716	<.0001	0.715	<.0001
TIMESQ	1.033	0.0026	1.019	0.088	1.019	0.0899	1.019	0.0895	1.020	0.0846	1.020	0.0849	1.019	0.0879	1.020	0.0873
BADHEALTH	1.175	0.0064	1.196	0.0049	1.198	0.0043	1.193	0.0054	1.183	0.0081	1.184	0.0079	1.192	0.0057	1.194	0.0052
SECOND	0.821	0.0002	0.857	0.0059	0.850	0.0036	0.851	0.0038	0.858	0.0067	0.851	0.0038	0.847	0.0028	0.856	0.0058
SUPER	0.462	<.0001	0.603	<.0001	0.594	<.0001	0.594	<.0001	0.601	<.0001	0.595	<.0001	0.595	<.0001	0.602	<.0001
SUPPORT	1.343	0.0002	0.873	***0.1782***					0.810	0.0374	0.811	0.0383	0.883	***0.213***	0.869	*0.1636*
BLCOL	1.083	0.1014	1.058	***0.2801***					1.041	***0.4422***					1.060	*0.2702*
SELFWORK	1.166	0.0575	1.160	0.0862	1.155	0.0942	1.154	0.0959	1.095	***0.2959***					1.163	0.0801
PUBLIC	0.565	<.0001	0.695	<.0001	0.696	<.0001	0.696	<.0001	0.675	<.0001	0.669	<.0001	0.685	<.0001	0.695	<.0001
PARTIME	2.514	<.0001	2.036	<.0001	2.044	<.0001	2.038	<.0001	1.862	<.0001	1.855	<.0001	2.028	<.0001	2.034	<.0001
PCTINC	0.475	<.0001	0.611	<.0001	0.645	<.0001	0.645	<.0001	0.616	<.0001	0.616	<.0001	0.612	<.0001	0.610	<.0001
HEAVY	1.274	0.2175	0.923	***0.7147***	0.951	0.8156										
LIGHT	0.613	0.0747	0.561	0.0533	0.563	0.0549										
CHILDONLY	1.380	<.0001							1.442	<.0001	1.454	<.0001				
CHILD&OLD	1.031	0.910							0.959	***0.8791***	0.970	***0.9104***				
OLDONLY	1.072	0.726							0.819	***0.387***	0.824	***0.4005***				
FIRST	0.998	0.991													0.859	*0.4832*
SEC	0.863	0.569													0.615	*0.1056*
sample size			25151		25151		25151		25151		25151		25151		25151	
AIC intercept			12716.00		12716.00		12716.20		12716.17		12716.17		12716.00		12716.166	
AIC covariate			12050.00		12049.00		12049.40		12010.15		12007.79		12051.00		12051.078	

Source: ECHP 1994-2001

Table 27: Hazard rate models for stopping work (effect of intensity, duration, and multi-roles for mid cohort)

Models	Bivariate(A)		Intensity Full control(B)		Multi-roles full control (C)		Duration Full control (D)		Duration Best model (E)	
Predictors	**Odds**	**Pr>Chi**	**Odds**	**Pr>Chi**	**Odds**	**Pr>Chi**	**Odds**	**Pr>Chi**	**Odds**	**Pr>Chi**
Intercept				<.0001		<.0001		<.0001		<.0001
TIME	0.645	<.0001	0.658	<.0001	0.659	<.0001	0.659	<.0001	0.658	<.0001
TIMESQ	1.053	<.0001	1.050	<.0001	1.050	<.0001	1.05	<.0001	1.050	<.0001
BADHEALTH	1.175	0.0285	1.461	<.0001	1.469	<.0001	1.461	<.0001	1.461	<.0001
SECOND	0.615	<.0001	0.705	<.0001	0.700	<.0001	0.705	<.0001	0.705	<.0001
SUPER	0.382	<.0001	0.593	<.0001	0.588	<.0001	0.592	<.0001	0.592	<.0001
SUPPORT	1.216	0.0045	0.728	0.0001	0.719	<.0001	0.722	<.0001	0.722	<.0001
BLCOL	1.334	<.0001	1.125	0.0053	1.125	0.0052	1.126	0.0051	1.126	0.0049
SELFWORK	1.165	0.0026	1.177	0.0021	1.176	0.0022	1.181	0.0017	1.180	0.0017
PUBLIC	0.346	<.0001	0.459	<.0001	0.458	<.0001	0.460	<.0001	0.460	<.0001
PARTIME	1.976	<.0001	1.578	<.0001	1.557	<.0001	1.578	<.0001	1.577	<.0001
PCTINC	0.437	<.0001	0.536	<.0001	0.537	<.0001	0.534	<.0001	0.534	<.0001
HEAVY	1.404	**0.0004**	1.061	*0.5648*						
LIGHT	0.966	0.7497	0.954	*0.6741*						
CHILDONLY	1.179	**<.0001**			1.059	*0.2021*				
CHILD&OLD	1.458	**0.0002**			1.114	*0.3157*				
OLDONLY	1.117	0.2988			0.966	*0.7557*				
FIRST	1.284	**0.0143**					1.163	*0.1542*		
SEC	1.083	0.4333					0.885	*0.2583*		
sample size			56172		56172		56172		56172	
AIC intercept			20663.84		20663.84		20663.84		20663.84	
AIC covariate			19363.31		19363.31		19360.39		19359.83	

Source: ECHP 1994-2001

Table 28: Hazard rate models for stopping work (effect of intensity, duration, and multi-roles for old cohort)

Models	Bivariate(A)		Intensity						Multi-roles						Duration	
			Full control(B)		Best model(C)		Best control(D)		full control(E)		Best model (F)		Best control(G)		Full control (H)	
Predictors	Odds	Pr>Chi	Odds	Pr>Chi	Odds	Pr>Chi	Odds	Pr>Chi	Odds	Pr>Chi	Odds	Pr>Chi	Odds	Pr>Chi	Odds	Pr>Chi
Intercept				<.0001		<.0001		<.0001		<.0001		<.0001		<.0001		<.0001
TIME	0.833	<.0001	0.861	0.0012	0.862	0.0012	0.86	0.0012	0.862	0.0012	0.862	0.0013	0.862	0.0012	0.861	0.0012
TIMESQ	1.024	0.0111	1.019	0.0585	1.019	0.0595	1.02	0.0609	1.018	0.0631	1.018	0.0645	1.019	0.0609	1.019	0.0598
BADHEALTH	1.427	<.0001	1.306	<.0001	1.307	<.0001	1.32	<.0001	1.314	<.0001	1.317	<.0001	1.320	<.0001	1.317	<.0001
SECOND	0.706	<.0001	0.778	<.0001	0.776	<.0001	0.77	<.0001	0.776	<.0001	0.772	<.0001	0.773	<.0001	0.775	<.0001
SUPER	0.574	<.0001	0.729	<.0001	0.725	<.0001	0.72	<.0001	0.724	<.0001	0.719	<.0001	0.720	<.0001	0.724	<.0001
SUPPORT	0.898	0.0584	0.846	0.014	0.845	0.0139	0.84	0.0085	0.837	0.0091	0.836	0.009	0.836	0.0085	0.835	0.0083
BLCOL	1.226	<.0001	1.019	*0.6905*					1.025	0.6157					1.022	0.6468
SELFWORK	1.184	0.0009	1.101	0.0847	1.102	0.0807	1.11	0.0605	1.109	0.0643	1.110	0.0604	1.110	0.0605	1.108	0.0644
PUBLIC	0.667	<.0001	0.798	0.0003	0.797	0.0003	0.80	0.0002	0.798	0.0003	0.797	0.0003	0.795	0.0002	0.796	0.0002
PARTIME	1.545	<.0001	1.431	<.0001	1.432	<.0001	1.42	<.0001	1.424	<.0001	1.426	<.0001	1.422	<.0001	1.421	<.0001
PCTINC	0.780	<.0001	0.898	0.0820	0.898	0.0808	0.89	0.0648	0.889	0.0556	0.888	0.0546	0.893	0.0648	0.892	0.0635
HEAVY	1.317	0.0020	1.176	0.0923	1.176	0.0923										
LIGHT	0.718	0.0032	0.760	0.0190	0.759	0.0185										
CHILDONLY	0.923	0.2697							0.873	0.0746	0.874	0.0774				
CHILD&OLD	1.351	0.0624							1.237	0.2101	1.238	0.2087				
OLDONLY	0.946	0.4826							0.907	0.246	0.907	0.2424				
FIRST	1.046	0.6728													1.021	0.8561
SEC	0.990	0.9150													0.940	0.5233
Sample size			16022		16022		16022		16022		16022		16022		16022	
AIC intercept			12680.80		12680.80		12681.00		12680.80		12681.00		12681.00		12680.84	
AIC covariate			12454.30		12452.5		12458.00		12459.30		12458.00		12458.00		12462.93	

Source: ECHP 1994-2001

Table 29: Hazard rate models for stopping work (comparison multi-roles effect between young and old cohorts)

Models	Bivariate(A)		Full control (B)		Best Model (C)		Without interactions(D)	
Predictors	**Odds**	**Pr>Chi**	**Odds**	**Pr>Chi**	**Odds**	**Predictor>Chi**	**Odds**	**Pr>Chi**
Intercept				<.0001		<.0001		<.0001
TIME	0.688	<.0001	0.775	<.0001	0.775	<.0001	0.779	<.0001
TIMESQ	1.040	<.0001	1.022	0.0036	1.022	0.0037	1.022	0.0035
BADHEALTH	1.581	<.0001	1.261	<.0001	1.262	<.0001	1.264	<.0001
SECOND	0.625	<.0001	0.828	<.0001	0.826	<.0001	0.821	<.0001
SUPER	0.386	<.0001	0.642	<.0001	0.638	<.0001	0.639	<.0001
SUPPORT	0.972	0.4440	0.825	0.0007	0.825	0.0007	0.826	0.0007
BLCOL	1.287	<.0001	1.021	*0.5524*				
SELFWORK	1.379	<.0001	1.088	0.0653	1.089	0.0635	1.097	0.0449
PUBLIC	0.478	<.0001	0.754	<.0001	0.753	<.0001	0.758	<.0001
PARTIME	2.021	<.0001	1.590	<.0001	1.589	<.0001	1.611	<.0001
PCTINC	0.568	<.0001	0.768	<.0001	0.768	<.0001	0.769	<.0001
COH3	2.043	<.0002	1.958	<.0001	1.972	<.0001	1.747	<.0001
CHILDONLY	0.875	<.0001	1.428	<.0001	1.434	<.0001	1.190	<.0001
CHILD&OLD	1.046	0.5752	0.973	0.9186	1.114	0.4527	1.122	0.4237
OLDONLY	1.165	0.0097	0.828	0.4076	0.892	0.1456	0.920	0.2873
CHILDONLY*COH3			0.583	<.0001	0.580	<.0001		
CHILD&OLD*COH3			1.213	*0.5471*				
OLDONLY*COH3			1.090	*0.7222*				
Sample size			41173		41173		41173	
AIC intercept			25866.57		25866.57		25866.57	
AIC covariate			24595.97		24590.81		24625.35	

Source: ECHP 1994-2001

Table 30: Hazard rate models for stopping work (comparison intensity effect between young and old cohorts)

Models	Bivariate(A)		Full control(B)		Best model(C)	
Predictors	**Odds**	**Pr>Chi**	**odds**	**Pr>Chi**	**Odds**	**Pr>Chi**
Intercept				<.0001		<.0001
TIME	0.688	<.0001	0.781	<.0001	0.781	<.0001
TIMESQ	1.040	<.0001	1.022	0.003	1.022	0.0032
BADHEALTH	1.581	<.0001	1.258	<.0001	1.259	<.0001
SECON	0.625	<.0001	0.827	<.0001	0.825	<.0001
SUPER	0.386	<.0001	0.644	<.0001	0.641	<.0001
SUPPORT	0.972	0.444	0.853	0.005	0.854	0.005
BLCOL	1.287	<.0001	1.021	*0.553*		
SELFWORK	1.379	<.0001	1.103	0.034	1.103	0.0333
PUBLIC	0.478	<.0001	0.767	<.0001	0.767	<.0001
PARTIME	2.021	<.0001	1.655	<.0001	1.654	<.0001
PCTINC	0.568	<.0001	0.772	<.0001	0.772	<.0001
COH3	2.043	<.0002	1.665	<.0001	1.681	<.0001
HEAVY	1.496	<.0001	0.965	*0.869*	1.130	0.1645
LIGHT	0.879	0.0846	0.555	0.048	0.718	0.0023
HEAVY*COH3			1.213	*0.414*		
LIGHT*COH3			1.359	*0.337*		
sample size			41173		41173	
AIC intercept			25866.57		25866.57	
AIC covariate			24633.81		24629.81	

Source: ECHP 1994-2001

Annex 18:

Models building

Model building depends on the hypothesis to test.

Hypothesis 1: Effect of caring

- Bivariate analysis (gross odds)
- Full-control (with all the independent variables presented before)
- Forward and backward methods were used to find a model were caring is still significant. It allows identifying the control variables that make it not significant.

Hypothesis 2: Caring characteristics on stopping work

The analyses are run separately for the 3 caring characteristics (intensity, duration, multi-roles). The predictors of care introduced in the model differ from one characteristic to another:

1) Intensity: light care, heavy care (reference category is no care for elderly)

2) Duration: first year of caring spell, caring at least for 2 years (reference category is not caring for elderly)

3) Multi-roles: caring for children only, caring for elderly only, caring both for children and elderly (reference category is not caring at all)

The use of different care specifications allows testing the effect of each predictor. One can see, for instance, if women providing heavy care have an hazard rate of stopping work significantly different than those who do not care for elderly. The effect of providing light care can also be tested.

- Bivariate analysis (gross odds)
- Full control models have been used to test the second hypothesis.

Hypothesis 3: The effect of the caring characteristics differ from one cohort to another

For the within cohort comparison (one best model for each caring characteristic –as in hypothesis 2 - of each cohort)

- Bivariate analysis
- Full control model

- Previous model without insignificant controls[14] (best model)
- Control for the best model (previous without the caring predictors)[15] allowing comparing AIC statistics and assessing if including the caring characteristics improves the fit of the model.

The between cohort comparison (one 'best model' for each caring characteristic)

- Bivariate analysis
- Full control model (cohort 1 is the reference category, interaction between cohort 3 and the care characteristics predictors are included)
- Previous model without insignificant controls or interaction terms[16] (best model)
- Previous model without interaction term[17].

Note

Alongside multi-roles, intensity and duration of care, 2 other characteristics have been tested: firstly, the interaction between *PARTIME* and *OLDCARE.* The expected relationship was that part-time workers have more time out of work for caring. Therefore, they should therefore be more able to manage their caring-working role. Small effectives among cohorts did not allow carrying out the analysis. Secondly, I would have like to investigate the effect of having a 'supportive' adult on the caring working relationship. This hypothesis was that living with another adult (non care-recipient) could lower the effect of caring on working because this third person can help with caring. The results of the regression with the interaction *SUPPORT*OLDCARE* were not clear: this term was significant on the all study sample but it became insignificant in the different cohorts. This could be due to its construction or to small numbers.

[14] Caring predictors are kept in the best model even if they are not significant because there are necessary to do meaningful interpretation.
[15] If none of the caring predictors are significant, this model is similar to the previous one.
[16] The 'main effect' of the caring characteristics (without interaction) is kept in the best model for allowing meaningful interpretations.
[17] If none of the interaction term cohort*caring characteristics is significant this model is similar to the previous one.

Annex 19:

Cross-tabs between caring characteristics and control variables

Table 31: Frequencies of predictors by caring characteristics: person-period of the all sample

	OLDONLY		***CHILD&OLD***		***CHILDONLY***		***FIRST***		***SEC***		***LIGHT***		***HEAVY***	
Event	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**1**
Predictors	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**
BADHEALTH	27.9	36.1	28.2	30.2	30.8	23.6	28.1	32.5	28.0	35.1	28.2	30.2	27.9	37.8
PCTINC	28.2	28.9	28.5	18.0	31.9	21.6	28.3	26.1	28.4	23.9	28.3	27.2	28.4	22.5
SUPPORT	88.0	77.6	87.6	86.4	84.2	93.8	87.8	80.1	87.8	81.4	87.6	88.0	88.0	73.5
BLCOL	39.9	40.8	39.9	40.2	40.6	38.8	39.9	40.1	39.9	41.0	40.1	34.5	39.7	46.8
SELF WORK	15.1	20.8	15.2	22.3	15.7	14.7	33.1	36.0	15.1	22.7	15.4	13.1	14.9	29.8
PUBLIC	33.1	34.0	33.1	37.1	31.7	35.9	51.7	44.2	33.1	34.4	33.0	39.7	33.3	30.4
SECON	34.1	22.0	33.7	31.8	31.7	37.0	33.8	27.6	34.0	23.9	33.9	25.9	33.9	25.2
SUPER	25.7	25.2	25.7	24.9	24.0	28.8	25.7	25.5	25.7	24.8	25.5	32.2	25.9	17.9
PARTIME	23.1	26.8	23.0	33.8	18.7	31.5	23.1	27.6	23.0	30.8	22.9	31.8	23.1	26.9

Source: ECHP 1994-2001

Table 32: Frequencies of predictors by caring characteristics: person-period of the young cohort

	OLDONLY		***CHILD&OLD***		***CHILDONLY***		***FIRST***		***SEC***		***LIGHT***		***HEAVY***	
Event	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**1**
Predictors	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**
BADHEALTH	17.6	25.8	17.6	27.7	17.5	18.2	17.6	25.3	17.6	28.1	17.6	31.9	17.7	21.3
PCTINC	25.8	26.3	25.9	11.4	29.1	18.8	25.8	23.4	25.9	17.3	25.8	24.9	25.9	17.1
SUPPORT	88.0	77.5	87.9	79.5	84.7	94.6	88.0	75.7	87.9	81.9	87.9	84.1	88.0	72.7
BLCOL	36.7	34.6	36.6	36.8	35.6	38.7	36.7	35.1	36.6	35.7	36.7	28.9	36.6	41.6
SELF WORK	7.8	9.8	7.7	16.4	6.2	11.3	7.7	11.7	7.7	12.9	7.8	7.6	7.7	16.5
PUBLIC	25.0	25.5	25.0	20.0	23.5	28.1	25.0	25.1	25.0	21.3	24.9	26.9	25.0	20.3
SECON	41.4	38.6	41.4	35.5	41.9	40.2	41.4	39.2	41.4	34.9	41.4	36.2	41.4	38.7
SUPER	24.7	24.2	24.7	26.8	24.2	25.8	24.8	23.2	24.7	28.1	24.7	31.2	24.8	19.4
PARTIME	17.7	17.9	17.6	35.5	12.8	28.3	17.6	25.6	17.7	22.1	17.7	22.6	17.6	25.7

Source: ECHP 1994-2001

Table 33: Frequencies of predictors by caring characteristics: person-period of the mid cohort

	OLDONLY		*CHILD&OLD*		*CHILDONLY*		*FIRST*		*SEC*		*LIGHT*		*HEAVY*	
Event	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**1**
Predictors	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**
BADHEALTH	27.8	33.4	28	27.6	31.4	23.7	27.9	30.2	27.8	31.2	27.9	29.1	27.8	32.4
PCTINC	28.8	30.7	29.2	18.3	33.9	22.5	28.9	26.6	29.1	23.6	28.9	26.7	29.0	23.1
SUPPORT	89.3	81.4	89.1	86.8	85.4	93.7	89.2	83.3	89.2	84.4	88.9	91.4	89.5	75.9
BLCOL	39.4	39.6	39.4	39.7	40.7	37.8	39.4	39.6	39.4	39.6	39.6	33.7	39.2	46
SELF WORK	14.4	17.2	14.3	21.2	14.6	14.5	14.4	18.0	14.3	20.0	14.7	11.7	14.1	27.0
PUBLIC	37.9	37.2	37.8	39.8	37.1	38.9	37.8	39.0	37.9	37.9	37.7	41.8	38.0	34.7
SECON	33.5	22.3	33.1	31.9	26.8	30.9	33.3	28.3	33.4	25.4	33.4	25.8	33.3	27.8
SUPER	28.7	27.5	28.7	25.8	16.8	32.2	28.7	27.1	28.7	26.4	28.4	34.2	29.0	18.7
PARTIME	23.6	22.6	23.3	33.0	31.6	44.3	23.5	25.5	23.4	29.1	23.3	29.9	23.5	24.8

Source: ECHP 1994-2001

Table 34: Frequencies of predictors by caring characteristics: person-period of the old cohort

	OLDONLY		*CHILD&OLD*		*CHILDONLY*		*FIRST*		*SEC*		*LIGHT*		*HEAVY*	
Event	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**1**	**0**	**1**
predictors	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**	**%**
BADHEALTH	46.1	42.5	45.7	50.2	45.8	45.7	45.9	42.0	45.9	44.6	46.5	32.2	45.3	54.4
PCTINC	30.3	27.0	30.1	21.1	31.0	22.3	30.1	26.4	30.3	25.9	30.0	29.2	30.4	23.2
SUPPORT	83.1	72.5	82.0	89.1	80.9	91.6	82.5	74.1	82.6	75.4	82.2	81.2	82.9	68.9
BLCOL	47.0	44.1	46.8	46.8	46.2	51.5	46.9	43.9	46.9	44.8	47.3	38.2	46.6	50.4
SELF WORK	30.2	28.7	30.0	34.7	29.8	32.3	30.1	28.3	30.0	30.4	30.7	18.3	29.5	40.2
PUBLIC	29.9	31.8	30.1	32.1	30.1	30.3	29.9	33.9	30.1	30.4	29.6	39.1	30.4	25.0
SECON	23.8	17.4	23.1	27.5	22.6	28.1	23.3	19.8	23.5	18.2	23.2	22.5	23.7	15.3
SUPER	16.6	22.3	17.2	17.4	17.6	13.6	16.9	22.4	16.9	21.0	16.6	28.0	17.3	15.6
PARTIME	30.2	34.9	30.5	37.7	30.0	36.0	30.5	34.1	30.3	36.2	30.2	39.2	30.6	31.7

Source: ECHP 1994-2001

Annex 20:

The dissertation month after month

September 2006: An idea…l

I arrived in the IMPALLA programme with the purpose of improving my statistical and research skills in preparation of a PhD. My idea was to orient the papers and the dissertation toward the topic I will develop in my PhD: carers for elderly. Learning longitudinal data analysis was my second purpose.

As the opportunity was given to us to choose our own topic, I started looking through the dataset questionnaire available in the CEPS. When I found a question on caring activity in the ECHP, I decided to work on it. This reaches my 2 goals: tackling the caring topic and doing longitudinal analysis.

October 2006: Precision of the topic and dissertation proposal

- Among the questions linked to caring activities in the ECHP, one deals with underemployment due to caring. I thought it could be interesting to use it as dependent variable for studying the impact of caring on working. The idea was to focus on the carers and to compare the gender differences in the caring effect, comparing 4 different countries (1 for each Welfare Regime: Italy, Denmark, Ireland and The Netherlands).

-I start looking for literature and the theoretical background on this topic.

-I deliver the proposal in the middle of the month.

-I write the 'social development' essay on the caring for elderly issue (*Paying for informal care in Flanders: A new feature of the second modernity?)* and I made use of the ECHP dataset and some of the caring variables for the statistics levelling-up paper (*Underemployment 'due' to caring activities for the elderly: Does gender matter? If so, how?*). Professor Hagenaars raises the question of the causality direction between caring and working.

November 2006: First contact with my supervisor, methodological readings & paper on the

- I started reading on survival analysis and looked into the documentation of ECHP (weight, attrition and so on).

-A first selection of the variables of interest was made.

- I start thinking about the different ways of selecting the sample

- I send a first e-mail to Professor Billiet introducing myself and asking his view about the choice (sample selection, countries to study, etc.) that I have made.

December 2006: Discussion with Professor Moors and change of the dependent variable

- At the end of the Professor Moors' lecture on discrete-time event history analysis, I discussed with him the choice of the dependent variable. One of the problems he raised was that the dependent variable deals with about opinion or behaviour and not with an objective status. I was therefore impossible to check if the respondent would have work more if he would not had been a carer. Also, the number of cases could be problematic. Some exploratory analyses confirm this 'fear' of small number of cases. I then changed the dependent variable: it would be a decrease in working-time (e.g. starting with full-time workers as study population and then move from to part-time or no work as event).
- The professor Moors lecture introduces me to the different issues link to event history analysis. It makes me wonder if it will not be better to work only on 1 country. I investigate the possibility to study only **Netherlands** and I read more specific literature on this country.

January 2007: Improving statistical skills

- In January, no time for dissertation, all the energy is directed toward the preparation of the statistical exams. It allows me to have a better understanding of the methods I plan to use in my dissertation.

February 2007: Netherlands focus

- The policy cycle individual paper gives me the opportunity to have a better knowledge of the informal care policies in the Netherlands *(Carers support policy in the Netherlands: The evolution of the Health System and its consequences on the elderly's carers: an evaluation).*
- I wonder what should be the best statistical method to use: multi-level analysis or event history analysis. I send a mail to professor Moors stating the pros and cons I see in each method regarding my research interest. I ask him if he could give me some advice for choosing one of the 2. He did not answer. The literature makes me choose event history analysis.

March 2007: Sample selection

- I continue to think about the selection: Do I have to work only on carers and also on non-carers (and to see the effect of caring)? Would it be better to focus on women? How many waves to consider for the analysis? Do I need a balanced sample (all the individuals are participating in the same number of waves) or can I work on an unbalanced one? Do I have enough transitions for each wave to study the phenomenon? After several discussions with Professor Billiet and 2 of his assistants and various exploratory analyses, the following was fixed: the study sample is unbalanced, composed of working women of all the countries participating since the first wave (not enough events in the Netherlands only, all the waves of ECHP will be used).

April 2007: From reduction of number working hours …to stopping work

- For the Social policy paper (*Women carers for the elderly and employment: Is the Netherlands a hybrid model?*), the transition from full-time to part-time or not working among working women in the Netherlands, Denmark, Italy and Ireland was analysed using simple logistic regressions on an individual file. The results obtained are not easy to interpret and the definition of full-time is difficult to fix when working on different countries. This experience makes me change the dependent variable one more time: the event will be stop working.

- Papers for auditing (Assignment 1. *Respite service for the elderly carers: Performance plan of a Welfare State organization.* Assignment 2. *Audit commission-UK (2004)"Support for carers of older people: Independence & well being 5": Assessment of the quality of a performance audit report)* and for Evaluation and tools (*Critical evaluation of an example of a program evaluation: 'Evaluation of the Australian Commonwealth Respite Care Program:A case study from Western Australia & the Australian Red Cross'* (Peter J. Hancock, Jayne A. Jarvis) In *Evaluation and Program Planning* 28 (2005) 301–311)are dealing with the informal care issue too.

May 2007: Last decision on selection of the study population, struggling with bias, construction of the person-period file and first draft

- Some last changes are brought: Luxembourg is excluded (the question on caring is not asked) and the analyses will start with the second wave (one important question on caring is not asked at the first wave).

- The meeting with professor Hagenaars leads me to investigate the potential problems that could bias the results. For selection bias problem, I ask advice to Professor Williams. Finally, I decided to not apply the Heckman correction.
- The person-period file is created (it takes 10 full days with the help of my father who is an computer engineer, to have a correct person-period file).
- I write the first draft of the dissertation.

June 2007: re-reading, correction and setup

July 2007: Delivery