

IZA DP No. 2010

**Informal Care and Employment in England:
Evidence from the British Household Panel Survey**

Axel Heitmueller
Pierre-Carl Michaud

March 2006

Informal Care and Employment in England: Evidence from the British Household Panel Survey

Axel Heitmueller

*Department for Work and Pensions, UK,
London Business School and IZA Bonn*

Pierre-Carl Michaud

*RAND Corporation
and IZA Bonn*

Discussion Paper No. 2010
March 2006

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
Email: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit company supported by Deutsche Post World Net. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Informal Care and Employment in England: Evidence from the British Household Panel Survey*

More than 40% of the respondents in the British Household Panel Survey provide informal care at least for one year within the period 1991-2003 and carers are usually less likely to hold simultaneously a paid job. There is little evidence on the mechanism that links informal care provision and labour market outcomes. This paper provides evidence on the pathways through which this pattern arises using a multivariate dynamic panel data model that accounts for state-dependence, feedback effects and correlated unobserved heterogeneity. We find evidence of a causal link from informal care to employment with employment rates reduced by up to 6 percentage points. However, this effect is only found for co-residential carers who account for one third of the population of carers and less than 5 percent of the overall labor force. For the same group, a significantly smaller link from employment to care provision is found. A micro-simulation exercise using the model estimates suggest that the overall potential pressure on the provision of informal care created by a rise in the employment rate is minimal.

JEL Classification: I0, J2, C3

Keywords: informal care, employment dynamics, aging, dynamic panel data models

Corresponding author:

Axel Heitmueller
London Business School
Regent's Park
London, NW1 4SA
United Kingdom
Email: aheitmueller@london.edu

* We thank Julie Zissimopoulos for numerous comments and participants of the ISER seminar at the University of Essex. The views expressed do not necessarily reflect the opinion of the Prime Minister's Strategy Unit.

1. Introduction

At least 40% of individuals in Britain look after sick, disabled, and elderly relatives and friends for at least one year of their lives according to the British Household Panel Study (BHPS). The 2001 Census shows that there were about 5.2 million informal carers in England and Wales. The General Household Survey estimates the total number of carers to be 6.8 million for the whole of Great Britain in 2000 (ONS, 2000). The majority provides care for parents, friends and family living outside their own homes. However, there is also a rising number of carers looking after someone in their own home reflecting the growing preference of elderly people for non-residential care (OECD, 2005).

Several social and economic factors are likely to impact on the informal care market in the short and medium term. Increasing longevity and rising rates of disability will undoubtedly continue to increase the demand for care services. Furthermore, changing family patterns such as lower marriage rates, fewer children, greater geographic mobility, and declines in intergenerational co-residence are also likely to contribute to changes in informal care patterns over time as the vast majority of informal carers look after parents and spouses (Grundy, 2000). At the same time, the provision of home care by the Social Service has been decreasing while higher home ownership rates among the elderly population might have increased the demand for these services. For example, in the UK between 2000 and 2003 the number of households in receipt of formal care has fallen by 9% from almost 400,000 to 363,000 (U.K. Department of Health, 2003). Similarly, a disproportionately lower increase in long-term care supply compared with the increase in the aging population has been observed for the US which may partly be due to an increase in informal care arrangements (Lakdawalla and Philipson, 2002).

Despite already higher than average labour market participation in Britain compared to most other European countries, the UK government is determined to increase it even further and exceed the employment targets set by the Lisbon Agenda in 2000. With individual's time being scarce, increasing demand for informal care provision may mitigate these efforts. Increasing formal care provision may be an alternative though little is known about substitutability of informal and formal care. Estimates of the replacement cost to e.g. the UK government for informal care range from £21 billion per

year in 1999 (Laing and Buisson, 2002) to as much as £57 billion per year (Carers UK, 2000).¹ Already, costs for long-term care in OECD countries range from 0.5 to 3.0 percent of the GDP and is estimated to increase further (OECD, 2005). Hence, any policy affecting labour supply and in particular the employment decision of women who are providing the bulk of informal care may add to the costs of formal care.

Yet, surprisingly little is known about the association between informal care and employment decisions – particularly so in Britain – beyond mainly qualitative evidence.² In this paper, we investigate the reasons behind the stylized fact that carers work less than other employees. In particular we address the question as to whether individuals give up work in order to engage in informal care or whether individuals take up care responsibilities in the absence of employment opportunities? This is a crucial distinction from a policy makers' point of view. If caring keeps people from working, policy should focus on the provision of formal care assuming that the two are substitutes. However, if on the other hand lack of work opportunities is the reason why people take up caring responsibilities, an expansion of the formal care market will not necessarily increase labour force participation. Similarly, an increase in the labour market participation may reduce the informal care provision at a time when the demand for care is increasing, putting additional pressure to develop formal care services further. Finally, if poor households are more prone to face caring responsibilities and fail to have opportunities to engage in labor market activities, policymakers might target such groups in order to improve their position.

But to evaluate different economic policies, we ought to step from correlations to causality. It turns out that there has been little work on disentangling the pathways through which this correlation arises. Very few empirical studies try to unravel the interaction between informal care and employment. Recent papers have shown that unobserved persistent factors account for part of the relationship (Leger, 2002; Stern, 1995). We follow a somewhat more general approach by looking at the sequence of events that lead individuals to be out of the labor force and caring. This involves

¹ Unfortunately it is not clear whether the figure considers that many disabled people have to contribute to their care costs. Yet, using very conservative assumptions about hourly care costs and only considering individuals caring for more than 20 hours a week it can be shown that replacement costs are very likely to be above £15bn a year.

² For a summary on UK informal care literature see research report by DWP (2005)

modeling the sequence of caring and employment spells and looking for systematic differences in transition rates conditional on current states. In that process, we allow for correlated unobserved heterogeneity therefore accounting for three possible channels through which the interdependence may arise. For example, this allows testing if for a given labour force status those who care face lower probabilities in the future compared to those who do not provide informal care and vice-versa for the effect of work on care. At the same time we can allow for unobservables “causing” both decisions. This framework is an adaptation of Granger non-causality to longitudinal data (Granger, 1969; Chamberlain, 1984).

The structure of the paper is as follows. In Section 2, we briefly survey the theoretical foundation for the relationship between work and caring. Section 3 presents the data while section 4 presents the empirical strategy used. In Section 5 we discuss the results and present some policy simulations. Section 6 concludes.

2. Theoretical Foundations and Empirical Regularities

The attachment of informal carers to the labour market has been the concern of policy makers for some time.³ Yet, informal care and its impact on labour market participation have widely been ignored in the British literature, Carmichael and Charles (1998, 2003) and Heitmueller (2005a) being the exception. However, there are number of U.S. studies on particular care types which are briefly summarized below. Formal models on informal care decisions more generally are however, rare and mainly confined to intra-household decision of children caring for their parents. Stern (1995) and Engers and Stern (2002) provide theoretical models to describe how children decide on the care for their elderly parents. Soldo and Wolf (1994) present a simple utility framework to model the intra family care decision of married women to their elderly parents and its impact on employment.

Theoretically, the impact of informal care on labour market decisions depends on a number of factors. Firstly, care decisions can be negotiated in an individual or family context depending on household size and care relation. This aspect is not further pursuit in the paper for lack of data on family decisions and the demand for care an individual

³ See <http://www.carers.gov.uk/>.

faces is treated exogenously.⁴ Secondly, informal care introduces a further time constraint on the individual complicating the usual trade-off between leisure and work or leisure and consumption. Trivially, informal care is a competing use of time but it has the same opportunity cost as leisure (i.e. the wage evaluated in terms of marginal utility of income). Hence, the effect on leisure and employment depends crucially on the income and substitution effect and the preference relation of leisure and informal care (substitute or complement). Finally, the magnitude of the care effect on employment and leisure will also depend upon the availability of a formal care market and the substitutability of informal for formal care.⁵

In particular, for a given level of leisure, tightening the time constraint leads to a loss in income (a decrease in hours worked) and raises the marginal utility of income. Therefore, to maintain the same consumption level, leisure has to be reduced. However, if leisure is a complement for care (e.g. informal carers require respite), care increases the value of leisure. In that case optimal leisure can increase with informal care if this effect outweighs the income effect created by the “time crunch”. On the other hand, if leisure is a substitute for care, leisure will decrease with care. That simple model therefore illustrates that the effect of informal care on employment is ambiguous and depends on preferences and the nature of the goods. Although it is more likely that care and leisure are complements rather than substitutes, particularly if care is provided at home, the effect remains indeterminate.

Furthermore, using this simple utility framework it can be shown how reverse causality may arise, from employment to care. For workers with high relative opportunity costs of time, leisure and informal care are relatively more costly. Coward and Dwyer (1991) find a strong negative correlation between opportunity costs and caregiving. For example, those who are employed and climbing the experience ladder are more likely to have high opportunity cost of time and may also be more likely to substitute formal for

⁴ In particular, the empirical model controls for several family characteristics such as household size and family bargaining may be partly captured as unobserved heterogeneity for which the model explicitly controls.

⁵ Again, this is not further pursued in the empirical part as mixed provision of informal and formal care are constrained through institutional settings in England. Recent evidence from Scotland where the government has introduced free personal care also show that this had no effect on the provision of informal care (Bell and Heitmueller, 2006).

informal care. As ability and formal care are often unobserved in the data this may lead to a bias in the estimated effect of care on employment.

On the other hand, those with poor employment prospects are likely to be the same individuals that are more likely to experience a demand for care in the future. This is quite plausible since disability is strongly related to socio-economic status and socio-economic status is quite persistent over generations. Therefore disadvantaged individuals, who are more likely to be out of the labor force may also be more likely to have a sick parent or spouse. In that case, third factors lead them to be out of the labor force and provide care but none of these statuses causes the other.

Other effects such as the desire to take breaks from caring responsibilities (respite effect) and lower expected earnings due to less job reliability are likely to also impact on labour supply (discrimination effect). Again, these factors are rarely captured in survey data.

Empirically, most studies identify a negative correlation between employment rates and caregiving responsibilities (e.g. Soldo and Wolf, 1994; Léger, 2000). Yet, a main difficulty is to determine the direction of causality as outlined in the above discussion. For example, Soldo and Wolf (1994) employ a two-stage, double sample-selection model to simultaneously estimate the relationship between informal care provision, labour market activity and hours of work supplied by married women in the US to address the endogeneity inherent in the decision process. They find no effect of care on hours worked or employment when looking at a cross-section of married women. However their identification strategy relies on strong assumptions. The authors are careful to point out that non-linearity is probably not a solid base for identification but they employ rather doubtful exclusion restrictions in order to identify causal effects (family characteristics and education). These are likely to be correlated with other unobservables that cause both employment and caring (Stern, 1995). The estimated effects on hours of work are large but highly variable (from -3.7 to 2.7) depending on which set of exclusion restrictions is use. These effects are not estimated precisely which could also be due to the use of an inefficient two-step method relative to a full information maximum likelihood procedure.

In order to address endogeneity concerns, a perhaps more promising strategy is to use panel data in order to control for third-factors leading to the association between caring and employment. Léger (2000) shows that controlling for potential endogeneity in living arrangements of adult children reduces but does not eliminate the impact of parental illness on the labour supply of women in the U.S. Heitmueller (2005a) finds that not controlling for unobserved time-invariant heterogeneity in the provision of informal care, the impact of care on labour supply is overestimated for co-residential carers in England and fails to find an impact of care on employment for other types of carer (e.g. extra-residential carers). Although Léger and Heitmueller's analysis address endogeneity due to third factors, they do not tackle the issue of bi-directional causality. The remaining effects of care on employment and hours could also be due to reverse causality of employment/hours on care. This bidirectional causality is enough to invalidate simple regressions of care on employment and vice-versa as it may overestimate each of the causal pathways. What we ultimately want to know to inform policy is how increases in formal care arrangements (that may reduce the demand for informal care), subsidies to informal care or increases in employment opportunities affect the supply of informal care and labor. Understanding pathways that lead carers to be without a job is key to answering these questions.

3. Data

The analysis in this paper is based on data from the British Household Panel Study (BHPS) from 1991 to 2003. Each year over 5,000 households consisting of roughly 10,000 individuals have been interviewed. Most of these are re-interviewed in subsequent years allowing us to follow their behavior over time for a maximum of 13 years. The BHPS offers a wide range of variables for our variables and is nationally representative. For the purpose of our analysis only individuals who are aged 16 to 64 (59 for women), and not working for the armed forces or in self-employment.⁶ We limit our analysis to England because of sampling weight problems due to incoming boost samples from Scotland and Wales in 1999. Participation in paid employment is defined by whether an

⁶ We make the cut at the state pension age because few individuals work past that age. This age is different for men (65) and women (60) in the U.K.

individual has done paid work in the week prior to the interview or has not done paid work but has had a job from which they were absent.

Finally, we work with the unbalanced sample after deletions of cases with missing information on important covariates for the analysis (e.g. care, education, etc.). Appendix A gives a summary of how the analytic file is constructed. In all descriptive figures, we use sampling weights. In Table 1 in Appendix A, we present the entries and exits from the panel between 1991 and 2003.⁷

Individuals are classified as carers if they provide an affirmative answer to at least one of the following questions:

Co-residential carer: *“Is there anyone living with you who is sick, disabled or elderly whom you look after or give special help to?”*

Extra-Residential carer: *“Do you provide some regular service or help for any sick, disabled or elderly person not living with you?”*

Two groups of carers are distinguished according to these questions, co-residential and extra-residential. The former refer to individuals caring within, the latter to individuals caring outside their own home. Figure 1 reporting age-cohort profiles shows how caring responsibilities usually increase with age and that caring is more prevalent among females. At age 25, the proportion of male and female carers is similar (5.6 vs. 5.7 percent). By age 50 however, more than a quarter of females are carers while this fraction never rises above 20 percent for male cohorts. Since the age profiles in Figure 1 are cohort-specific, we can also look of at the same age, if caring increases or decreases over cohort. For females, there appears to be no real cohort differences. The only exception is perhaps for the group born between 1966 and 1975 which appears at age 45 to exhibit some difference from the preceding 1956-1965 cohort. For males, the

⁷ In 1991, 4,291 respondents enter the panel, of which 618 exit the following wave. In total, 5,248 respondents entered the panel and 5356 have exited following waves. As a result there are 3,565 respondents left in the 2003 wave. Given our selection criteria, exits are mostly due to individuals becoming eligible for state pension but also respondents due to respondents quitting the panel. Attrition rates range from 10 percent to 20 percent per wave. The entries are mostly respondents becoming 16 years old enter our analysis and a refreshment sample that enters in 1997. We do not consider the case of non-random attrition. It is unclear how this is related to informal care. The proportion of carers among those who quit before 2003 is relatively the same as that of those who remain in the panel.

cohort profiles vary a lot, making it difficult to see any clear trends in the fraction of carers over cohorts. A priori, the large increasing trend over time in labor force participation of females does not appear to be associated with any equivalently large decreasing effect on caring. At the same time where the supply of individuals wishing to be carers may have fallen (because of the increase in labour force participation), the demand for carers may have increased even more from the aging of a significant share of the population.

Figure 2 gives an overview of the different types of carers we find in the BHPS sample. Three points are worth noting. There is a larger proportion of female carers than male carers (16.2 percent vs. 11.2 percent). The majority are extra-residential carers. About three quarters of carers look after individuals not living with them (11.8 percent (7.6 percent) for females (males)). Finally, a negligible proportion (less than 1 percent) of respondents do both extra and co-residential caring. Presumably, these two types of care involve different considerations. Co-residential carers usually care for close relative such as a spouse or a parent while extra-residential carers mostly care for parents or other relatives. Furthermore, co-residential caring appears to be time intensive. The BHPS asked a question about the hours spent caring. There is a significantly larger proportion of co-residential carer who report caring for more than 20 hrs a week (46 compared to 5 percent for extra-residential carers). Therefore, the difference between the two groups has a lot to do with the intensity of caring activities. Table 2 gives a socio-demographic portrait of carers in comparison to non-carers. Carers, whom we know are generally older, are mostly married individuals with significantly lower education. One out of three carers has no diploma (23% of non-carers have no diploma). They are also less likely to have young children at home (8% compared to 14% for non-carers). Finally, they are more likely to say that their health limits their ability to participate in the labor force (17 compared to 10 percent for non-carers). Of course, these are all differences that do not account for the simple fact that carers are older than non-carers as seen in Figure 1 already. But these differences persist even when controlling for age.

Table 3 reports on how much economic resource carers live on compared to non-carers. Median annual household income (adjusted using OECD equivalence scale)⁸ is somewhat lower for carers than non-carers (£22,522 vs. £20,136). This represents a 10 percent difference. Most of the difference is made up by lower earnings from paid work (£26,945 vs. £23,689). This difference has two components: the earnings differential for carers vs. non carers and earnings differential for other earners in the household. The first component reflect the possibility that carers are mostly non-workers, work less hours or that other characteristics associated with informal care account for permanent differences in earning potential. The second component can reflect a non-labor income effect on the employment and/or caring decision. At the same time, it can also reflect the lower earnings of household members who are disabled. The effect of non-labor income is unclear and depends on the dynamics of informal care and work decisions.

The difference in household earnings is compensated to some extent by the fact that households with carers collect more benefits than non-carer households. The fraction of carers receiving benefits is 7 percentage points higher than for non-carers and the median carer receiving benefits gets £3,482 per year. From 1991 to 2003, there has been a means-tested caring allowance (not always under the same name) of £2,500 per year.⁹ Last but not least, unobserved differences in benefit take-up for other programs can also explain the difference in the amount of benefits households with carers receive compared to non-carers. A surprising finding is that carers receive lower transfers from relatives compared to non-carers. Finally, median household with carers and non-carers have approximately the same amount of capital income.

We next consider how care is related to employment and hours worked. Labour market participation is defined by whether a respondent has done paid work in the week

⁸ The OECD scale is 0.5 for additional adults in the household and 0.3 for children. Co-residential recipients of care presumably use up more of the resources than a healthy adult. We did not attempt to correct the scale for that. Since this would presumably increase the household equivalence scale, we view the differences in household income found in Table 3 as lower bounds.

⁹ The BHPS allows determining whether somehow receives the carers allowance or not. A small number (2.5%) of male carers report receiving such an allowance while this figure is somewhat larger for females (7.3%). The low fraction of households receiving the carers allowance is potentially related the income threshold for the carers allowance is quite low (approximately £5,000 per year). Furthermore, receipt of the carers allowance can decrease other benefits the household receives such the financial relief provided might not be large. Therefore it is unclear how much incentive the carers allowance provides.

prior to the interview or has done no paid work but has had a job from which he/she was temporarily absent. We classify reported weekly hours of work into two categories, (1-29) and (30+) along with a third category for missing hours for those who report doing paid work. There is a large share of respondents with missing hours (approx. 40 percent) and the censoring can go either way (those with large hours or low hours can decide not to report them). In Table 4, we compare hours and participation for carers and non-carers. There is a large difference in labour force participation rates for both males and females. For males, the difference is 9.7 percentage points while for females it is somewhat smaller (7.6 percent). However, there is not a large difference in the fraction performing part-time work which suggests that the extensive margin really captures the first order effect. This might not be really surprising for men. Less than 2 percent work less than 30 hours per week.¹⁰ Although working carers may have a preference for working less, it may be that it cannot be chosen because of demand side constraints or hours restrictions. Hence, we focus on the relationship between the extensive margin (participation) and caring. Soldo and Wolf (1994) do not report large differences in hours worked for carers and non-carers in the U.S. However, the difference in terms of employment rates is remarkably similar to ours (5 to 10 percentage point difference)

Table 5 presents results from probits of the caring decision on the participation decision controlling for demographics but also non-labor income (excluding earnings from other household members and benefit income who are clearly endogenous if these are being cared for). Explicitly, this exercise is a test of conditional independence. If we do not reject the null of conditional independence, then there is no need to look into causality. To control for non-labor income, we use a strategy proposed by Hyslop (1999) and include both the individual sample average over time for non-labor income (permanent non-labour income) along with the transitory component (the deviation from the sample average) in the regression. This allows controlling for any persistence in the errors due to unobserved differences across respondents in non-labor income. We allow for clustering of the unobservables at the individual level which is the same as allowing

¹⁰ Of course, censoring appears to be somewhat different between carers and non-carers. A higher fraction of carers report hours worked. Suppose censoring is from below. Then we underestimate the percentage of workers working less than 30 hours a week. If censoring is from above we underestimate the percentage of those working more than 30 hours a week. The overall effect is non-trivial and depends on the censoring points along with the relative dominance of censoring from above and below.

for random effects in the probit model but with a different normalization (Arulampalam, 1999) The results show that we can easily reject conditional independence between caring and participation decision. The conditional difference in the probability of caring between workers and non-workers is relatively large (-3.8 percent (-2.1 percent) for females (males)) given that the observed probability of caring in any period is 16 percent for females and 11 percent for males. Results for other regressors confirm in part differences from descriptive statistics of Figure 1 and Table 2. But there are reasons to believe that a dynamic analysis is needed in order to understand the relationship between care and employment.

The labor force participation process is known to be dynamic with interruptions quite frequent particularly for females (Heckman and Willis, 1977). We also know that it is quite a persistent process where both unobservables correlated over time and true differences in the probability of working depending on past employment history create dependence (e.g. Hyslop, 1999). It turns out that there is little known about the dynamic properties of care-giving spells, how frequent they are?, how persistent? etc. For example, there is evidence that while mental distress is high at the onset of a caring spell, individuals adapt fairly rapidly to their new role reducing perceived distress levels (*habituation* effect) (Heitmueller, 2005b). This may be due to learning effects or simply because of better support for long-term carers. This evidence therefore suggests that caring may become easier over time the longer the caring spell. Table 6 reports some feature of the distribution of caring spells. Each year, an average of 5.1 percent (6.3 percent) of male non-carers become carers while the reverse transition is made in higher proportion among males (39.1 vs 32.2 percent). We observe 42.8 percent (32.8 percent) of females (males), spending at least a year as a carer out of the 13 years we survey most respondents. Of those who care at least once over 13 years in the panel, 79.3 percent of females (males) care for more than one year. Although caring appears to be persistent, there is a relatively large proportion of carers who make more than 2 transitions (29 percent of females and similarly for males). In the next section, we present an econometric model that uses those dynamics to identify causal pathways through which carers end up working less than non-carers.

4. Empirical Strategy

There is no consensus on the type of model that should be used to estimate determinants of care and employment decisions. The family bargaining approach involves other members of the family and hence characteristics of care recipients. Therefore, estimating a structural model of such family decisions would require data on the family (number of siblings, their income, etc.) and not only on the household (see Engers and Stern, 2002). Unfortunately, this information is not available in the BHPS.

A unitary structural model, not involving decisions of the care recipients and other family members such as that proposed in Soldo and Wolf (1994) would require that we observe when care needs are present since this is the only way through which caring becomes desirable (health of the potential recipient enters the utility function). Such information is available in the BHPS only for other household members living with the respondent.

One would face two difficulties in modeling such choices structurally using that approach. First, we only observe characteristics if the potential care recipient lives with the household. But most care recipients will move in once the respondent has decided to care therefore introducing a sample selection problem since we do not really observe characteristics of potential care recipients living outside the household. Furthermore, co-residential caregiving accounts for less than 25% of overall caregiving, leaving out of the analysis the vast majority of carers.

Our approach relies instead on using the sequence of events that lead carers to be out of the labor force and vice-versa. By measuring transition rates conditional on past caring and employment states, we can use panel data to identify causal pathways from transitions between the two states. This reduced-form approach has the advantage of not requiring information on family characteristics and not requiring the use of instruments to identify the effect of one decision on the other. Although family background information is potentially informative about choices, it will not affect an analysis based on transition rates if such characteristics are constant over time. We can control for unobserved heterogeneity responsible for sorting across the two states (non-workers into caregiving

and workers out of caregiving) and test whether past caring spells modify the transition rate into employment and out of employment. Therefore, the identification strategy is to use “time” or the sequential nature of outcomes as an instrument while carefully allowing for sorting based on unobservable third factors causing both decisions.

4.1 The Model

Each respondent i has at time t , a decision to make regarding whether or not to provide informal care and participate in paid work. Denote c_{it} the decision to care, assumed to be dichotomous (0,1) and similarly h_{it} the decision to participate in paid work. We focus on the extensive margin because descriptive evidence appears not to reveal important differences in terms of hours worked. We propose to use a bivariate dynamic model that takes the form

$$\begin{aligned}
 c_{it}^* &= x_{it}'\beta_c + \gamma_h h_{it-1} + \delta_c c_{it-1} + u_{it}^c \\
 h_{it}^* &= x_{it}'\beta_h + \gamma_c c_{it-1} + \delta_h h_{it-1} + u_{it}^h \\
 c_{it} &= I(c_{it}^* > 0), h_{it} = I(h_{it}^* > 0) \\
 i &= 1, \dots, N; t = 1, \dots, T
 \end{aligned} \tag{1}$$

where x_{it} is a vector of observable characteristics from the respondent, assumed to be strictly exogeneous with respect to u_{it}^c and u_{it}^h , unobservables. With γ_c, γ_h set to zero, this represents a pair of panel data dynamic binary choice models as used in Heckman (1981) and Hyslop (1999). In such model, δ_c, δ_h capture state-dependence defined as the difference in caring and employment propensities between carers/non-carers and workers/non-workers in the previous period. Those who care have higher probabilities to care next period, perhaps because of learning. Similarly, those who work may have higher probabilities to work next period, perhaps due to adjustment costs, human capital depreciation while out of the labor force or learning-by-doing on the job. There is strong evidence to suggest that there is state-dependence in employment probabilities (Hyslop, 1999; Honoré and Kyriazidou, 2000; Michaud and Tatsiramos, 2005).

We will argue that the parameters γ_c, γ_h define causal pathways. They represent respectively the difference in employment propensity between carers/non-carers conditional on past employment states. With error terms $u_i^k (k = c, h)$ independent over time, these causal parameters could be estimated from a pair of probits where perhaps one would allow for a contemporaneous correlation in these error terms across equations. However, we will also allow these error terms to be correlated over time, presumably because of the presence of unobserved heterogeneity. In such case, γ_c, γ_h as well as δ_c, δ_h cannot be identified solely from differences in conditional transition rates. Unobserved heterogeneity will imply sorting of respondents over time in a way that conditional transition rates will differ without there necessarily being a true difference in individual transition probabilities conditional on the state occupied. Suppose there are two types carers and two types of workers,

- $a_{h,1}$ and $a_{h,0}$ represent employment probabilities for high and low type workers respectively ($a_{h,1} > a_{h,0}$). These differences can be due to ability, taste or opportunities.
- $a_{c,1}$ and $a_{c,0}$ represent caring probabilities for high-risk carers and low-risk carers. For example, these differences can be due to family unobservables (few siblings, etc.).

Suppose for the sake of the argument that only the pairs $(a_{h,1}, a_{c,0})$ and $(a_{h,0}, a_{c,1})$ exist in the population of interest and denote their respective share in the population as p and $(1-p)$. The observed employment and caregiving rate in the initial period is

$$\begin{aligned} p_{h,0} &= pa_{h,1} + (1-p)a_{h,0} \\ p_{c,0} &= (1-p)a_{c,1} + pa_{c,0} \end{aligned} \quad (2)$$

In the second period, the group of those employed last year, caring this year is composed of the two types, in proportions $pa_{h,1}a_{c,0}$ and $(1-p)a_{h,0}a_{c,1}$. Therefore the conditional probability of caring given employment last period is given by

$$P(c_t = 1 | h_{t-1} = 1) = \frac{P(c_t = 1, h_{t-1} = 1)}{P(h_{t-1} = 1)} = \frac{pa_{h,1}a_{c,0} + (1-p)a_{h,0}a_{c,1}}{pa_{h,1} + (1-p)a_{h,0}} \quad (3)$$

while the probability of caring in the group that did not work last period is given by

$$P(c_t = 1 | h_{t-1} = 0) = \frac{P(c_t = 1, h_{t-1} = 0)}{P(h_{t-1} = 0)} = \frac{p(1 - a_{h,1})a_{c,0} + (1 - p)(1 - a_{h,0})a_{c,1}}{1 - P(h_{t-1} = 1)} \quad (4)$$

It can be shown easily that $P(c_t = 1 | h_{t-1} = 1) < P(c_t = 1 | h_{t-1} = 0)$ as long as there is heterogeneity in these transition probabilities ($p \in (0,1)$ and $a_{k,1} \neq a_{k,0}, k = h, c$) (Heckman and Willis, 1977). The construction used here will *de facto* create strong differences because we have allowed perfectly negatively correlated types (there are no high-risk carers and high opportunity costs workers). To see that note that the pool of non-workers has a higher fraction of high caring types $(1-p)(1 - a_{h,0})$ because those have low employment probabilities compared to $(1 - p)a_{h,0}$ for workers. Therefore, more carers will emerge of that group compared to the group of workers. Since employment is negatively correlated with caring propensity, the selection implies that previous employment is informative about future caring propensity. But this is not because employment affects future caring propensity but because this pool of workers are less likely to be high-risk carers and more likely to be high opportunity cost types. Non-workers do not face different incentives to care than workers and therefore working does not have a causal effect on caring next period. Hence, the use of lagged states to infer causality breaks down when there is correlated unobserved heterogeneity unless it is properly accounted for. This is similar to the spurious vs. true state-dependence distinction in univariate models (Heckman and Willis, 1977; Heckman, 1981). The difference $P(h_t = 1 | h_{t-1} = 1) - P(h_t = 1 | h_{t-1} = 0)$ can be positive even in the absence of true state-dependence effects. Low probability types transit more rapidly to the pool of non-workers and therefore employment rates out of that state are lower than those in the pool of previous workers. Yet, such a distinction between true state-dependence/lagged causal effects and unobserved heterogeneity is crucial for designing policy.

An example given by Heckman and Willis (1977) is that of the study of unemployment dynamics. If state-dependence is predominant, policies that target the

long-term unemployed might be more effective at reducing unemployment while if heterogeneity is more important, it will point to a better screening mechanism of applicants and human capital type policies to raise the "employability" of those with disadvantaged profiles. The same distinction might also be important for informal care. One driving force behind the positive state-dependence in informal care may be the above mentioned learning effect i.e. the respondent learns his ability to care and consequently continues to care in successive periods because he becomes more productive. In contrast, an increase in the care needs of recipients that is potentially unobserved may be an example of unobserved heterogeneity. Again, this will decide about the target group of policy interventions.

To account for unobservables, we assume the unobservable $u_{it}^k (k = c, h)$ is decomposed in a time-invariant unobserved heterogeneity term α_i^k and a time-variant shock ε_{it}^k . Unobserved heterogeneity is distributed independently across respondents with variance σ_α^k and with a correlation across decisions ρ_α . Therefore, this allows assessing if types are negatively correlated or not. We assume a similar covariance structure for ε_{it}^k allowing for shocks to affect each decision simultaneously.

We use two set of assumptions in order to regarding unobserved heterogeneity. The first one is a fixed effect model where we require only that x_{it} be strictly exogenous with respect to ε_{it}^k but not α_i^k . This leads to a fixed effect formulation and because of the incidental parameter problem (Neyman and Scott, 1948) we estimate a linear probability specification.¹¹ The other approach is to assume strict exogeneity of x_{it} with respect to u_{it}^k which is a stronger assumption, generally referred to as the random effects model.

Unobserved family characteristics that are fixed over time can be accounted for using the

¹¹ In most non-linear settings, the parameters to estimate increase with sample size but it is impossible to estimate consistently parameters of interest neglecting the nuisance parameters (the fixed effects). In the binary dynamic fixed effect case, Chamberlain (1985) proposed a conditional logit approach for estimation of the dynamic model when there are no regressors while Honoré and Kyriazidou (2000) proposed a similar estimator for the more general case by imposing strong requirements on the distribution of covariates. In both cases, in addition to their other disadvantages (such as not providing average partial effects), there is no multivariate generalization that would enable to estimate the model in (1).

fixed effect approach while in the random effects model, these will probably lead to biased estimates of some of the other time-invariant characteristics. It is unclear however, how this will affect the lagged caring and employment effects as their identification relies mostly on variation over time in some observed/unobserved characteristics. Comparing fixed and random effect estimates provides a relatively good check on that possibility. Hyslop (1999) and Alessie et al. (2004) use similar strategies.

It turns out that there is a direct correspondence between the model presented here and the model used in Soldo and Wolf (1994). Their cross-sectional model takes the form¹²

$$\begin{aligned}
c_i^* &= x_i' \eta_c + \theta_h h_i^* + v_i^c \\
h_i^* &= x_i' \eta_h + \theta_c c_i^* + v_i^h \\
c_i &= I(c_i^* > 0), h_i = I(h_i^* > 0) \\
i &= 1, \dots, N;
\end{aligned} \tag{5}$$

Causal responses are defined in terms of a contemporaneous effect of the latent index on the other decision. It is well known that identification of such model will necessitate exclusion restrictions and in a cross-sectional dataset it is difficult to think of plausible exclusions (Stern, 1995). With panel data available, one latent causal response model that is represented by the dynamic model we propose is given by

$$\begin{aligned}
c_{it}^* &= x_{it}' \eta_c + \theta_h h_{it}^* + \phi_c c_{it-1} + v_{it}^c \\
h_{it}^* &= x_{it}' \eta_h + \theta_c c_{it}^* + \phi_h h_{it-1} + v_{it}^h \\
c_{it} &= I(c_{it}^* > 0), h_{it} = I(h_{it}^* > 0) \\
i &= 1, \dots, N; t = 1, \dots, T
\end{aligned} \tag{6}$$

As we argued earlier, including lagged indicators is guided by the earlier literature on employment. For caring, it remains to be seen if such dynamics do exist. By solving this structural model for each latent index, we get the model in (1). In such case, causal

¹² Their model also includes an hours of work equation which we do not discuss here.

response as estimated by Soldo and Wolf (1994) can be recovered from the ratio of reduced form parameters

$$\begin{aligned}\theta_c &= \gamma_c / \delta_h \\ \theta_h &= \gamma_h / \delta_c\end{aligned}\tag{7}$$

Although potentially attractive, such structural interpretation would rely entirely on the exclusion of lagged causal effect from the structural model (h_{it-1} in eq. for c_{it} and vice-versa). Lagged employment status potentially affects future caregiving decisions and vice-versa. Therefore, such exclusion restrictions appear overly restrictive. Without other exclusion restrictions one cannot say if rejection of $H: \gamma_c = 0$ is due to a contemporaneous effect of the index of a true structural lagged effect. But one can conclude to a causal pathway from care to employment. This is the main point we wish to emphasize. A first step in understanding better the interaction between these variables is to identify causal pathways, whether they are contemporaneous or not. This strategy is a direct application of Granger causality applied to panel data. As Heckman (2000) notes, Granger causality interpreted in a structural sense would define a different causal parameter than what we would traditionally define as the causal effect of caring on employment in the current period. It defines the response of the probability of an outcome to a change in another outcome in the preceding year rather than in the current year. Trivially, estimating a contemporaneous effect would require an exclusion restriction.

Furthermore, defining a causal effect in terms of a response to a change in a latent index rather than the outcome itself is not a mere choice of convenience in such models. Schmidt (1981) shows how a simultaneous bivariate binary model (where both binary outcomes affect each outcome) has an incomplete probability representation when it comes to estimation of the parameters of interest. Although potentially a coherent representation of behavior, such model cannot be estimated by maximum likelihood and alternative solutions have only recently be proposed for cross-sectional models (Tamer, 2003). In the case of (5), if we replace latent dependent on the right hand side by observed binary outcomes and if causal parameters are negative, one can easily show that the sum of the probabilities or observed each joint outcome will be less than one leading to an incomplete probability representation.

There are two caveats with respect to the methodology we use. First, there is a range of parameters for which finding the absence of any causal pathway would not necessarily imply that there is no such pathways. In a traditional vector autoregression, the absence of a lagged response can be the result of the absence of dynamics themselves, therefore hiding potential contemporaneous effects. We would argue that this is not likely to be a problem in the current case, particularly for employment as from other studies we know there are strong and persistent dynamic effects. In the event where dynamics are not strong, the reduced form to the structural model (6) would have correlated transitory errors ε if contemporaneous causal effects were present. Of course, these could not be told apart without exclusion restrictions but the null of no causal pathway could still be tested ($H_0 : corr(\varepsilon_{it}^c, \varepsilon_{it}^h) = 0$).

The second caveat relates to the absence of expectations, or future values of the indicators in such model. For example, one could fear that future carers may decide to work more before a caring spell to smooth out the decrease in time available to work when they will have to care. We would argue that such expectation effects could in principle affect decisions of how many hours to work in a year if respondents can anticipate that they will care in future years. But whether they would go from not working to working in anticipation of such future caring spell appears less probable. One would need a strong intertemporal substitution motive to do that when there are presumably high adjustment costs to going in and out of the labor force every year. Furthermore, it is unclear how a respondent can predict future caring spells. With information on potential care recipients, this could in principle be assessed (by including future characteristics of potential care recipients), but this information is missing in the BHPS. In the end, we must note that if such effects were present, this would reverse the causality inferred from Granger causality tests. It would be future caring spells that would cause current employment and not caring that would be affected by past employment.

4.2 Estimation

The model proposed in (1) under the random effect formulation can be estimated by maximum simulated likelihood. However, because of the presence of both

lagged response variables and unobserved heterogeneity, ML estimates will be inconsistent when the time dimension of the panel is small relative to the cross-sectional size (Heckman, 1981). This is commonly known as the initial condition problem. We follow Alessie et al. (2004) by approximating the recursive solution to the first period outcomes to

$$\begin{aligned} c_{i0} &= I(x_{i0}'\beta_{c0} + \lambda_{cc}\alpha_i^c + \lambda_{ch}\alpha_i^h + u_{i0}^c > 0) \\ h_{i0} &= I(x_{i0}'\beta_{c0} + \lambda_{hc}\alpha_i^c + \lambda_{hh}\alpha_i^h + u_{i0}^h > 0) \end{aligned} \quad (8)$$

No restrictions are imposed on the way x is related to initial outcomes and the only link across initial condition equations and those of other period is the presence of unobserved heterogeneity. In particular x_{i0} loads with different parameters than in future periods such that this represents a flexible approximation to the distribution of initial outcomes. Eq. (5) makes clear that without unobserved heterogeneity, there is no initial condition problem. In that case, the log-likelihood of initial conditions is separable from the one from period 1 onwards. If unobserved heterogeneity is present, the contributions to the log-likelihood are not independent.

The joint likelihood derived in Appendix B is simulated using the GHK (Geweke, Hajivassiliou and Keane) simulator (see Hajivassiliou et al., 1996). We simulate the likelihood using 40 Halton draws which are known to reduce the variance and improve the coverage of the random draws (Train, 2002).¹³ We then maximize this simulated likelihood over the parameter space using the BFGS subroutine available in Ox. Average partial responses are computed for each variable by integrating over the estimated distribution of unobserved heterogeneity and standard errors are obtained by Monte Carlo simulation using the estimated distribution of the parameters.

In the fixed effect formulation which uses the linear probability models, we use moments of the form

¹³ Increasing the number of draws did not change results. A normal rule of thumb is that 10 times less Halton draws for the same mean squared error criterion is needed (Train, 2002).

$$\begin{aligned}
E(\Delta \varepsilon_{it}^k c_i^{t-2}) &= 0, E(\Delta \varepsilon_{it}^k h_i^{t-2}) = 0 \\
E(\Delta \varepsilon_{it}^k \Delta x_{it}) &= 0 \\
k &= c, h, t = 1, \dots, T \\
y_i^{t-s} &= (y_{it}, \dots, y_{it-s})'
\end{aligned} \tag{6}$$

as proposed by Arrellano and Bond (1991) to estimate by generalized method of moments the parameters in (1). We make use of mean stationarity restrictions to complement the estimation in order to improve on the efficiency of the estimator (Blundell and Bond, 1998). We estimate each equation separately ignoring the potential gains in efficiency from estimating them jointly. These linear probability models have the usual caveats (unbounded probabilities, heteroscedasticity, etc.) but provide robust alternatives to check on the non-linear results that depend to a greater extent on the functional forms assumed and the strict exogeneity assumption.

4.3 Specification and Stratification

We include standard demographics such as age, cohort and education attainment in the vector of observable characteristics x . In addition, we include the number of kids in the household of different age groups as childcare will presumably interfere with the capacity to provide informal care to relatives, particularly for women. We also include household non-labor income as a measure of other resources which can be used to buy formal care but also affect labor supply. Furthermore, we include housing status which in addition to non-labor income and education attainment will capture differences across the socio-economic ladder. Following Hyslop (1999), the non-labor income measure is included in two forms. One is a permanent non-labor income consisting of the individual mean (over time) in household non-labor income and the other is the per-period non-labor income. This attempts to control for the potential endogeneity between non-labor income and unobserved heterogeneity. The assumption is that heterogeneity will be correlated with the permanent component but not with the transitory part. This general identification strategy follows from Chamberlain's (1984) suggestion of using correlated random effects to ease up the restrictions imposed by random effect models.

We stratify the estimations by gender for two reasons. First, there are reasons to believe that there are genuine differences between male and female as to why and how care is provided to relatives. There are also potentially gender differences in the employment dynamics. For example, both the level and the persistence of employment for men is much stronger than for women. Transitions out of work for men are usually towards unemployment, for short periods, while for women it is more likely to reflect exits out of the labor force. The second reason is that both spouses of a couple are present in the data so that we wish to abstract from having to consider the obvious correlation in their decisions. Even by stratifying the estimation, we are still potentially neglecting important interactions in the household as to who provides care and how the other spouse responds to this. Engers and Stern (2002) look at those decisions in a bargaining framework. However, its interaction with employment decisions of the family network beyond the scope of the present paper. The BHPS, given its limited information on relatives living outside the household is not the adequate dataset to model such interactions. More extensive information could be obtained from the English Longitudinal Survey of Ageing (ELSA) which is however a cross-sectional dataset for the time being (as of January 2006).

5. Discussion of Results

We first discuss results for caregiving (either co-residential or extra-residential) before looking into the desegregation. Marginal effects and other parameter estimates along with their standard errors are presented for females in Table 7 and Table 8 for males. Initial condition parameter estimates can be found in Appendix C.

Observable characteristics associated with employment probabilities show no surprise and are consistent with other studies using a similar specification for the U.K. and other countries (Hyslop, 1999; Michaud and Tatsiramos, 2005). Employment probabilities are usually lower for those with higher permanent household non-labor income either showing a true income effect or the association between taste for work and saving at the household level. The transitory component is also negative and significant for both males and females reflecting potential true adjustments of labor supply to

changes in non-labor income. Health problems are negatively associated with employment will the usual positive education gradient is observed. Finally, as expected, the presence of children in young age (less than 4 years old) as quite a bearing on female's employment possibilities while it has no association with employment for males.

As for the characteristics associated with caring decisions, the education gradient appears stronger for females. Females with graduate degrees tend to care much less than those with no studies. In part, this probably reflects a socio-economic gradient but potentially also the fact that highly educated individuals have probably higher opportunity costs of time and therefore more likely to use private formal care. Interestingly this gradient does not exist for males. Permanent non-labor income is positively associated with care. Although it is unclear whether this relationship is causal, the fact that employment is negatively associated with non-labor income leaves room for other activities to be substituted for leisure. The original model by Becker (1965) certainly predicts that in most cases, positive non-labor income shocks diverts time from market activities towards home production (which may include informal care as seen in section 2). For males, caring also appears to be positively associated with bad health (marginal effect = 0.0178 percentage points, s.e = 0.008). Although marginally statistically significant, this association may suggest that opportunity costs really play a role in the caring decision. Alternatively, it could also be the result of the correlation of health across generations. Those who have parents in bad health may be more likely to be in bad health themselves through various channels including socio-economic status and other health behaviors transmitted across generations.

More importantly, however, the results suggest no effects of care on work and only modest effects from employment to informal care. For example, women who are working in the current year have a 2.1 percentage points lower probability to care in the next period. Since the average caring propensity for a female is 16%, this implies a 13.1% relative difference in the probability to care next year. The effect for males is similar in magnitude. This is because although the lagged employment effect appears smaller, the caring state-dependence effect is also small such that the contemporaneous effect is much larger.

Unobserved heterogeneity, particularly family related characteristics and care recipient characteristics are likely to be related to both decisions and since those are generally time-invariant (how many siblings, etc.) it is important to allow a correlation between those factors across decisions. Since the variance of the error term in a binary model is not identified, it has been normalized to one. Hence, we estimated the share of this variance due to transitory unobservables assumed uncorrelated over time. For both care and work and males and females, the share of the variance due to transitory shocks is large (between 0.72 and 0.85 with small standard errors) but still leaves a considerable part for time-invariant unobservables to explain the variance across respondents. Surprisingly, we do not estimate a sizeable nor precise correlation between these unobservables across decisions ($\rho_{UH} = -0.053$ for females and 0.016 males). This would suggest that a large part of the cross-sectional correlation across decisions is already captured by observable characteristics included in the specification. Therefore, this suggest that causality is probably not due to omitted third factors that do not vary over time.

However, for males, we do find a negative and significant correlation between the transitory shocks ($\rho_{\epsilon} = -0.087$, $se = 0.036$). This can be due to at least two factors. First, the reduced-form model in (2) suggest that even if the structural errors v are not correlated, the transitory shocks u can be correlated if causal effects are present. Therefore, one could interpret this correlation as potentially indicating effects that operate contemporaneously or simply a correlation in structural shocks.

Since approximately 20 percent of the variance in unobservables is due to time-invariant omitted factors, this explains some of the persistence in both employment and caring spells. But state-dependence appears equally if not more important. We estimate that females employed in a given year, have a 42 percentage points ($se = 0.015$) higher probability to be employed next year compared those out of the labor force. The estimate for males is almost the same. As for state-dependence in caring spells, the estimates suggest that it is considerably important given the relatively low average caring propensities. Women have a 24.8 percentage points higher chance to care in the next period compared to those not currently caring. For males, the estimate is somewhat

smaller (15.6 percentage points) although males have in general lower caring propensities than females such that in relative terms these effects may be quite similar.

As mentioned previously, there are reasons to believe that these results may differ by care type since these decisions may involve different considerations and different abilities to conciliate work along with caregiving. Ideally, since there are three states (co-residential, extra-residential and no care), one would model all three states simultaneously. However, given the relatively small number of transitions between care types this is not feasible with the current set of data. Instead we focus on the same model as in (2) where we simply look at each decision separately and its interaction with the employment outcome. Given the low cross-equation correlations for the overall care category discussed above this seems a valid alternative. In doing so, we mix in the pool of non-carers, those that do not care and those that provide another type of care. However imperfect, this strategy is still likely to reveal differences across groups of carers regarding their dynamics with employment outcomes.

The results in Table 9 to 12 suggest important differences in the dynamics depending on the type of care provided. For females, co-residential carers appear to have lower probabilities of being employed in the following year (-5.9 percentage points, $se = 0.024$). This is also found for males although it is estimated with much less precision. Because co-residential carers account for a small group of carers (less than 5%), this effect was likely to be hidden in the overall effect of caregiving. Indeed, results in Table 11 suggest no effect for extra-residential female carers (0.9 percentage points, $se = 0.012$) and also male carers. As for the effect of employment on the probability to provide care, these are more stable across type of care but lack precision in general (e.g. female co-residential -0.78 percentage points, extra-residential -1.1 percentage points).

To check on the robustness of these results, we also estimated linear probability models. Results for the causal effects can be found in Table 13 and 14. Although the effects generally compare well with those estimated using the random effect model, the Sargan test of overidentifying restrictions rejects strongly the specifications. In one case, we even estimate a positive effect of employment on caring for men while we found a negative effect in the random effect models. These differences and the general rejection of the linear probability model can be due to both the linearity assumed but also

misspecification of the dynamics. However, we do not pursue any refining the dynamics here because it is relatively difficult to expand the binary choice models in the current setting to one that features higher order dynamics. It is also unclear if the potential need for further lags in the linear probability model is not simply due to the misspecification of the functional form. Finally, because these linear probability models use lags of a binary variable as instruments for the dynamics and these dynamics appear to be persistent, it is possible that these are relatively weak instruments for the first order dynamics that we assumed.

Taken together the results tend to suggest that co-residential carers are the only group for which robust evidence of causal pathways is found. For extra-residential carers, the correlation in unobservables in addition to the causal pathways are estimated imprecisely. Although it appears to run both ways for co-residential carers, in relative terms, the effect of employment on caregiving appears somewhat stronger. Since co-residential carers are a small group it remains to be seen how much of an impact the causal mechanisms have on the association between caregiving and employment at the aggregate level. Next, we use a micro-simulation exercise to gauge these aggregate effects although we anticipate the effects to be small. The dynamic representation of the model can reveal important compositional differences that may change simulated outcomes.

6. Simulations

In order to grasp the magnitude of the estimated pathways and see their importance for public policy, we simulate the outcomes of two changes in 2004 and follow their dynamic impact over time. Given the above results we focus on co-residential caring only.¹⁴ Three simplifying assumptions are made throughout the simulation. First, fertility is held constant, i.e. no births occur after 2003 in our sample. This assumption is not expected to interfere much with the simulated effects since most carers are aged 40 years and older (yet we still allow children born before 2003 to affect caring and employment probabilities). Second, we keep residency fixed so that there is no transition from home ownership to rent and vice-versa. Finally, transitory non-labor

¹⁴ Results for overall carers and extra-residential carers can be obtained from the authors upon request.

income is set to zero. Each run takes the following form: the simulation is started in 2004. The change introduces a permanent increase of 10 percentage points in *individual* probabilities of employment starting in 2006. Because of the dynamic multiplier effect, such changes can decrease the provision of co-residential care. One can think of this change as one that would stimulate employment perhaps following new targets set by the government.¹⁵ The other simulation is one where we increase co-residential care by 10 percentage points starting in 2004. Though substantial, this change would be consistent with the increased demand created by ageing if formal care does not step up to meet these needs.

Figures 4 and 5 report the effect on employment of raising co-residential caregiving by 10 percentage points for both males and females. The effect on employment is quite small (around 1 percentage points in 2010). Since this group is small, the aggregate effect is not surprising. Therefore it is not likely that this can lead to substantial changes in employment rates. For example, a 5 percentage point decrease in employment would require an approximately 50 percentage points increase in co-residential caregiving. Since employment is not responsive to extra-residential caregiving, the overall pathway going from caregiving to work is bound to be small.

Figure 5 and 6 show the simulated effects of raising employment rates on caregiving. Even more so than with the reverse pathway, there is virtually no change in the provision of co-residential caregiving. Since the effect for co-residential carers is again small, the aggregate effect is even smaller. Therefore, these simulation results do not indicate a big tradeoff between employment and caring at the aggregate. However, it emphasizes that some groups are more affected than others (co-residential carers).

¹⁵ Each simulation run takes account of three sources of uncertainty. The first refers to the transitory shocks that affect both decisions. The second refers to the unknown type (unobserved heterogeneity) of each respondent and the last refers to the parametric uncertainty given that parameters of the model are estimated. We consider these three types of uncertainties when computing the share of respondents employed and caring in each year. This allows computing standard errors for the simulation and to form confidence bands for these predictions. The mean predicted outcomes and associated standard errors are estimated from Monte Carlo simulations using the estimated parameters that quantify the uncertainty. In order to keep the graphical representation simple, standard errors are not reported but have generally been found to be very small.

7. Conclusion

This paper set out to study the link between employment and informal care provision and in particular disentangle cause and effect. From a policy maker's point of view it is vital to understand the interdependency between informal care and the labour market for at least two reasons. First, most industrial countries face an ageing population which will increase the demand for care unless longer life expectancies are synonymous to healthier life expectancies. At the same time formal care budgets are on the decline. Second, the Lisbon Agenda aspires European countries to increase labour market participation particularly female participation. The British government cherishes an even more ambitious aspiration of an 80 percent employment rate in the years to come. Hence, if there is an association between informal care provision and employment ensuring sufficient care and increasing employment rates are not to be considered independent policy options.

Employing a dynamic simultaneous equation framework we find no link between overall informal care provision and participation in paid employment. The same holds true for the large number of extra-residential carers. However, we find a significant link between employment and co-residential caring. In particular, there is evidence that suggests that caring reduces employment probabilities by up to 6 percentage points. Yet, employment has only a very moderate effect on co-residential care provision.

Using a set of care and employment reform scenarios to simulate the dynamics over time we find that relatively large changes to individual care and employment probabilities lead to disproportionately smaller aggregate effects on care and employment.

Despite its size, co-residential carers are a key care group as they are more likely to provide intensive care both in terms of hours and tasks. Consequently they are less likely to exercise choice over their care provision compared to the large group of the extra-residential cares which often provides relatively basic support to friends and relatives. Hence, it is not surprising to find the strongest association between care and employment for co-residential carers. Though the interdependencies between work and care shown in this paper are moderate, this may change in the years to come as a result of

an increase in the demand for informal care either due to demographic and health effects or the a continuation in the decrease in formal care provision.

These findings do not contradict earlier qualitative studies that found a strong link between care provision and employment. The quantitative results in this paper reflect on the average carer rather than selected individual cases. Hence, even though the informal care population as a whole in England does not suffer substantial employment disadvantages on average, some individual carers might do warranting targeted policy interventions.

Despite the clear merits of our specification, there are several caveats. First, we model individual care and employment decisions treating demand for care – which may be the result of family bargaining – as given. The BHPS provides limited information on the care receiver and other comprehensive longitudinal UK data is scarce. Yet, comparing carers and non-carers who do face a demand for care may be a fruitful exercise and may provide additional and relevant policy insights. Second, results in this paper hold in first-order dynamic models we have estimated. Some evidence is presented in the linear probability models that these dynamics may be misspecified. Finally, this paper has focused on the extensive margins of employment and care though different care types are distinguished which are correlated with care intensity. Further work may explore the causal impact of informal care on hours worked and the intensity of care on employment.

References

- Alessie, R., S. Hochguertel and A. van Soest (2004): "Ownership of Stocks and Mutual Funds: A Panel Data Analysis", forthcoming *Review of Economics and Statistics*.
- Arellano, M. and S. Bond (1991): Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations, *Review of Economic Studies*, 58, pp. 277-297.
- Arulampalam, W. (1999): "A Note on Estimated Coefficients in Random Effect Probit Models", *Oxford Bulletin of Economics and Statistics* 61, pp. 597-602.
- Becker (1965): "A Theory of the Allocation of Time", *Economic Journal*, vol 75, no. 299.
- Bell, D. and A. Heitmueller (2006): "Free Personal Care in Scotland: What happened to informal care provision?" University of Stirling mimeo.
- Blundell, R. and S. Bond (1998): "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models", *Journal of Econometrics*, 87, pp. 115-143.
- Carmichael, F. and Charles, S. (1998): "The labour market costs of community care", *Journal of Health Economics* 17 (6), pp. 747-765.
- Carmichael, F. and Charles, S. (2003): "The opportunity costs of informal care: does gender matter?", *Journal of Health Economics* 22 (5), pp. 781-803.
- Chamberlain, G. (1985): "Heterogeneity, Omitted Variables and Duration Dependence", in *Longitudinal Analysis of Labor Market Data*, eds, Heckman, J.J. and Singer, B., Econometric Society Monograph chapter 1, pp. 3-38.
- Chamberlain, G. (1984): "Panel Data", *Handbook of Econometrics II*, eds. Z. Griliches, M. Intrilligator, Elsevier.
- Coward, R. T. and Dwyer, J. W. (1991): "A Multivariate Comparison of the Involvement of Adult Sons Versus Daughters in the Care of Impaired Adults." *Journal of Gerontology: Social Sciences* 46(5), pp 259-269.
- Department for Health (2003): *Community Care Statistics 2003*
- Department for Work and Pensions (2005): *Carer's aspirations around work and retirement. Research Report No 290*
- Engers, M. and S. Stern (2002): "Long-term care and family bargaining", *International Economic Review* 43(1), pp. 73-113.

- Granger, C.W.J. (1969): Investigating Causal Relations by Econometric Models and Cross-Spectral Methods, *Econometrica*, 37, pp. 424-438.
- Grundy E. (2000): "Co-residence and mid-life children with their elderly parents in England and Wales: Changes between 1981 and 1991", *Population Studies* 54:193-206.
- Hajivassiliou, V., McFadden, D. and P. Ruud (1996): "Simulation of Multivariate Normal Rectangle Probabilities and their Derivatives: Theoretical and Computational Results", *Journal of Econometrics* 72, pp. 85-134.
- Heckman, J.J. (1981): "The Incidental Parameters Problem and the Problem of Initial Condition in Estimating a Discrete-Time Data Stochastic Process", in *Structural Analysis of Discrete Data with Econometric Applications*, C.F. Manski and D. McFadden eds. , MIT Press, Cambridge, pp. 179-195.
- Heckman, J.J. and Willis, R. (1977): "A Beta-Logistic Model for the Analysis of Sequential Labor Force Participation by Married Women", *Journal of Political Economy*, 85(1). pp. 27-58.
- Heitmueller A. (2005a): "The Chicken or the Egg? Endogeneity in labour market participation of informal carers in England", mimeo, London Business School
- Heitmueller A. (2005b): "Spell-Persistence of Informal Care Related Mental Distress in Britain 1991-2003", mimeo, London Business School
- Honore, B.E. and Kyriazidou, E. (2000): "Panel Data Discrete Choice Models with Lagged Dependent Variables", *Econometrica*, 68, pp. 839-874.
- Hyslop, D.R. (1999): "State Dependence, Serial Correlation and Heterogeneity in Intertemporal labour Supply Participation of Married Women", *Econometrica* 67:6, pp.1255-1294.
- Ilmakunnas, S. and S. Pudney (1990), "A model of female labour supply in the presence of hours restriction", *Journal of Public Economics* 41: 183-210.
- Laing and Buisson (2002): *UK Domiciliary Care Markets 2002*. Fifth Edition, Laing & Buisson, London
- Lakdawalla, D. and Philipson, T. (2002): "The Rise in Old-Age Longevity and the Market for Long-Term Care", *American Economic Review*, 92(1), pp: 295-306
- Leger, P.-T. (2000): "Parental Illness and the Labour Supply of Adult Children" SEDAP Research Paper 21
- Michaud, P.-C. and K. Tatsiramos (2005): "Employment Dynamics of Married Women in Europe", RAND Working Paper 213.

Neyman, J. and Scott E. L. (1948): "Consistent Estimation from Partially Consistent Observations." *Econometrica*, 16, pp: 1-32

OECD (2005): "Ensuring quality long-term care for older people". Policy Brief March 2005

ONS (2002): Carers 2000. Office for National Statistics, London

Soldo B. J. and Wolf D.A. (1994): "Married Women's Allocation of Time to Employment and Care of Elderly Parents", *Journal of Human Resources*, 29, pp. 1259-1276

Stern, S. (1995). "Estimating Family Long-Term Care Decisions in the Presence of Endogeneous Child Characteristics." *Journal of Human Resources* 30(3): 551-580.

Train, K.E. (2002): Discrete Choice Methods with Simulation, Cambridge University Press, 321 p

Van Soest A., Woittiez I. and A. Kapteyn (1990) "Labour supply, income taxes and hours restrictions in the Netherlands", *Journal of Human Resources* 25: 517-558.

Appendix A Derivation of the Analytic File

Of the 156,211 person-year interviews between 1991 and 2003, we drop 33,274 person-year observations that are younger than 16 or older than 59 (women) and 65 (men). We drop 42,597 person-year observations from Scotland and Wales when restricting the sample to England. We also drop military workers (97 person-year) and self-employed workers (7,177 person-year). We do not consider 13,042 observations from individuals who have missing intermediate years in their interview sequence. We drop 630 for missing health information, 236 person-year for missing housing information and 6 person-years for missing income, 1192 person-years with missing education, 24 person-year for which we cannot determine caring status. The final sample has 53,198 person-year observations. This is the sample used for descriptive statistics. For estimation, we only consider respondents who remain in the panel for more than 3 year. 4,998 observations are lost from that requirement.

Table 1: Panel Entry and Exit

year	Panel Status				Total
	Enter>1	Enter=1	Continue	Exit	
1991	3,673 85.60	618 14.40	0 0.00	0 0.00	4,291 100.00
1992	364 8.59	202 4.77	3,296 77.75	377 8.89	4,239 100.00
1993	255 6.28	144 3.55	3,329 82.02	331 8.15	4,059 100.00
1994	237 5.98	139 3.51	3,257 82.25	327 8.26	3,960 100.00
1995	244 6.33	118 3.06	3,289 85.30	205 5.32	3,856 100.00
1996	286 7.18	163 4.09	3,295 82.75	238 5.98	3,982 100.00
1997	727 15.94	253 5.55	3,349 73.43	232 5.09	4,561 100.00
1998	275 6.15	124 2.77	3,713 82.97	363 8.11	4,475 100.00
1999	185 4.41	26 0.62	3,732 88.88	256 6.10	4,199 100.00
2000	281 6.40	195 4.44	3,186 72.52	731 16.64	4,393 100.00
2001	265 6.83	150 3.86	3,161 81.43	306 7.88	3,882 100.00
2002	209 5.59	101 2.70	3,049 81.61	377 10.09	3,736 100.00
2003	307 8.61	0 0.00	3,258 91.39	0 0.00	3,565 100.00
Total	7,308 13.74	2,233 4.20	39,914 75.03	3,743 7.04	53,198 100.00

Notes: See Appendix A for details on the construction of the analytic file. Enter=1 refers to respondents entering the panel and leaving next wave. Enter>1 refers to those who go on to stay for more than one wave.

Appendix B Log Likelihood Derivation

Application of the GHK simulator requires that we derive the error distribution of the composite error term $u_0 = (u_0^1, \dots, u_0^T)'$ where $u_t = (\alpha^h + \varepsilon_t^h, \alpha^c + \varepsilon_t^c)'$ and

$$u_0 = (\lambda_{hh}\alpha^h + \lambda_{hc}\alpha^c + \varepsilon_0^h, \lambda_{ch}\alpha^h + \lambda_{cc}\alpha^c + \varepsilon_0^c)'$$

We partition the covariance matrix into four region

$$\Omega_u = \begin{bmatrix} \Sigma_0 & \iota_T' \otimes \Sigma_s \\ \iota_T \otimes \Sigma_s & \Sigma_{-0} \end{bmatrix}$$

Where Σ_0, Σ_s are 2x2 matrices corresponding to $E u_0 u_0'$ and $E u_0 u_s'$ for $s=1, \dots, T$, ι_T represents a column vector of ones and Σ_{-0} is a $2T \times 2T$ matrix standing for $E(u_{-0} u_{-0}')$.

For $t > 0$, the covariance matrix takes the simple form

$$\Sigma_{-0} = I_T \otimes \Omega_\varepsilon + J_T \otimes \Omega_\alpha$$

With I_T , the identity matrix of dimension T and $J_T = \iota_T \iota_T'$ whereas the covariance matrices of each component of the error term are given by

$$\Omega_\varepsilon = \begin{bmatrix} \sigma_{\varepsilon^h}^2 & \cdot \\ \rho_\varepsilon \sigma_{\varepsilon^h} \sigma_{\varepsilon^c} & \sigma_{\varepsilon^c}^2 \end{bmatrix}, \Omega_\alpha = \begin{bmatrix} \sigma_{\alpha^h}^2 & \cdot \\ \rho_\alpha \sigma_{\alpha^h} \sigma_{\alpha^c} & \sigma_{\alpha^c}^2 \end{bmatrix}$$

given the normality assumptions. Since such models are identified up to scale we let $\sigma_{\alpha^j}^2 + \sigma_{\varepsilon^j}^2 = 1$, such that each of these variances represent their share in the variance of u . As for initial condition errors it directly follows from their form that $\Sigma_0 = \lambda \Omega_\alpha \lambda' + \Omega_{\varepsilon_0}$ where

$$\lambda = \begin{bmatrix} \lambda_{hh} & \lambda_{hc} \\ \lambda_{ch} & \lambda_{cc} \end{bmatrix}, \Omega_{\varepsilon_0} = \begin{bmatrix} 1 & \cdot \\ \rho_0 & 1 \end{bmatrix}.$$

The remaining part is $\Sigma_s = \lambda \Omega_\alpha$. Since each respondent is in the panel for T_i , the covariance matrix will be different for each respondent. Denote the observed decision vector

$$y_i = (h_{i0}, c_{i0}, \dots, h_{iT_i}, c_{iT_i})'$$

The probability of observing that sequence given the process $x_i = (x_{i0}, \dots, x_{iT_i})$ and parameters is

$$P(y_i | x_i; \theta) = \int_{u \in B(x_i, \theta)} dF(u_i; \Omega_{ui})$$

where $B(x_i, \theta)$ represents the set of errors consistent with observed choices. We use the GHK simulator to simulate $P(y_i | x_i; \theta)$ which we then plug in the maximum likelihood estimator. That estimator has been shown to be consistent for R the number of replications going to infinity. It is asymptotically normal and asymptotically equivalent to maximum likelihood for R increasing slower than the square root of the sample size.

Appendix C Initial Condition Results

Female Carers

	Work Eq.		Caring Eq.	
	P.Est	S.Dev	P.Est	S.Dev
Constant	0.648	0.0829	-1.38	0.0967
age 30-40	0.448	0.0864	0.163	0.0937
age 40-50	0.413	0.0887	0.393	0.0861
o level	0.445	0.0789	-0.185	0.0854
a level	0.729	0.113	-0.294	0.125
degree	1.12	0.13	-0.183	0.124
high d	1.21	0.404	-0.392	0.397
kid 0-4	-0.962	0.085	-0.265	0.106
kid 5-11	-0.437	0.0756	-0.0816	0.0857
rent	-0.441	0.0706	0.0354	0.0813
hlimit w	-0.565	0.101	0.374	0.108
permanent	-0.558	0.0914	0.369	0.091

Female Co-Residential Carers

	Work Eq.		Caring Eq.	
	P.Est	S.Dev	P.Est	S.Dev
Constant	0.66	0.0846	-2.96	0.258
age 30-40	0.454	0.088	0.188	0.176
age 40-50	0.422	0.0903	0.385	0.156
o level	0.45	0.0802	-0.337	0.152
a level	0.745	0.116	-0.543	0.25
degree	1.14	0.133	-0.501	0.251
high d	1.24	0.413	-3.87	19.8
kid 0-4	-0.979	0.0871	-0.373	0.217
kid 5-11	-0.442	0.0768	-0.0618	0.158
rent	-0.451	0.0719	0.505	0.144
hlimit w	-0.569	0.103	0.256	0.183
permanent	-0.569	0.093	0.349	0.179

Female Extra-Residential Carers

	Work Eq.		Caring Eq.	
	P.Est	S.Dev	P.Est	S.Dev
Constant	0.645	0.0826	-1.52	0.105
age 30-40	0.446	0.0862	0.132	0.0998
age 40-50	0.411	0.0885	0.331	0.091
o level	0.444	0.0786	-0.155	0.0906
a level	0.727	0.113	-0.272	0.133
degree	1.12	0.129	-0.0605	0.128
high d	1.2	0.403	-0.219	0.4
kid 0-4	-0.961	0.0847	-0.214	0.113
kid 5-11	-0.436	0.0753	-0.0976	0.0918
rent	-0.437	0.0703	-0.13	0.0885
hlimit w	-0.562	0.101	0.367	0.115
permanent	-0.555	0.091	0.335	0.0931

Male Carers

	Work Eq.		Caring Eq.	
	P.Est	S.Dev	P.Est	S.Dev
Constant	0.782	0.0927	-1.6	0.111
age 30-40	0.821	0.111	-0.0425	0.116
age 40-50	0.834	0.116	0.277	0.104
o level	0.203	0.0887	-0.278	0.106
a level	0.677	0.122	-0.337	0.133
degree	0.9	0.138	-0.184	0.136
high d	0.716	0.293	-0.512	0.312
kid 0-4	0.141	0.115	-0.235	0.138
kid 5-11	-0.113	0.0936	0.00465	0.104
rent	-0.729	0.0852	0.122	0.0968
hlimit w	-1.15	0.114	0.384	0.117
permanent	-0.536	0.0887	0.365	0.0938

Male Co-Residential Carers

	Work Eq.		Caring Eq.	
	P.Est	S.Dev	P.Est	S.Dev
Constant	0.792	0.095	-2.87	0.246
age 30-40	0.828	0.114	0.00957	0.203
age 40-50	0.853	0.119	0.336	0.175
o level	0.208	0.0903	-0.404	0.18
a level	0.695	0.124	-0.619	0.239
degree	0.922	0.141	-0.406	0.245
high d	0.735	0.297	-0.782	0.588
kid 0-4	0.149	0.117	-0.363	0.25
kid 5-11	-0.118	0.0951	0.1	0.178
rent	-0.738	0.0875	0.437	0.155
hlimit w	-1.16	0.117	0.71	0.176
permanent	-0.546	0.0906	0.243	0.177

Male Extra-Residential Carers

	Work Eq.		Caring Eq.	
	P.Est	S.Dev	P.Est	S.Dev
Constant	0.78	0.092	-1.81	0.127
age 30-40	0.814	0.111	-0.092	0.129
age 40-50	0.821	0.115	0.209	0.114
o level	0.199	0.0883	-0.207	0.116
a level	0.674	0.121	-0.228	0.144
degree	0.896	0.137	-0.122	0.147
high d	0.712	0.292	-0.244	0.319
kid 0-4	0.141	0.114	-0.191	0.154
kid 5-11	-0.113	0.0932	-0.0381	0.117
rent	-0.721	0.0844	-0.0172	0.11
hlimit w	-1.15	0.113	0.305	0.13
permanent	-0.534	0.088	0.404	0.0969

Tables and Figures

Table 2: Characteristics of Carers and Non-Carers

Variables	non-carers	carers	total
health limits work	0.10	0.17	0.11
Married	0.56	0.70	0.58
No Education	0.23	0.33	0.25
O-level Education	0.39	0.37	0.38
A-level Education	0.19	0.16	0.19
Degree Education	0.17	0.13	0.16
Higher degree Education	0.02	0.02	0.02
At least one kid aged 0-4	0.14	0.08	0.13
At least one kid aged 5-11	0.21	0.19	0.21
At least one kid aged 12-15	0.16	0.17	0.17
At least one kid aged 16-18	0.06	0.05	0.06
Household size	3.11	3.09	3.10
House owned	0.77	0.77	0.77
Renter	0.21	0.21	0.21
N	45,999	7,199	53,198

Notes: Authors' calculations from the BHPS. Sample weights used for the pooled sample (1991-2003)

Table 3: Economic Resources of Carers and Non-Carers

Annual	Non-carers		Carers	
	median	%>0	median	%>0
Household Income				
Earning	26,945	0.92	23,689	0.87
Benefit	1,369	0.63	3,482	0.70
Transfer	1,632	0.10	999	0.10
Capital	338	0.70	485	0.72
Total (equivalized)	22,522	100	20,136	100

Notes: Annual Income expressed in 2003 British Bounds using the RPI. Total income is equivalized using the OECD equivalence scale (0.5 additional adults, 0.3 for child). Sample weights used.

Table 4: Hours worked and Informal Care

hours	Males		Females	
	non-carers	carers	non-carers	carers
no hours	20.3	30.0	27.1	35.7
0-30 hours	1.2	1.4	7.7	8.3
30+ hours	43.3	34.7	20.3	16.7
hours missing	35.3	33.9	45.0	39.4
Total	100	100	100	100

Notes: Hours missing refers to respondents that report having a paid job but do not report hours. Sample weights used.

Table 5 Probit Marginal Effects for Caring

caring	Females		Males	
	M.E.	t-stat	M.E.	t-stat
age 31-40	0.077	5.95	0.028	2.16
age 41-50	0.174	9.28	0.104	5.72
age 51+	0.225	9.19	0.122	5.55
(birth year-1900)/10	-0.103	-1.75	-0.038	-0.78
o level	0.013	1.15	0.012	0.99
a level	0.007	0.43	0.014	1.11
degree	-0.031	-2.32	-0.003	-0.2
g degree	-0.056	-1.72	-0.024	-1.08
health limits work	0.028	2.29	0.025	2.05
kids 0-4	-0.012	-1.33	-0.020	-2.16
kids 5-11	0.018	2	0.006	0.77
own house	0.026	1.04	0.038	1.68
rent house	0.004	0.18	0.020	1.04
non-labor income	-0.002	-0.37	0.001	0.19
permanent nl income	0.028	2.42	0.026	2.57
participation	-0.038	-4.39	-0.021	-2.15
N	28620		24578	
obs. P	0.16		0.11	
pseudo R2	0.073		0.051	

Notes: Maximum likelihood probit estimates. Marginal effects computed at the mean for continuous regressors and as discrete differences for discrete regressors. Errors are clustered at the respondent level to reflect the grouped nature of the data.

Table 6: Informal Care Dynamics

	female	male	total
% carer->non-carer	32.3	39.1	34.8
% non-carer->carer	6.3	5.1	5.7
% caring once	42.8	32.8	38.2
% years caring			
1-10%	20.7	22.1	21.3
11-20%	14.5	19.7	16.6
21-30%	13.0	13.7	13.3
31-40%	13.5	12.1	13.0
41-50%	7.5	5.1	6.5
51-60%	6.9	7.3	7.1
61-70%	8.2	6.4	7.4
71-80%	3.7	4.4	4.0
81-90%	3.5	3.1	3.3
91-100%	8.5	6.3	7.6
# transitions			
0	10.2	7.5	9.1
1	31.7	32.9	32.2
2	30.8	31.1	31.0
3	12.3	13.7	12.8
4	8.0	8.2	8.1
5	4.1	3.6	3.9
6	1.7	2.3	2.0
7	0.8	0.4	0.6
8	0.5	0.2	0.4

Notes: % carers-> non-carers report conditional relative fraction going from carer to non-carer in the next wave. % caring once reports the fraction who report being carer once over the years. Conditional on caring once, the next rows report the fraction of years respondents were carers. Finally, the distribution of the # of transitions is reported. Sample weights used.

Table 7 Results for Work and Caring - Females**Transition Equation Marginal Effects**

	Work Eq.		Caring Eq.	
	P.Est	S.Dev	P.Est	S.Dev
age 30-40	0.0293	0.0123	0.0564	0.0101
age 40-50	0.0595	0.0175	0.112	0.0166
age >50	-0.0156	0.0249	0.128	0.0219
byear/10	0.142	0.075	-0.161	0.104
o level	0.0986	0.0125	-0.00368	0.00786
a level	0.0913	0.0126	-0.000275	0.00984
degree	0.154	0.011	-0.0306	0.0089
high d	0.155	0.0162	-0.0575	0.0145
kid 0-4	-0.219	0.0113	-0.00607	0.00665
kid 5-11	-0.0482	0.00936	0.00919	0.00609
rent	-0.0973	0.0108	0.0113	0.00709
hlimit w	-0.203	0.0146	0.00844	0.00725
nlinc/1000	-0.0259	0.006	-0.00721	0.00989
permanent	-0.0995	0.0151	0.0755	0.0205

Lagged Variable Marginal Effects

	Work Eq.		Caring Eq.	
	P.Est	S.Dev	P.Est	S.Dev
work(t-1)	0.422	0.0158	-0.0208	0.00748
care(t-1)	-0.0115	0.0126	0.248	0.0127

Covariance Parameters (Work,Care)

	P.Est	S.Dev
sh.Trans(h)	0.79	0.0147
sh.Trans(c)	0.769	0.0136
rho UH	-0.0529	0.0539
rho eps	-0.0259	0.0275
rho Init	-0.0441	0.0655
L(h,c)	1.37	0.102
L(h,h)	0.0948	0.0898
L(c,h)	-0.109	0.0847
L(c,c)	1.3	0.101
LogLike/N	-5.14	

Notes: Maximum simulated likelihood estimates using the BFGS algorithm and the GHK simulator with 40 Halton draws along each dimensions. Intercepts not included. Marginal Effects are computed at mean characteristics. For discrete outcomes, they are computed using discrete Differences in the probabilities. Standard errors are computed using 1000 draws from the estimated asymptotic variance of the ML estimates. Sh.Trans refer to the share of the transitory variance in the total variance of the unobservable. Rho UH refers to the correlation between time invariant unobserved heterogeneity in care and employment while rho eps refers to the same correlation in the transitory errors. Rho Init refers to the correlation between errors in the initial condition equation. Finally the four parameters L(x,k) refer to the initial parameters lambda in eq. 5.

Table 8 Results for Work and Caring – Males**Transition Equation Marginal Effects**

	Work Eq.		Caring Eq.	
	M.Est	S.Dev	M.Est	S.Dev
age 30-40	0.107	0.0102	0.0262	0.0101
age 40-50	0.115	0.0127	0.0698	0.0155
age >50	0.0467	0.0215	0.0746	0.0194
byear/10	0.339	0.0684	-0.0614	0.111
o level	0.0447	0.011	0.00335	0.00896
a level	0.0364	0.0119	0.00787	0.00995
degree	0.107	0.00969	-0.00841	0.00973
high d	0.0995	0.012	-0.0214	0.0157
kid 0-4	0.00639	0.0116	-0.0123	0.00664
kid 5-11	-0.0101	0.0106	0.00549	0.00644
rent	-0.113	0.011	0.00992	0.00703
hlimit w	-0.284	0.0174	0.0178	0.0083
nlinc/1000	-0.035	0.00613	-0.000283	0.0114
permanent	-0.0766	0.0125	0.0957	0.0226

Lagged Variable Marginal Effects

	Work Eq.		Caring Eq.	
	M.Est	S.Dev	M.Est	S.Dev
work(t-1)	0.425	0.0208	-0.0154	0.00786
care(t-1)	-0.0249	0.0148	0.156	0.0118

Covariance Parameters

	P.Est	S.Dev
sh.Trans(h)	0.851	0.0183
sh.Trans(c)	0.72	0.0166
rho UH	0.0162	0.0715
rho eps	-0.0869	0.0363
rho Init	-0.0613	0.0844
L(h,c)	1.47	0.17
L(h,h)	-0.0356	0.157
L(c,h)	0.0392	0.0888
L(c,c)	1.1	0.0997
LogLike/N	-3.95	

Notes: Maximum simulated likelihood estimates using the BFGS algorithm and the GHK simulator with 40 Halton draws along each dimensions. Intercepts not included. Marginal Effects are computed at mean characteristics. For discrete outcomes, they are computed using discrete Differences in the probabilities. Standard errors are computed using 1000 draws from the estimated asymptotic variance of the ML estimates.

Table 9 Results for Work and Co-Residential Carers - Females

Transition Equation Short-Run Marginal Effects

	Work Eq.		Caring Eq.	
	M.Est	M.Dev	M.Est	M.Dev
age 30-40	0.0309	0.012	0.0187	0.00566
age 40-50	0.0612	0.0172	0.0188	0.00843
age >50	-0.0136	0.0247	0.0336	0.0131
byear/10	0.148	0.0755	-0.014	0.166
o level	0.0996	0.0125	-0.00895	0.00387
a level	0.0925	0.0125	-0.00967	0.004
degree	0.156	0.0109	-0.0159	0.00349
high d	0.158	0.0165	-0.0195	0.0046
kid 0-4	-0.22	0.012	0.00329	0.0035
kid 5-11	-0.0492	0.00944	0.013	0.00356
rent	-0.0963	0.0105	0.0132	0.00386
hlimit w	-0.201	0.0147	0.00199	0.00334
nlinc/1000	-0.0264	0.00617	0.00947	0.0101
permanent	-0.101	0.0158	0.0461	0.0297

Lagged Variable Short-Run Marginal Effects

	Work Eq.		Caring Eq.	
	M.Est	S.Dev	M.Est	S.Dev
work(t-1)	0.421	0.0156	-0.0078	0.00369
care(t-1)	-0.0589	0.0242	0.146	0.0193

Covariance Parameters (Work,Care)

	P.Est	S.Dev
sh.Trans(h)	0.788	0.0147
sh.Trans(c)	0.638	0.0304
rho UH	-0.125	0.069
rho eps	-0.0988	0.0462
rho Init	-0.2	0.119
L(h,c)	1.38	0.104
L(h,h)	0.0987	0.151
L(c,h)	-0.208	0.101
L(c,c)	1.71	0.197
LogLike/N	-3.42	

Notes: Maximum simulated likelihood estimates using the BFGS algorithm and the GHK simulator with 40 Halton draws along each dimensions. Intercepts not included. Marginal Effects are computed at mean characteristics. For discrete outcomes, they are computed using discrete Differences in the probabilities. Standard errors are computed using 1000 draws from the estimated asymptotic variance of the ML estimates.

Table 10 Results for Work and Co-residential Carers – Males

Transition Equation Short-Run Marginal Effects

	Work Eq.		Caring Eq.	
	M.Est	S.Dev	M.Est	S.Dev
age 30-40	0.108	0.011	0.0148	0.00655
age 40-50	0.116	0.0132	0.0351	0.0114
age >50	0.0483	0.0223	0.0535	0.0167
byear/10	0.348	0.0718	0.38	0.161
o level	0.0438	0.0115	-0.00879	0.00431
a level	0.0349	0.012	-0.00642	0.00443
degree	0.108	0.00983	-0.0157	0.00402
high d	0.101	0.0135	-0.00864	0.0078
kid 0-4	0.00636	0.0115	-0.00143	0.00404
kid 5-11	-0.0102	0.0105	0.00903	0.0039
rent	-0.113	0.0117	0.0105	0.00441
hlimit w	-0.285	0.0179	0.012	0.00511
nline/1000	-0.0348	0.00647	0.0266	0.0133
permanent	-0.0793	0.0129	0.04	0.0343

Lagged Variable Short-Run Marginal Effects

	Work Eq.		Caring Eq.	
	M.Est	S.Dev	M.Est	S.Dev
work(t-1)	0.431	0.0206	-0.00554	0.00443
care(t-1)	-0.0476	0.0272	0.0922	0.0136

Covariance Parameters

	M.Est	S.Dev
sh.Trans(h)	0.853	0.0182
sh.Trans(c)	0.626	0.0295
rho UH	-0.0495	0.0927
rho eps	-0.0243	0.0537
rho Init	-0.284	0.158
L(h,c)	1.53	0.183
L(h,h)	-0.217	0.242
L(c,h)	0.218	0.112
L(c,c)	1.46	0.168
LogLike/N	-2.72	

Notes: Maximum simulated likelihood estimates using the BFGS algorithm and the GHK simulator with 40 Halton draws along each dimensions. Intercepts not included. Marginal Effects are computed at mean characteristics. For discrete outcomes, they are computed using discrete Differences in the probabilities. Standard errors are computed using 1000 draws from the estimated asymptotic variance of the ML estimates.

Table 11 Results for Work and Extra-Residential Carers - Females

Transition Equation Short-Run Marginal Effects

	Work Eq.		Caring Eq.	
	M.Est	S.Dev	M.Est	S.Dev
age 30-40	0.0291	0.0118	0.0388	0.00898
age 40-50	0.0583	0.0171	0.0892	0.0148
age >50	-0.0185	0.024	0.0915	0.0191
byear/10	0.15	0.0713	-0.195	0.111
o level	0.0985	0.012	0.00175	0.0069
a level	0.0909	0.0122	0.00565	0.00901
degree	0.154	0.0105	-0.0177	0.00793
high d	0.157	0.0158	-0.0415	0.0129
kid 0-4	-0.219	0.0117	-0.0127	0.006
kid 5-11	-0.0496	0.00955	-0.000818	0.00538
rent	-0.0981	0.0106	-0.00022	0.00597
hlimit w	-0.203	0.0149	0.00564	0.00672
nlinc/1000	-0.0264	0.00614	-0.014	0.0116
permanent	-0.0995	0.0147	0.0797	0.0223

Lagged Variable Short-Run Marginal Effects

	Work Eq.		Caring Eq.	
	M.Est	S.Dev	M.Est	S.Dev
work(t-1)	0.42	0.016	-0.0105	0.00694
care(t-1)	0.00949	0.0123	0.213	0.0122

Covariance Parameters

	P.Est	S.Dev
sh.Trans(h)	0.788	0.0148
sh.Trans(c)	0.783	0.014
rho UH	-0.0207	0.0562
rho eps	-0.00141	0.0291
rho Init	-0.00536	0.0702
L(h,c)	1.37	0.1
L(h,h)	0.118	0.0913
L(c,h)	-0.0152	0.0954
L(c,c)	1.29	0.113
LogLike/N	-4.857	

Notes: Maximum simulated likelihood estimates using the BFGS algorithm and the GHK simulator with 40 Halton draws along each dimensions. Intercepts not shown. Marginal Effects are computed at mean characteristics. For discrete outcomes, they are computed using discrete differences in the probabilities. Standard errors are computed using 1000 draws from the estimated asymptotic variance of the ML estimates.

Table 12 Results for Work and Extra-Residential Carers – Males**Transition Equation Short-Run Marginal Effects**

	Work Eq.		Caring Eq.	
	M.Est	S.Dev	M.Est	S.Dev
age 30-40	0.107	0.0107	0.0182	0.00838
age 40-50	0.113	0.0132	0.0464	0.013
age >50	0.0455	0.0223	0.0419	0.0156
byear/10	0.337	0.0703	-0.232	0.124
o level	0.0441	0.0114	0.0178	0.00803
a level	0.0354	0.0118	0.0227	0.00929
degree	0.108	0.0096	0.0131	0.00908
high d	0.0998	0.0136	-0.00711	0.0148
kid 0-4	0.00641	0.0118	-0.0122	0.00559
kid 5-11	-0.0102	0.0106	-0.00606	0.00498
rent	-0.114	0.0116	0.00106	0.00583
hlimit w	-0.285	0.0176	0.0122	0.00633
nlinc/1000	-0.0346	0.00601	-0.00918	0.0122
permanent	-0.0787	0.0122	0.109	0.0239

Lagged Variable Short-Run Marginal Effects

	Work Eq.		Caring Eq.	
	M.Est	S.Dev	M.Est	S.Dev
work(t-1)	0.424	0.0215	-0.00894	0.00664
care(t-1)	-0.0102	0.0158	0.119	0.0113

Covariance Parameters

	P.Est	S.Dev
sh.Trans(h)	0.849	0.0182
sh.Trans(c)	0.732	0.0178
rho UH	0.0615	0.0764
rho eps	-0.0789	0.04
rho Init	-0.0302	0.09
L(h,c)	1.44	0.166
L(h,h)	0.15	0.163
L(c,h)	-0.0465	0.102
L(c,c)	1.1	0.116
LogLike/N	-3.599	

Notes: Maximum simulated likelihood estimates using the BFGS algorithm and the GHK simulator with 40 Halton draws along each dimensions. Intercepts not included. Marginal Effects are computed at mean characteristics. For discrete outcomes, they are computed using discrete differences in the probabilities. Standard errors are computed using 1000 draws from the estimated asymptotic variance of the ML estimates.

Table 13 Fixed Effect Dynamic Model --- Women

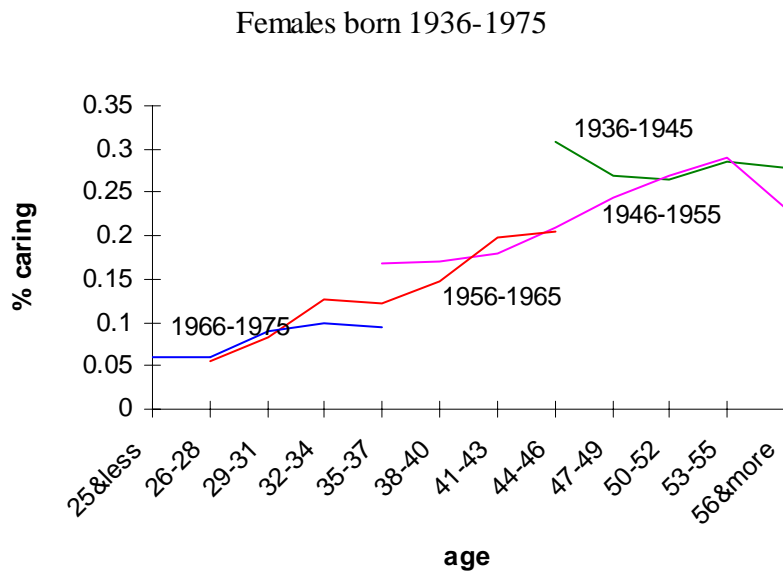
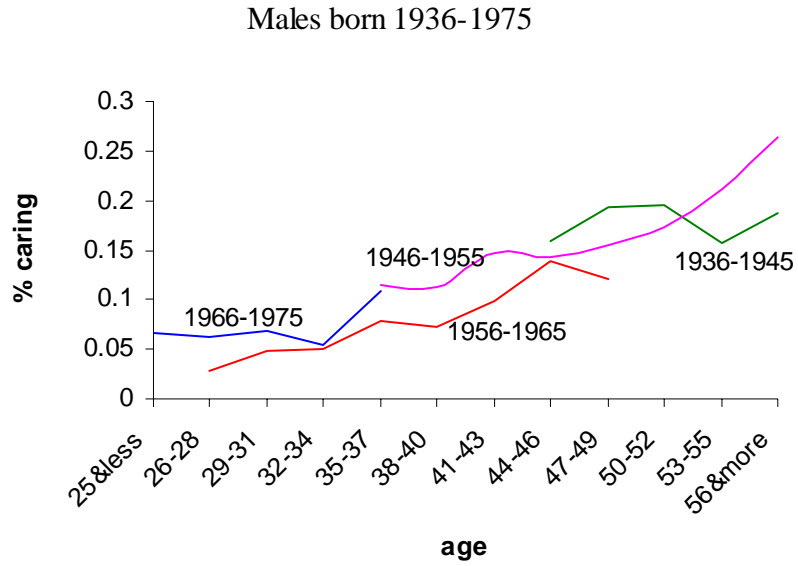
Employment Equation			
Variable	caring	co-res	extra-res
work t-1	.46218831***	.47268112***	.45692154***
caring t-1	-.00338926		
co-res t-1		-.05433085***	
ext-res t-1			.0065071
sargan	232.68927	241.15537	239.01948
N	3080	3080	3080
Caring Equation			
Variable	caring	co-res	extra-res
work t-1	-.02830001*	-.00426316	-.01423669
caring t-1	.34521094***		
co-res t-1		.50386318***	
ext-res t-1			.31062564***
sargan	272.46848	176.52574	282.82922
N	3080	3080	3080

legend: * p<0.05; ** p<0.01; *** p<0.001

Table 14 Fixed Effect Dynamic Model --- Men

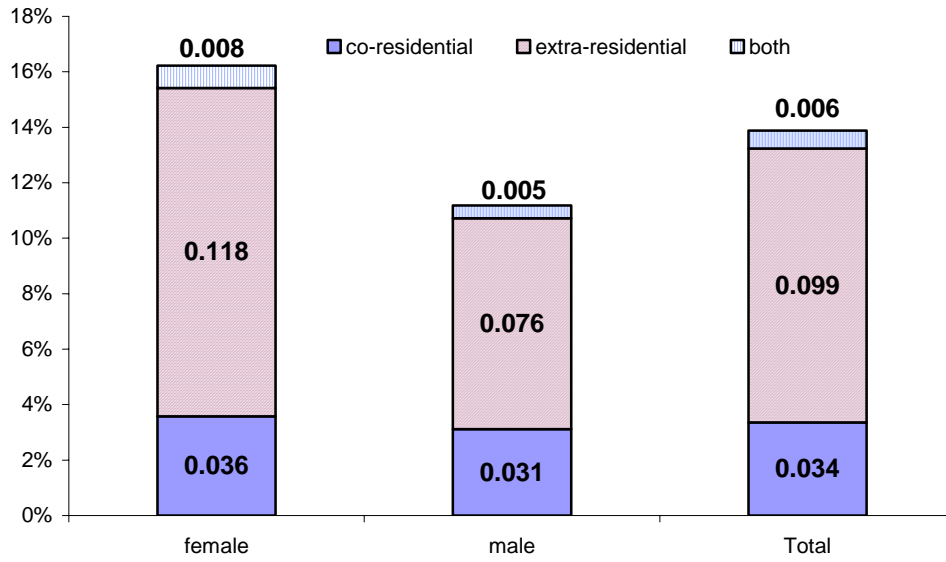
Employment Equation			
Variable	caring	co-res	extra-res
Work t-1	.32501188***	.34615819***	.32448966***
caring t-1	-.00084176		
co-res t-1		.00417977	
ext-res t-1			-.0065412
sargan	157.13417	167.88563	163.20097
N	2692	2692	2692
Caring Equation			
Variable	caring	co-res	extra-res
Work t-1	.04114374**	.00825243*	.01894592
Caring t-1	.25942603***		
Co-res t-1		.39238985***	
Ext-res t-1			.23667054***
sargan	213.3446	157.29204	191.32712
N	2692	2692	2692

Figure 1: Age-Cohort Profiles of Informal Care



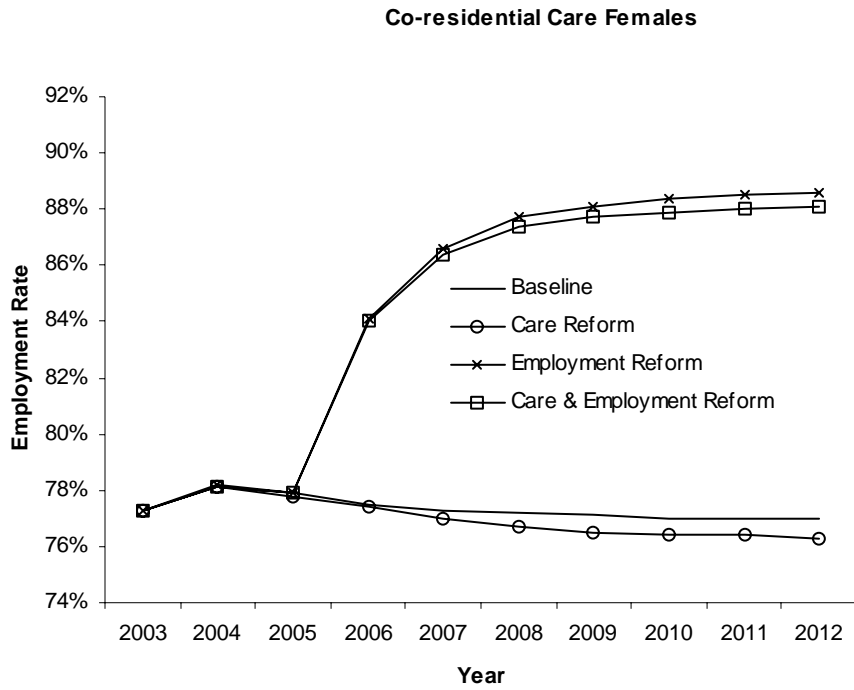
Notes: Authors calculation from the BHPS. Cohort profiles report fraction caring (weighted) within 3 year age group of a particular cohort.

Figure 2: Type of Carers by Gender



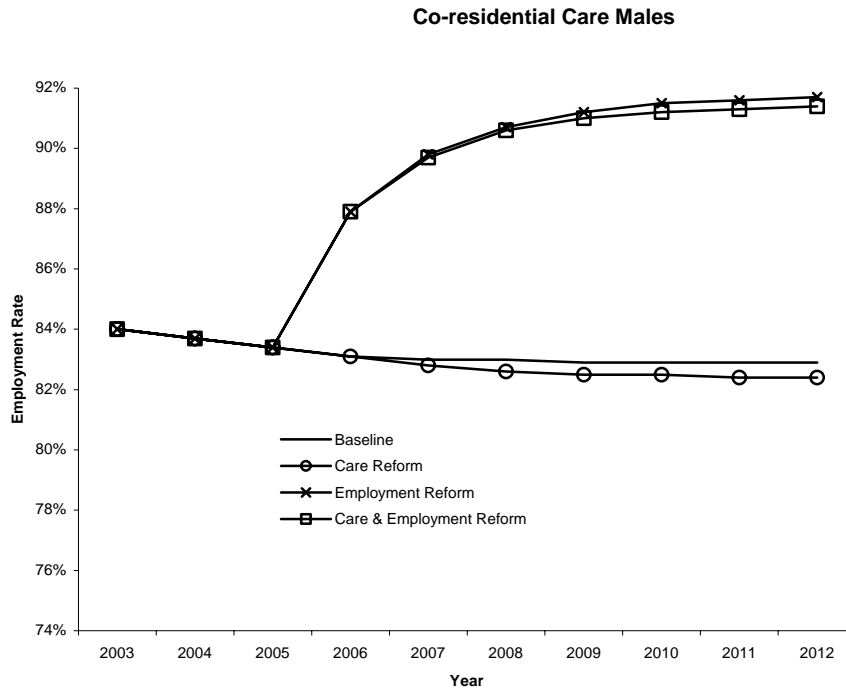
Notes: Authors' calculations from the BHPS. Weighted fractions for the pooled sample.

Figure 3: Pseudo Care and Employment Reforms – Employment Effect (Females)



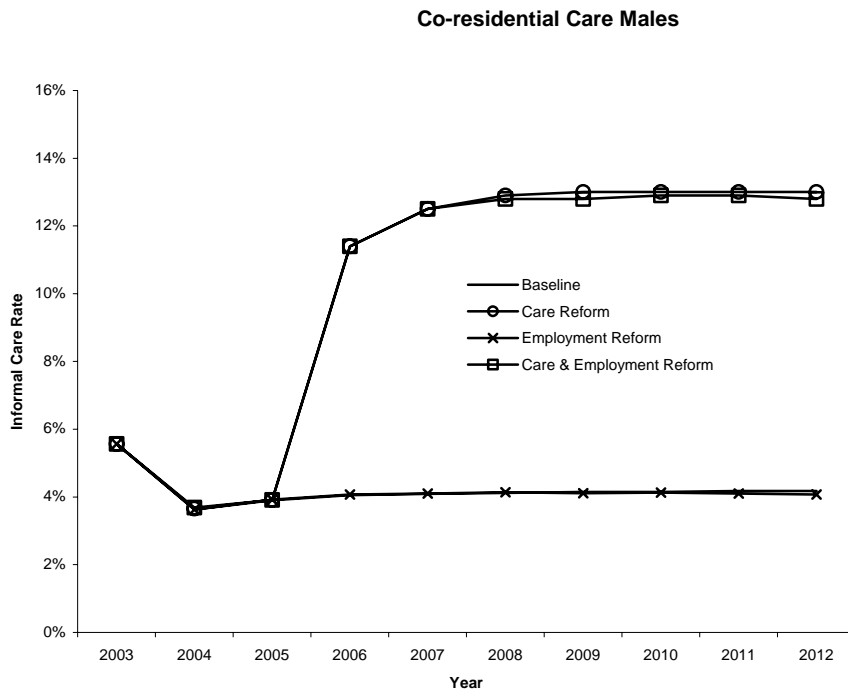
Note: Care and Employment reform refer to a 10 percent increase in respective rates based on individual level predictions from regression model.

Figure 4: Pseudo Care and Employment Reforms – Employment Effect (Males)



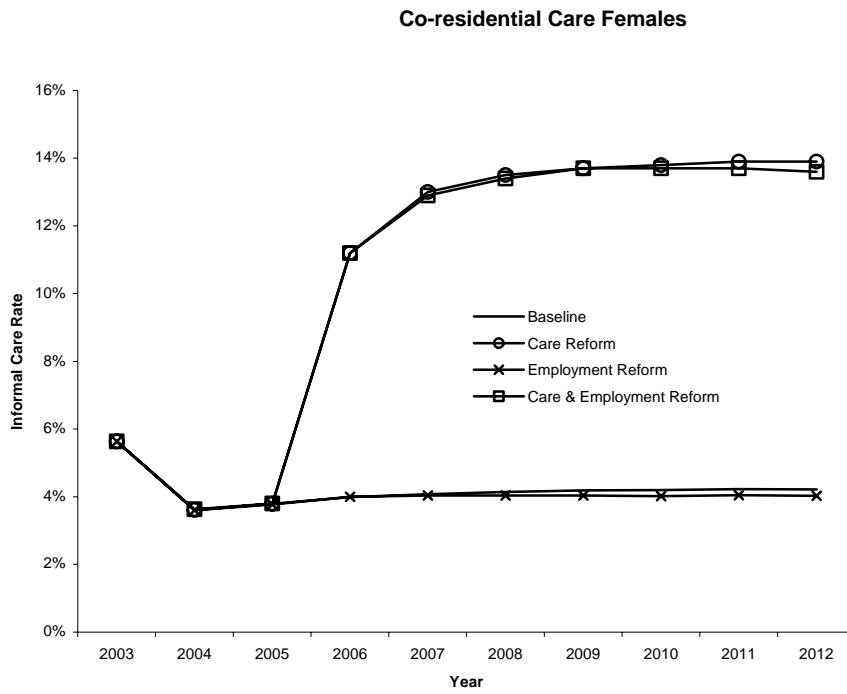
Note: Care and Employment reform refer to a 10 percent increase in respective rates based on individual level predictions from regression model.

Figure 5: Pseudo Care and Employment Reforms – Care Effect (Males)



Note: Care and Employment reform refer to a 10 percent increase in respective rates based on individual level predictions from regression model.

Figure 6: Pseudo Care and Employment Reforms – Care Effect (Males)



Note: Care and Employment reform refer to a 10 percent increase in respective rates based on individual level predictions from regression model.