Network for Studies on Pensions, Aging and Retirement

NETSPAR ACADEMIC SERIES

# The Dynamic Effects of Informal Caregiving on Caregivers' Health

Jannis Stöckel and Judith Bom

DP 02/2020-011

# The Dynamic Effects of Informal Caregiving on Caregivers' Health

Jannis Stöckel[a*] and Judith Bom[a]

[a] Erasmus School of Health Policy & Management, Erasmus University Rotterdam, the Netherlands

[*] Corresponding author Email address: stockel@eshpm.eur.nl (Jannis Stöckel)

**Abstract**

We estimate the long-run and dynamic health effects of providing informal care on caregivers' health in the United Kingdom. Using propensity score matching to address the endogeneity of informal care provision we estimate static and sequential matching models exploring health effects of multiple years of informal caregiving and the persistence of initial caregiving effects for up to five years after care provision started. Further, we purify the caregiving-effect from the family-effect, whether individuals suffer because they care *for* or care *about* someone, for a subsample of within-household caregivers. Our results suggest substantial negative health effects in the mental domain and asymmetrically experienced by caregivers of both genders providing more than 20 hours of weekly care. These effects are independent from the family effect. Lastly, our sequential matching results indicate that the mental health effects of care provision seem to persist for caregivers providing multiple years of care.

*Keywords:* Informal care, mental health, physical health, propensity score matching, family effect

*JEL Classifications:* I10, I18, C21, J14

**Acknowledgements**

We gratefully acknowledge comments provided by Cameron Birchall, Joaquim Vidiella-Martin, Sara Rellstab, Pieter Bakx, Eddy van Doorslaer, Erik Schut, Job van Exel, Werner Brouwer, participants at the LSE Care Policy Evaluation Centre seminar, the KU Leuven Public Economics seminar, the 2019 lowlands Health Economic Study Group, the 2019 UK Health Economists' Study Group summer meeting, the 2019 International Health Economics Congress and the ESE/ESHPM PhD Research Symposium. This study received financial support from the Network for Studies on Pensions, Ageing and Retirement Grant *optimal saving and insurance for old age: The role of public-long term care insurance* (Bom) and the *Smarter Choices for Better Health Initiative* financed by the Erasmus Trustfonds (Stöckel). All remaining errors are our own.

# Introduction

Ageing populations pose a serious challenge to national health care systems in developed economies. Across OECD member countries, the share of individuals aged 65+ is expected to rise from 17.4% in 2017 to 27.1% in 2050. The United Kingdom (UK) is an exemplary case, by 2050 more than a quarter of the population is expected to be 65+ and over 10% of the population is predicted to be 80+ (OECD, 2019). These demographic changes are one of the drivers for the increasing demand for long-term care (LTC) (de la Maisonneuve & Martins, 2014).

One solution to meet the growing LTC demand is to increasingly rely on informal care, care provided by friends or relatives. Informal care is often preferred by the care recipient and is from a governmental perspective a low-cost alternative to formal care. In addition, there is evidence that (partially) substituting formal by (unskilled) informal care does not jeopardize care recipients' health outcomes. Receiving informal care can lower overall medical expenditures (van Houtven & Norton, 2004, 2008), decrease the likelihood of infections and bedsores (Coe et al., 2019) and improve care recipients mental health (Barnay & Juin, 2016). In the UK, informal care already plays a crucial role in meeting current care demand, in 2017 more than 18% of the 50+ population provided informal care in contrast to the OECD average of 13.5% (OECD, 2019).

Despite its benefits, there are serious concerns regarding the impact of informal care on labor market and health outcomes of the caregiver. To make informed decisions on adapting current policies to future demands a thorough understanding of such spillover effects is crucial due to their implications e.g. in the appraisal of medical interventions (Basu & Meltzer, 2005). Previous studies either found no or negative effects of informal care provision on labor market outcomes (see Lilly et al. (2007) and Bauer and Sousa-Poza (2015) for reviews of the literature) and considerable health effects for the caregiver due to the often considerable mental and sometimes physical strain (see Bom et al., 2019a for a review).

Up to now most literature has however focused on the immediate impact of care provision, whereas it is important to understand how these effects develop over time. This is especially the case as many individuals provide several years of care. According to the UK Census of 2011 men and women at age 50 are likely to spend respectively 5.9 and 4.9 years of their remaining life as informal caregivers (ONS, 2017). It is furthermore important to focus on health outcomes as conflicting hypotheses regarding the impact of duration of caregiving on health exist.[1] The first theory, as presented by Townsend et al. (1989), Haley & Pardo (1989) and Pinquart & Sorensen (2003a) in their overviews, is called the *wear-and-tear theory* implying that coping resources might decline over time thereby worsening the impact of care provision. The *trait hypothesis* on the other hand suggests that the caregiving burden might be constant as caregivers maintain a constant level of adaptation which is determined by their own pre-existing resources such as coping skills. Lastly, the *adaptation theory* argues that individuals might learn to adapt to the situation and hence experience less burden over time.

Lacey at al. (2018) studied the relation between care patterns and depression in the UK and found a correlation between long-term or repeated care and depression among women. To study the causal impact of care provision on health one however must account for endogeneity concerns resulting from the selection of individuals into providing informal care. To our knowledge thus far only Schmitz and Westphal (2015) and De Zwart et al. (2017) have studied longer-term effects of informal care provision in a causal framework. Using German panel data and focusing on female caregivers, Schmitz and Westphal (2015) find negative mental health effects that persist for up to three years after care provision and vanish thereafter. De Zwart et al. (2017) used panel data from multiple continental European countries to explore the effect of spousal caretaking among the elderly population. They report negative effects on mental health and increased medical consumption, however only in the first year after care provision. The disappearance of health effects over time could mean that caregiving effects do not last or that individuals find ways to cope with them. However, these results might also be driven by selective attrition, as the authors note that individuals with demanding caregiving tasks are more likely to drop out of the panel.

To better understand the longer-term health effects of care provision this paper explores the health effects of providing informal care in the UK context using data from the Understanding Society (USoc) longitudinal survey. We estimate both (i) the immediate and longer-term health effects of providing informal care for up to 5 years after the initial caregiving decision and (ii) the effect of

[1] A related stream of literature focuses on the longer-term impact of care provision on labor market outcomes. Schmitz & Westphal (2017) studied the German context and found informal care provision to have a longer-term impact on labor market outcomes, this effect did not differ dependent on the duration of care provision (e.g. individuals that provided 1 year of care compared to multiple years of care provision). Rellstab et al. (2020) studied the Dutch context and did not find any impact of care provision on labor market outcomes, which as they argue might be explained by the generous social support system in the Netherlands.

providing additional years of informal care. These effects, and their relation to care intensity and caregiver characteristics, might help policymakers to gauge the potential consequences of informal care provision and to identify the subgroups of caregivers that are in largest need of support.

We utilize the first eight waves of USoc, a representative panel-survey of the adult UK population, to explore caregiving effects in the UK. Our study extends the literature on the longer-term health effects of providing informal care for the caregiver in several ways. First, our large sample of detailed individual-level information on caregivers and care recipients allows us to explore the heterogeneity of caregiving effects along different groups of caregivers (e.g. compare the impact of different intensities of care provision). In addition, we check whether these caregiving effects might be partly explained by the so-called family effect. The family effect refers to the mental strain associated with caring about a close relative in need of care, which is distinct from the caregiving effect that results from providing care to someone in need (Bobinac et al., 2010). Second, we estimate the health effects of multiple years of care provision using a dynamic matching approach (Lechner, 2009). In doing so, we follow the work of Schmitz and Westphal (2017) who used dynamic sequential matching techniques to investigate labor market effects of providing informal care in Germany. The added benefit of using this technique is that we can investigate the impact of an additional year of care provision to determine whether caring tasks become more or less straining over time. Lastly, to our current knowledge we provide the first causal estimates for caregiving effects for the UK context. We thereby provide new evidence on the existence of caregiving effects and their magnitude from an institutional context in which informal care plays a central role in the delivery of social care services (Comas-Herrera et al., 2010).

In line with previous studies we find strong negative health effects in the mental health domain. We find that these effects are concentrated among high-intensity caregivers and remain persistent for multiple years. Additionally, the analyses suggest that these effects are not driven by the deteriorating health of a relative or the relationship between caregiver and recipient, and hence are independent from the family effect. Lastly, the estimates from our dynamic matching procedure indicate that the mental health effect of care provision seems to remain persistent over longer care trajectories but provide no clear-cut evidence for a marginal health effect of providing an additional year of care. Using alternative outcome measures, a broad subjective well-being measure and a mental health screening questionnaire, we find our results to be consistent across measures and indicative for their economic relevance.

# Institutional Background

Formal LTC in the UK is organized in a mixed-system, referring to a combination of universal and means-tested benefits and programs. Health services provided by the National Health Service (NHS) are free at the point of delivery and predominantly financed from general taxation. The health-related components of LTC, which mostly entail home nursing, are funded via the NHS, when granted by the GP (Comas-Herrera et al., 2010).

Other types of LTC, such as residential care and help with personal and domestic care at home, are the responsibility of local authorities (Glendinning, 2013). Access to these services is dependent on needs-assessments that are conducted at the local level. This care is offered via a safety-net structure, meaning that individuals are required to deplete their wealth before becoming eligible for publicly funded care (Colombo et al., 2011).[2] This system ensures that publicly funded LTC services are only provided to those with severe needs and who are unable to pay for care themselves (Fernández et al., 2009).[3] In 2015 the UK spent about 1.5% of GDP on LTC-services, about 23% of these overall LTC expenditures related to social care (ONS, 2015).

As formal LTC services are solely targeted at a specific group, a large part of LTC is provided informally. This is in line with the main philosophy of the current UK system which places responsibility for non-health related LTC as much as possible with individuals and their family (Comas-Herrera et al., 2010). As mentioned, in 2017 more than 18% of the UK 50+ population provided informal care (OECD, 2019). Additionally, more than a third of all caregivers do so for more than 20 hours per week according to data from the 2011 UK Census (ONS, 2013).

In response to this large dependence on informal caregivers in the UK LTC system, various policies to support informal caregivers (e.g. by providing information or support groups) are in place. The Care Act, which was implemented in 2014, gave caregivers the right to receive a needs-

[2] The current regime considers both income and assets (including housing wealth for nursing home care under certain circumstances). Individuals with assets above GBP 23.250 are not eligible for support. Those with assets between GBP 14.250 - GBP 23.250 (approximately $18.448 - $30.100) are required to pay a certain part of the care themselves. Individuals with assets below GBP 14.250 will have their costs completely covered (NHS, 2018).

[3] In case of self-funding one should expect costs of about £15/hour (approximately $19) for home care (Age UK, 2019a) and £600 and £800/week (approximately $777 and $1036) for care homes and nursing homes (Age UK, 2019b).

assessment and corresponding support services (European Commission, 2018). Reaching caregivers with the designed support options is however difficult and specifically mentioned as a duty of local councils. Only 6% of caregivers receive any form of local authority support (Yeandle, 2016). Financial support is offered to informal caregivers via a “carer's allowance” (Carers UK, 2016). This allowance, amounting to £66.15 (approximately $86) a week in 2020 (UK Government, 2020) is only paid to caregivers who meet certain conditions.[4] We do not expect this allowance to act as a financial incentive to provide care as the amount is rather small and take up of this allowance is low (Colombo et al., 2011). As the changes in informal care support policies are relatively small and take-up of most services is low, we assume that throughout the studied years the incentive structure in which informal caregivers operate has stayed relatively stable.

[4] Individuals can receive the carer’s allowance when they (i) are aged 16 or over (ii) provide at least 35 hours of care a week; (iii) earn less than £123 per week (approximately $152); (iv) are not full-time students or studying for more than 21 hours a week; (v) normally live in the UK and have been in the UK for at least 2 of the last 3 years (UK Government, 2020)

# Methods

The decision to provide informal care is not necessarily random. A specific group of individuals 'selects into' informal caregiving, thereby creating endogeneity when studying the health impact of informal care provision. We aim to overcome this problem by matching individuals on observable characteristics that affect health outcomes and the decision to provide informal care. To do so, we follow the intuition regarding the caregiving decision as proposed by Schmitz and Westphal (2015). They define three stages that separately affect one's caregiving decision. The first stage is care obligations, as the most important determinant of informal care provision is the presence of a family member in need and the presence of alternative potential caregivers. The second category, willingness to provide care, refers to personality traits and socio-economic characteristics, as these affect someone's inclination towards providing care. Lastly, the ability to provide care refers to someone’s caregiving ability based on one’s own health status.

Our empirical strategy builds upon the potential outcomes framework by Rubin (1974) and addresses the endogeneity of providing informal care using regression adjusted propensity score matching (Rubin, 1979). The main assumption underlying propensity score matching is that the conditional independence assumption (CIA) holds. The CIA in our context states that after conditioning on a set of observable variables, the potential health outcomes for both caregivers and non-caregivers are the same in the absence of informal care provision at all considered time periods. This implies that differences in health outcomes between caregivers and non-caregivers can be attributed to the provision of informal care alone. Following Lechner (2009), we make this assumption more credible by exploiting the panel structure of our data to match individuals upon information from the period directly preceding informal care provision. The advantages of this strategy are that (i) providing care cannot affect the covariates and (ii) the previous caregiving status captures most of the unobserved heterogeneity.

*Static Matching*

Our first aim is to estimate the static long-term impact of becoming an informal caregiver. This static approach means that we match starting caregivers with non-caregivers in the first time point and follow these two groups over time. We identify individuals as treated when we observe their transition into caregiving, everyone who does not report any care-episode is included in the control group (untreated). From here onward we will refer to the period directly preceding informal care provision as $t_{-1}$ and the period of first reporting informal care as $t_0$.

Propensity scores of providing informal care are estimated using probit models. We estimate the propensity of providing care at $t_0$ conditional on the variables affecting the transition into care provision at $t_{-1}$. We use these propensity scores to match treated to untreated individuals. To make use of the large amount of information available in the dataset we use a kernel matching approach that uses weighted averaging on the untreated sample to form the counterfactual group.[5] In contrast to alternative matching algorithms, this approach includes nearly all untreated individuals leading to usage of more information and hence a lower variance. This approach is furthermore preferred over for example nearest neighbor matching due to the large number of controls we match upon and our treated to control ratio. We assess the common support, whether there is sufficient overlap in characteristics between the treated and untreated individuals, as the risk of kernel matching lies in the increased chance of including "bad matches", untreated individuals that are highly dissimilar to the treated group, in the estimation (Caliendo & Kopeining, 2008). Furthermore, as we do not match on actual covariates but on propensity scores, we assess whether balance of covariates is achieved after the matching procedure. We do so by using the standardized bias (Rosenbaum & Rubin, 1985).

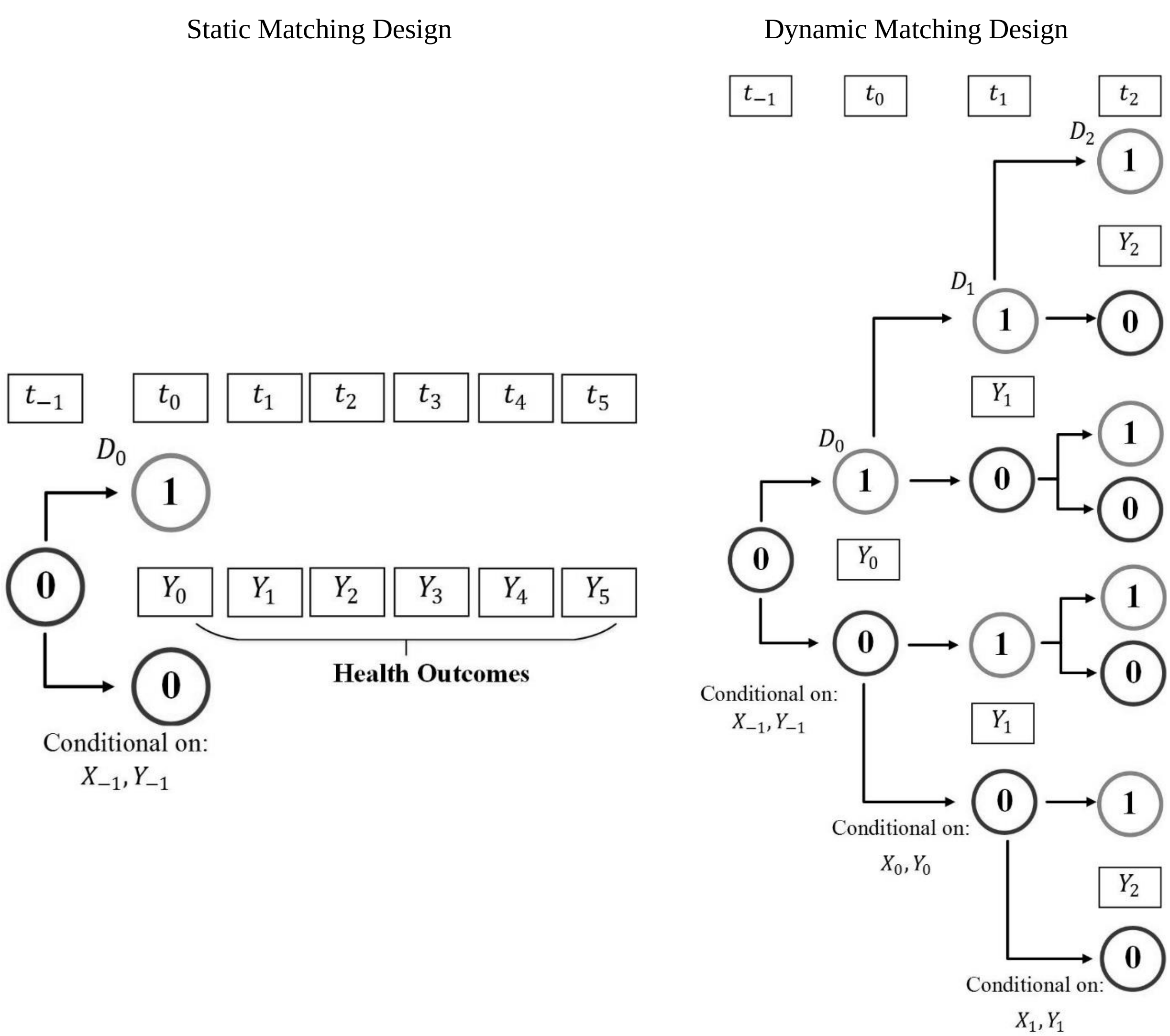
Finally, the average treatment effect on the treated (ATT) is estimated by regressing health outcomes on the treatment indicator (providing care) and all control variables used in the propensity score estimation with individuals in the control group weighted by their estimated kernel weights. By regressing on the control variables alongside the treatment indicator we aim to correct for remaining residual differences in the covariate distributions between the treatment and control group (Lechner, 2009; Rubin 1973). We do not use the covariates from later waves as these might be affected by treatment. The health impact of providing care is estimated for the immediate

[5] We make use of the Stata command psmatch2 (Leuven & Sianesi, 2003) using an Epanechnikov kernel with a bandwidth of 0.03. The bandwidth choice is a trade-off between a small variance and an unbiased estimate of the true density function (Caliendo & Kopeining, 2008). While not reported in detail we have tested our specification for changing bandwidths, e.g. higher and lower bandwidth values of 0.01 and 0.06, with negligible impact on our results.

time after care provision and up to five years afterwards. Figure 1 provides a graphical representation of the static and dynamic matching designs.

To study longer-term health effects, we follow the health of individuals starting care provision at $t_0$ over time. As mentioned, in this static matching approach, we only focus on initial care provision and disregard the different caregiving paths underlying these long-term results. Our treatment group hence contains both individuals that stopped providing care in $t_1$ and those who continued caregiving for various years. In our main analysis, our control group consists of individuals that never provide care.

Figure 1: Static and Dynamic matching design. (Own illustrations – based on Schmitz & Westphal 2017)



*Dynamic Sequential Matching*

Next to the static matching procedure, we estimate the impact of uninterrupted caregiving spells lasting several years using a dynamic matching procedure following the work of Lechner & Miquel (2010) and Schmitz & Westphal (2017). In contrast to the longer-term effects of providing care for at least one year, we now specifically follow individuals that provide multiple years of care to find out: is it much worse to provide care for 2 years than for 1 year? And much worse to provide care for 3 years than for 2 years? This helps to understand how the health effect of care provision is affected by duration of care. This is interesting as different theories regarding the longer-term health impact of caregiving exist and (mental) health outcomes might either decrease due to adaptation or increase as individuals run out of resources to cope with care provision.

As most individuals provide care for a limited amount of time, we solely follow individuals up to three years of care provision. To estimate the dynamic caregiving effects we have to match individuals on their propensity of providing care at all decision nodes, as caregiving throughout multiple years is a repeated decision. Based on various elements, for example one's own health, one may decide to continue or quit providing care. At every decision point (data wave) we therefore re-estimate the propensity of providing care to ensure comparability between the persisting caregivers and non-caregivers.

To explain the approach, we provide an example showing the steps undertaken to estimate the marginal effect of providing two years of care instead of one. The treatment group in this example comprises everyone that provided informal care in both waves ($t_0$ and $t_1$), whereas the control group consist of everyone that provided care in the first wave ($t_0$) but not in the second ($t_1$). In the dynamic matching design in Figure 1 this refers to comparing the group that followed the path *0-1-1* with the group that followed the route *0-1-0*.

Consider a binary indicator $D_t$ encoding whether care is provided in period $t$. Just like in the static estimations, we start our analysis by estimating the propensity of providing informal care at the first node ($D_0 = 1$) conditional upon not providing care in the period before, and pre-treatment health outcomes and other covariates using a probit model. The propensity of providing informal care at the first node is written as: $\Pr(D_0 = 1 | X_{-1,} Y_{-1})$.

Now, in extension to the static estimation which stops after matching individuals in the first node, we also estimate the decision taken at the second node ($D_1 = 1$). Here a caregiver decides whether to continue care provision or not. We estimate the propensity scores of both options conditional upon already being a caregiving and on health and the other observables both at the first as well as on the second node. The propensity of providing informal care at the second node after being a caregiver in the first period is defined as: $\Pr(D_1 = 1| D_0 = 1, X_{-1}, Y_{-1}, X_0, Y_0)$ The propensity of discontinuing care provision is defined as: $\Pr(D_1 = 0| D_0 = 1, X_{-1}, Y_{-1}, X_0, Y_0)$.

We use these scores to calculate inverse probability weights (IPW). IPW estimates might be sensitive to very high or low weights from individuals with very high or low propensity scores (Robins et al., 2000). As mentioned by Lechner (2009) the commonly used solution to this problem is to remove observations with extreme weights, we hence prevent the results from being sensitive to these propensity scores, by dropping all scores for the first decision that are smaller than 5% or larger than 95% of the estimated propensity score distribution.[6] Furthermore, for all scores we condition upon common support: in case no untreated counterparts with a similar propensity score for our treated respondents are present, these treated observations are excluded from the analysis.

Based on the estimated propensity scores we can calculate inverse probability weights for both the treatment and the control group. These are defined as follows:

$$\frac{1}{\left(\Pr\,(D_0 = 1|\, X_{-1}, Y_{-1})\right) * (\Pr(D_1 = 1 \mid D_0 = 1, X_{-1}, \mathrm{Y} - 1, X_0, Y_0))}\ \textit{for the treatment group}$$

$$\frac{1}{\left(\Pr\,(D_0 = 1|\, X_{-1}, Y_{-1})\right) * (\Pr(D_1 = 0 \mid D_0 = 1, X_{-1}, \mathrm{Y} - 1, X_0, Y_0))}\ \textit{for the control group}$$

We estimate the dynamic average treatment effect on the treated (those who provide two years of care) by regressing health on the treatment while controlling for any remaining differences by adding all covariates from the previous wave and weighting the data using the calculated inverse probability weights. We hence estimate, in this example, the health effects in $t_1$ of providing care in $t_0$ and $t_1$ compared to only providing care at $t_0$.

[6] To check the robustness of this approach we also estimate our results while (1) dropping all scores for the first decision that are smaller than 1% or larger than 99% of the estimated propensity score distribution and (2) dropping scores for all decision nodes that are smaller than 1% or larger than 99% of the estimated propensity score distributions. Qualitatively our results are robust to these different specifications (results available upon request).

This sequential matching strategy was proposed by Lechner (2009) to estimate treatment effects in settings with dynamic treatment durations. While it follows a similar intuition as the static matching procedure, identification is based on an augmented version of the CIA: the weak dynamic conditional independence assumption. Consider the case above comparing outcomes of two and one years of informal care. The weak conditional independence assumption combines two parts. Firstly, the initial conditional independence assumption stating that potential outcomes in $t_0$ and $t_1$ are independent of treatment status in $t_0$ once we match upon observables at $t_{-1}$. Secondly, that potential outcomes in $t_0$ and $t_1$ are independent of continued treatment in $t_1$ once we condition on control variables and outcomes at both $t_{-1}$ and $t_0$ and treatment status at the initial node $t_0$.

# Data

We use data from the Understanding Society (USoc) dataset, also known as the United Kingdom Household Longitudinal Study (UKHLS; University of Essex, 2019); an annually conducted representative panel survey of the adult UK population (aged 16+). The USoc started in 2009 with approximately 40,000 self-completing respondents across 30,000 households as the successor of the British Household Panel Survey (BHPS), which ended in 2008. In 2010, members of the last BHPS-wave were invited to join the USoc after which an additional 8,000 individuals joined. The USoc uses an overlapping panel design in which waves are collected over a 24-month period with individuals being interviewed every 12 months. This paper uses the complete dataset spanning all nine completed waves conducted between 2009 and 2019.

Informal caregivers are identified using their answer to the following question, *"Is there anyone living with you who is sick, disabled or elderly whom you look after or give special help to (for example a sick, disabled or elderly relative/husband/wife/friend etc.)?"*. Individuals providing care outside their own household are identified based on their answer to the question, *"Do you provide regular service or help for any sick, disabled or elderly person not living with you? [Exclude help provided in course of employment]"*. Apart from being able to identify individuals providing care inside and outside their own household the annual questionnaire also includes questions on care intensity (hours of care per week) and the relationship between the caretaker and care recipient within the household and for the first two care recipients outside the household.

We explore differences in the impact of caregiving dependent on the reported hours of care per week. Based on these reported hours of care provision we split our sample of caregivers in three, making a distinction between low intensity (<10 hours of care per week), medium intensity (between 10 – 20 hours of care per week) and high intensity caregivers (more than 20 hours of care per week). When evaluating our results, one however must be aware of a potential downward bias in our estimates due to an underrepresentation of caregivers in the upper end of the intensity distribution. The share of high intensity caregivers in our sample (12.8%) is namely much lower compared to the UK Census of 2011 which indicates that nation-wide about 37% of the caregivers provide care for more than 20 hours a week (ONS, 2013).

We construct two distinct datasets to implement the static and dynamic matching procedures. Individuals who identified as caregivers in their first observation period are excluded as for these we cannot observe the transition from non-caregiving into caregiving. For the static estimation we include all individuals that provide information for at least two time points on their health outcomes ($t_{-1}$ and $t_0$) and provide full information on all covariates used in the propensity score estimation at $t_{-1}$. Individuals remain in the sample during the subsequent time points $t_1$ to $t_5$ in case information on all outcome-variables is available. For the dynamic estimation procedure, data requirements are more restrictive as we re-estimate propensity scores at each decision node. For this analysis, only individuals with complete information on all control variables for three waves ($t_{-1}$ to $t_1$) and full information on outcome variables for four consecutive periods are included in the sample.

*Health outcomes*

Various studies report the impact of care provision on mental and physical health (e.g. Pinquart & Sörensen, 2003b). To identify potential changes in both health domains we make use of two measures from the SF-12 in which individuals self-report on 12 questions related to various aspects of their own health in the past four weeks. From the survey we derive two health scales namely the physical (PCS) and mental (MCS) component summary scales, both summary measures are constructed using different subscales related to physical and mental health.[7] The two health scales are validated for the UK context and range from 0 to 100, where a higher score represents a better health status. Both MCS and PCS scales are transformed to have a mean of 50 and standard deviation of 10 (Ware et al., 1995).

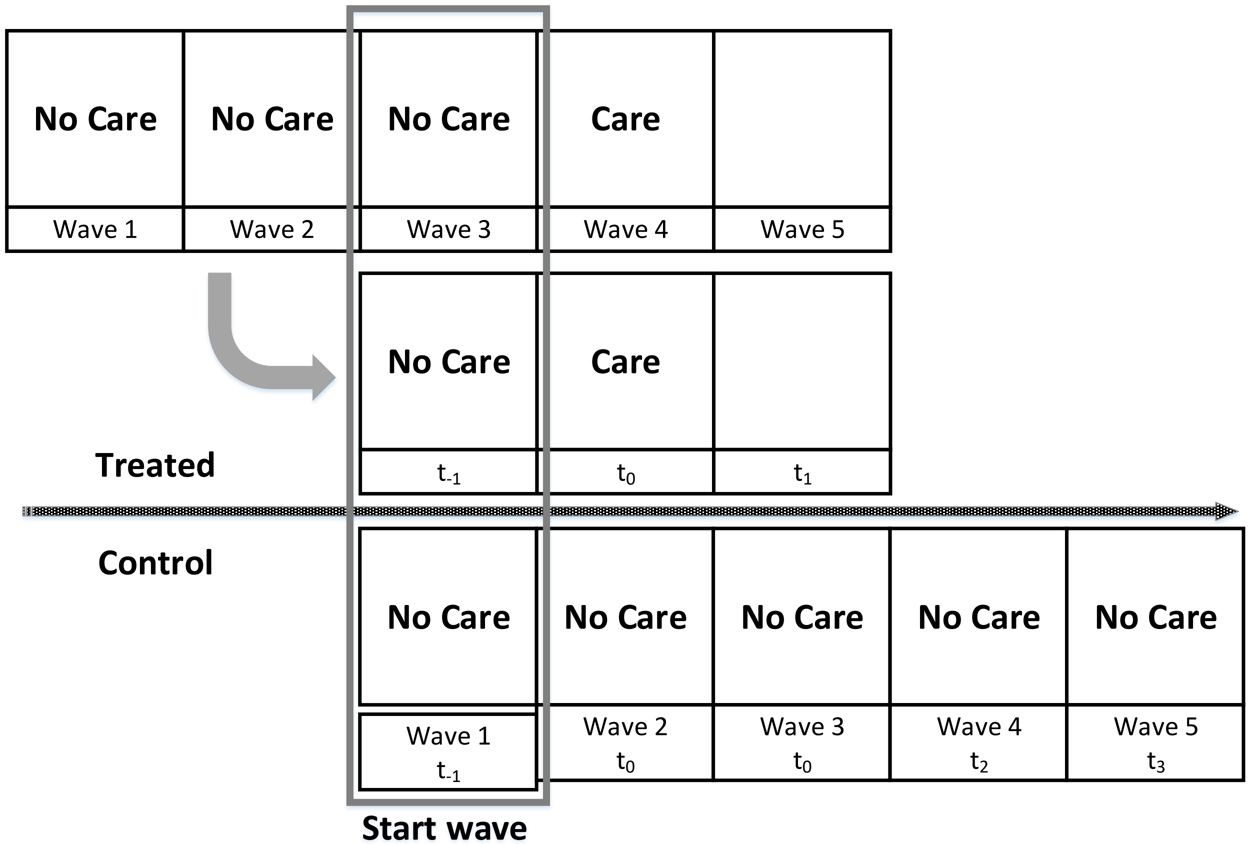
*Time Structure*

For the static matching procedure, we define a relative time variable whose value depends on an individual's first reported care-episode. Figure 2 provides a visualization of the created time structure. Among caregivers $t_{-1}$ is defined as the period before the first reported caregiving episode. For everyone in the control group, $t_{-1}$ is normalized to the individual's first appearance as a survey participant as they do not report any care episode during their participation. This time

[7] The PCS comprises the subscales: Physical functioning, Role-Physical, Bodily Pain and General Health. The MCS comprises the subscales: Vitality, Social Functioning, Role-Emotional and Mental Health.

structure is chosen to maximize the number of treated individuals we can observe, however, because of this approach we lose observations and less precisely identify the estimated treatment effects in the later time points.

Figure 2: Static Dataset – Time structure (Own visualization)




The analysis sample for the dynamic specification uses an augmented time-structure to allow for the modelling of all decision nodes between $t_{-1}$ and $t_2$ and the comparison of various care trajectories. The time variable is normalized to $t_{-1}$ being the entry wave of an individual into the panel. We cannot reorder the time structure as done in the static sample as estimating propensity scores in the later decision nodes requires observations of individuals following different care trajectories. For example, estimating the propensity of not providing care at $t_1$ conditional upon not providing care at $t_0$, requires observing individuals that start care provision at $t_1$.

# Results

*Matching Quality*

The descriptive statistics for the static matching sample are depicted in Table 1. Before the propensity score matching there is strong imbalance between the covariates of the control and treatment groups. The matching succeeds in correcting this imbalance.[8] For a detailed overview of the results of the propensity score estimation and the distribution of estimated propensity scores across both groups please see Appendix A3. Only four caregivers are identified as off-support and therefore dropped from the analysis.

Table 1: Descriptive Statistics – Treatment and Control Groups

| | **Unmatched** | | | | **Matched Controls** | | **Standardized Bias** | |
|---|---|---|---|---|---|---|---|---|
| | **Treated** | | **Control** | | | | | |
| | Mean | SD | Mean | SD | Mean | SD | Unmatched | Matched |
| **Care Obligations** | | | | | | | | |
| Mother alive | 0.587 | 0.492 | 0.697 | 0.460 | 0.586 | 0.492 | -23.0 | 0.1 |
| Age of mother | 66.717 | 9.911 | 61.140 | 9.717 | 66.585 | 9.627 | 38.2 | 0.7 |
| Father alive | 0.446 | 0.497 | 0.631 | 0.482 | 0.441 | 0.497 | -25.5 | 0.6 |
| Age of father | 66.500 | 8.208 | 62.529 | 8.846 | 66.248 | 7.569 | 30.8 | 1.9 |
| Both parents alive | 0.388 | 0.487 | 0.596 | 0.491 | 0.384 | 0.486 | -28.4 | 0.5 |
| Living siblings | 0.865 | 0.342 | 0.885 | 0.319 | 0.864 | 0.343 | -4.1 | 0.2 |
| Living partner | 0.703 | 0.457 | 0.635 | 0.481 | 0.700 | 0.458 | 9.6 | 0.3 |
| Age of partner | 51.662 | 12.387 | 45.609 | 11.552 | 51.517 | 11.947 | 34.2 | 0.6 |
| **Willingness to Care** | | | | | | | | |
| Age | 50.666 | 16.157 | 41.288 | 17.146 | 50.589 | 16.296 | 37.4 | 0.2 |
| Female | 0.592 | 0.491 | 0.527 | 0.499 | 0.591 | 0.492 | 8.8 | 0.1 |
| Tertiary Education | 0.354 | 0.478 | 0.386 | 0.487 | 0.346 | 0.476 | -4.4 | 1.0 |
| Secondary Education | 0.431 | 0.495 | 0.439 | 0.496 | 0.433 | 0.496 | -1.0 | -0.2 |
| Primary Education | 0.215 | 0.411 | 0.175 | 0.380 | 0.221 | 0.415 | 6.7 | -0.9 |
| Employed | 0.485 | 0.500 | 0.567 | 0.496 | 0.471 | 0.499 | -11.0 | 1.7 |
| Self-Employed | 0.078 | 0.268 | 0.067 | 0.250 | 0.077 | 0.266 | 2.8 | 0.2 |

[8] We calculate the standardized bias for each covariate in the model by taking the difference in means between the treatment and control group and dividing it by the standard deviation of the control group (Rosenbaum & Rubin, 1985). We follow the rule of thumb suggested by Caliendo & Kopeinig (2008) which states that there is sufficient balance when the bias is below 3-5%.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Working Full-Time | 0.397 | 0.489 | 0.502 | 0.500 | 0.382 | 0.486 | -14.1 | 1.8 |
| Unemployed | 0.045 | 0.208 | 0.051 | 0.219 | 0.048 | 0.215 | -1.6 | -0.9 |
| Retired | 0.261 | 0.439 | 0.141 | 0.348 | 0.264 | 0.441 | 21.0 | -0.4 |
| Student | 0.034 | 0.181 | 0.102 | 0.303 | 0.034 | 0.182 | -17.2 | -0.2 |
| Homecarer | 0.060 | 0.238 | 0.049 | 0.216 | 0.065 | 0.247 | 3.3 | -1.2 |
| Disabled | 0.037 | 0.189 | 0.023 | 0.150 | 0.040 | 0.197 | 5.7 | -1.0 |
| Income (logarithmic) | 7.258 | 0.610 | 7.231 | 0.624 | 7.228 | 0.634 | 2.9 | 2.8 |
| HH Income Fraction | 0.549 | 0.313 | 0.518 | 0.327 | 0.550 | 0.321 | 6.3 | -0.3 |
| Single | 0.154 | 0.361 | 0.249 | 0.433 | 0.153 | 0.360 | -15.5 | 0.2 |
| Partnership | 0.108 | 0.311 | 0.160 | 0.367 | 0.110 | 0.313 | -10.0 | -0.4 |
| Separated/Divorced | 0.099 | 0.298 | 0.073 | 0.260 | 0.101 | 0.302 | 6.3 | -0.5 |
| Widowed | 0.044 | 0.205 | 0.043 | 0.202 | 0.045 | 0.208 | 0.4 | -0.4 |
| Married | 0.595 | 0.491 | 0.475 | 0.499 | 0.590 | 0.492 | 16.1 | 0.5 |
| Children in Household | 0.310 | 0.463 | 0.395 | 0.489 | 0.315 | 0.465 | -11.8 | -0.6 |
| Children < 14 in Household | 0.275 | 0.447 | 0.356 | 0.479 | 0.278 | 0.448 | -11.6 | -0.4 |
| Region: North-East | 0.043 | 0.204 | 0.037 | 0.189 | 0.044 | 0.206 | 2,21 | -0.2 |
| Region: North-West | 0.103 | 0.304 | 0.114 | 0.318 | 0.102 | 0.303 | -2.4 | 0.1 |
| Region: Yorkshire | 0.069 | 0.253 | 0.077 | 0.267 | 0.067 | 0.249 | -2.1 | 0.5 |
| Region: East-Midlands | 0.082 | 0.274 | 0.077 | 0.266 | 0.082 | 0.274 | 1.3 | 0.0 |
| Region: West-Midlands | 0.085 | 0.279 | 0.074 | 0.262 | 0.088 | 0.283 | 2.8 | -0.6 |
| Region: East England | 0.095 | 0.294 | 0.092 | 0.289 | 0.095 | 0.293 | 0.9 | 0.1 |
| Region: South-East | 0.124 | 0.330 | 0.136 | 0.342 | 0.126 | 0.332 | -2.3 | -0.3 |
| Region: South-West | 0.100 | 0.299 | 0.089 | 0.284 | 0.103 | 0.304 | 2.5 | -0.6 |
| Region: Wales | 0.074 | 0.262 | 0.067 | 0.250 | 0.073 | 0.260 | 1.9 | 0.2 |
| Region: Scotland | 0.084 | 0.278 | 0.090 | 0.287 | 0.082 | 0.275 | -1.5 | 0.4 |
| Region: Northern Ireland | 0.054 | 0.226 | 0.046 | 0.211 | 0.052 | 0.222 | 2.3 | 0.6 |
| Region: London | 0.087 | 0.281 | 0.101 | 0.301 | 0.087 | 0.281 | -3.3 | 0.0 |
| Living in Urban Area | 0.726 | 0.446 | 0.758 | 0.428 | 0.728 | 0.445 | -4.9 | -0.2 |
| Big-5: Openness | 4.577 | 1.325 | 4.592 | 1.272 | 4.566 | 1.314 | -0.8 | 0.5 |
| Big 5: Conscientiousness | 5.570 | 1.097 | 5.444 | 1.086 | 5.543 | 1.110 | 7.7 | 1.4 |
| Big 5: Extroversion | 4.618 | 1.289 | 4.588 | 1.299 | 4.630 | 1.345 | 1.6 | -0.6 |
| Big 5: Agreeableness | 5.691 | 1.018 | 5.573 | 1.027 | 5.680 | 1.003 | 7.7 | 0.6 |
| Big 5: Neuroticism | 3.532 | 1.447 | 3.556 | 1.432 | 3.526 | 1.473 | -1.2 | 0.2 |
| **Ability to Care** | | | | | | | | |
| Self-Assessed Health | 2.578 | 1.054 | 2.423 | 1.054 | 2.607 | 1.131 | 9.9 | -1.6 |
| SF12 - Mental Score | 50.006 | 9.860 | 51.109 | 9.080 | 49.832 | 10.436 | -7.9 | 1.0 |
| SF12 - Physical Score | 49.435 | 10.854 | 51.568 | 9.970 | 49.140 | 11.751 | -13.9 | 1.6 |
| Chronic-Illness/Disability | 0.374 | 0.484 | 0.287 | 0.453 | 0.384 | 0.486 | 12.5 | -1.2 |
| Functional Limitations | 0.552 | 1.296 | 0.441 | 1.221 | 0.589 | 1.345 | 5.9 | -1.6 |
| Satisfaction with Health | 4.729 | 1.722 | 5.091 | 1.599 | 4.691 | 1.764 | -14.7 | 1.3 |
| Satisfaction with Income | 4.505 | 1.689 | 4.639 | 1.637 | 4.455 | 1.741 | -5.5 | 1.7 |

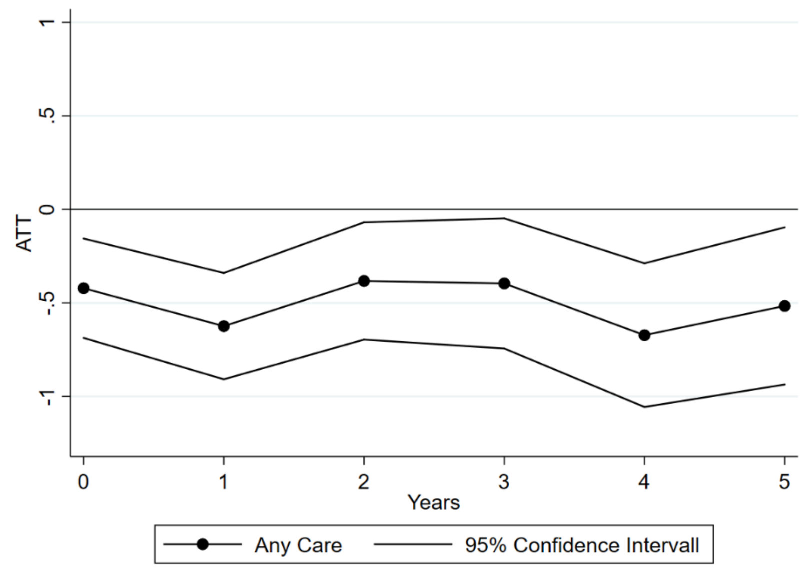
| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Satisfaction with Leisure | 4.788 | 1.676 | 4.773 | 1.622 | 4.779 | 1.728 | 0.6 | 0.3 |
| Satisfaction with Life | 5.198 | 1.472 | 5.369 | 1.370 | 5.165 | 1.547 | -8.1 | 1.3 |
| GHQ Score | 10.250 | 3.019 | 10.412 | 2.729 | 10.226 | 2.990 | -3.8 | 0.5 |
| **Number of Individuals** | | **6,852** | | **12,970** | | **12,970** | | |

*Static results – Treatment Effects by Caregiving Intensity*

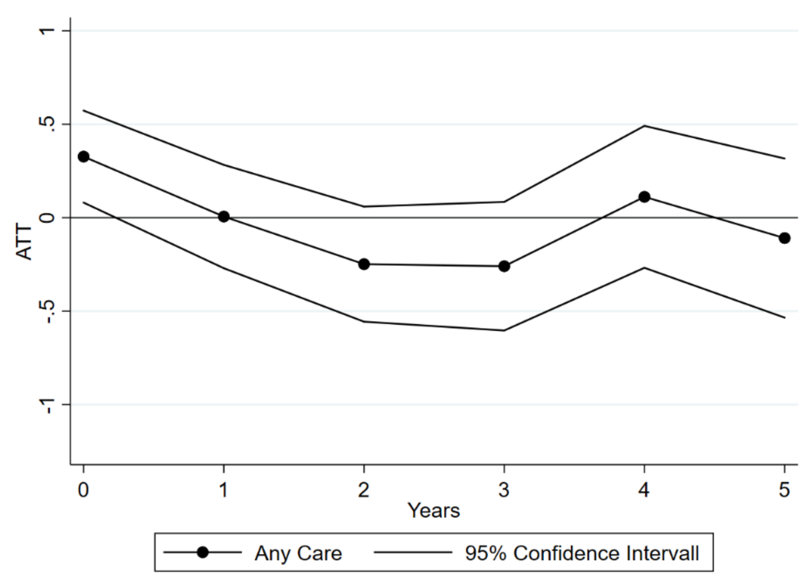
The results of the static matching procedure will be presented graphically. All underlying estimates can be found in Appendix Table A1.1, A1.2 and A1.3. In the presentation of our static matching results in the main text we do not differentiate the estimates ATTs by gender, these results are however available in Table A1.1. In the baseline analysis we estimate the effect of any informal care provided irrespective of the reported care intensity. Figure 3 depicts the estimated ATT on both the (a) mental and (b) physical health scores across time without differentiating the reported intensity level of informal care. Point estimates are depicted by the connected dots while the corresponding confidence interval is indicated by the solid line. In the mental domain we estimate small immediate negative effects of -0.421 (p<0.01) at $t_0$ and -0.624 (p<0.001) at $t_1$. These effects persist also in latter periods up until $t_5$ while ranging between -0.383 (p<0.05) at $t_2$ and -0.672 (p<0.001) at $t_4$. In the physical domain baseline estimates indicate a small positive effect of 0.327 (p<0.001) at $t_0$ but no effects thereafter.

Figure 3: Baseline Results (any care intensity)

a) Mental Health



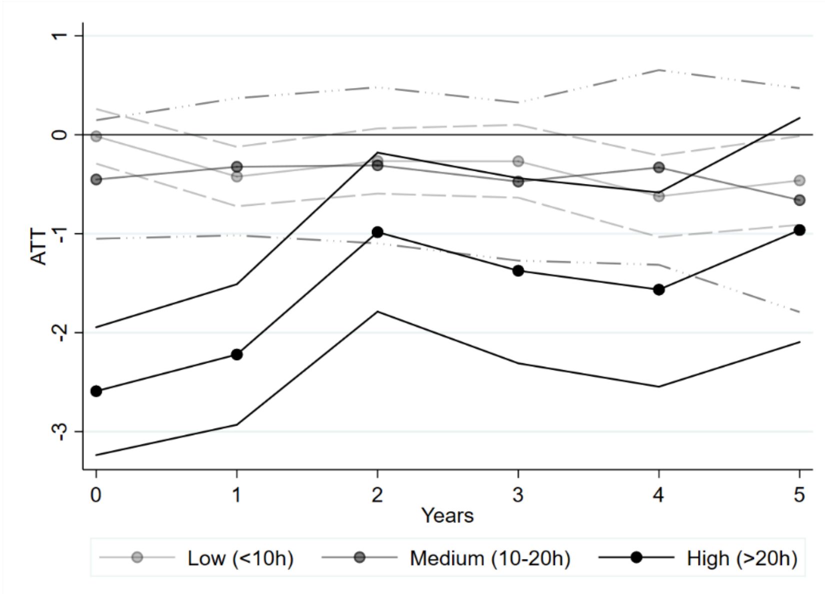

b) Physical Health



To explore heterogeneities in the estimated treatment effects we subdivide caregivers into three treatment groups according to the reported time spent caring per week. Figure 4 plots the estimation results when stratifying the treatment groups by care intensity for mental health (a) and physical health (b). Low intensity care provision for less than 10 hours per week is depicted in light grey, medium intensity care provision between 10 and 20 hours per week in dark grey, and high intensity caregiving of 20 hours and more in black.[9]

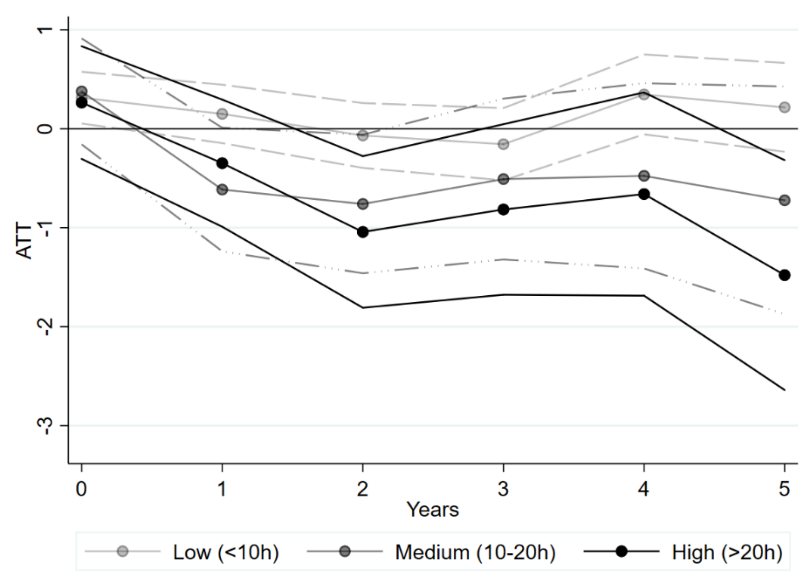
Figure 4: Treatment Effect by Care Intensity

a) Mental Health




[9] A small number of caregivers (92) provide no clear information on care intensity. These are excluded here.

b) Physical Health




The large heterogeneity in the estimated treatment effects underline the importance of care intensity for the experienced health effects. In the mental domain we find small effects for low intensity caregivers. Although the coefficient is nearly zero and insignificant at $t_0$ we find a negative effect of -0.423 (p<0.01) at $t_1$ and -0.622 (p<0.01) and -0.462 (p<0.05) at $t_4$ and $t_5$ with coefficients for the intermediate years being of similar size and direction but not significantly different from zero. For medium intensity caregiving the coefficients closely mirror those for low intensity care but are insignificant due to considerably larger standard errors. Among individuals caring for more than 20 hours per week we observe strong initial negative effects of -2.591 (p<0.001) at $t_0$ and -2.221 (p<0.001) at $t_1$. While these effects decrease for subsequent periods, they remain largely persistent with -0.984 (p<0.05), -1.374 (p<0.01), and -1.565 (p<0.01) at $t_2$ to $t_4$. The coefficient remains negative at $t_5$ and is of comparable size with -0.964 but only significant at the 10% level.

In the physical domain the pattern across care intensity levels is somewhat different. For low intensity caregivers we find a small but significant positive immediate effect of 0.314 (p<0.05) at $t_0$ while for the other intensity groups the coefficient is similar in size but insignificant. At subsequent periods the estimated effects vary considerably. Among low intensity caregivers coefficients are insignificant while varying around zero. For medium intensity care coefficients are larger and consistently negative, ranging from -0.476 to -0.773, with a significant effect of -0.760 ($p<0.05$) at $t_2$, while for high intensity caregivers they follow a similar pattern but are larger in size and indicating negative effects of -1.040 ($p<0.05$) and -1.479 ($p<0.05$) at $t_2$ and $t_5$. Especially the coefficients for high intensity care seem to follow a downward trend over time with differences emerging not immediately after care provision started but only in subsequent periods. We provide evidence in the Appendix (see Figures A2.1 and A2.2) that this pattern is partially driven by age-dependent physical health trends captured inadequately in the matching.[10] For the remainder of the discussion of our static matching results in the main body of the paper we will focus on the mental health outcomes. We do so as the observed pattern indicates that health effects disproportionally occur in the mental health domain while this pattern is also highly robust across all specifications. All corresponding results for physical health outcomes are reported in the Appendix. Additionally, in Appendix A1 we show that the reported heterogeneous health effects of care provision seem indeed driven by differences in care-intensity not the relationship between the caregiver and care-recipient.

*Static results – Family Effect*

Our results so far indicate the presence of strong negative caregiving effects in the mental health domain, especially among high intensity caregivers. However, an alternative explanation could be that this effect is not caused by the act of caregiving itself but rather by worries related to the health state and/or wellbeing of the care-recipient. This confounding effect has been identified as the so-called family effect (Bobinac et al., 2010), the impact of individuals caring *about* the care-recipients rather than caring *for* them. From a public perspective this distinction is highly important as both would require fundamentally different responses from policymakers. While the former might be mitigated through respite care allowing caregivers to manage the caregiving burden the latter might require counselling services instead.

[10] As propensity scores are a summary measure estimated based on many covariates these age-related trends are not guaranteed to be perfectly captured. For example, a younger individual might receive a high propensity score due to his/her physical health being low and/or other strong predictors but would be faced with an entirely different physical health trajectory in the short and medium term compared to an older individual. As illustrated in the Appendix this age-dependent trend is not a problem for mental health.

To explore the extent to which our estimated effects are driven by the caregiving effect or the family effect, we attempt to purify our estimates of the former from the latter. To achieve this, we make use of the fact that many caregivers provide informal care within their own household to individuals providing health information themselves. We restrict our analysis sample to those individuals cohabiting with living parents and/or a partner, with no parents alive outside their household. Among these we exclude individuals for which there is no available health information for living partners and parents at time points $t_{-1}$ and $t_0$. Based on the provided information we create a binary indicator for a cohabiting family member experiencing a health shock between these two periods. Health shocks are defined as a drop in either MCS or PCS of at least 10 points, equivalent to one standard deviation. If family members provide health information at $t_{-1}$ and are not self-reporting in $t_0$ due to (temporary) illness or old age as reported by a proxy respondent this is also coded as a health shock.[11] In addition, we use health information of partners and parents provided at $t_{-1}$ when re-estimating the propensity scores for this subsample.[12]

The outlined approach allows us to purify the effect of providing informal care from the family effect to the best of our abilities and given the available data. However, this means that we are not estimating both effects separately or quantify their relative magnitude. Instead the family effect is decomposed into an unobservable component related to the latent health of family members, captured in the propensity score re-estimation, and the observable health shock. The analysis sample is reduced considerably as we additionally drop propensity scores below the 5$^{th}$ and above the 95$^{th}$ percentile, to insure covariate balance. This leaves us with 2,878 individuals in the control group and 501 caregivers of which 165 provide high intensity care.

[11] Not doing this would lead to an underestimation of its direct effect as we only include less severe health shocks.

[12] The health information used includes MCS, PCS, self-assessed health, number of functional limitations, and the presence of a long-standing illness or disability for all cohabiting partners/parents.

Figure 5: Caregiving and Family Effects (high intensity care)

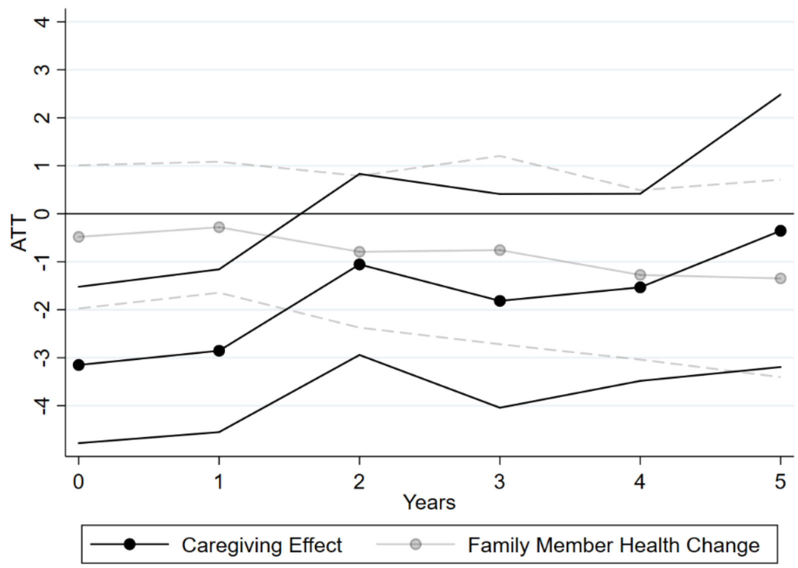



Figure 5 depicts the estimated caregiving effects for high-intensity caregivers while accounting for a health shock to a family-member. The estimated caregiving effects depict a similar pattern as before. The coefficient indicates a strong initial negative effect of -3.151 (p<0.001) at $t_0$ and -2.855 (p<0.01) at $t_1$. For later periods the coefficients remain negative and closely mirror the pattern reported for the full sample of high-intensity caregivers although they are insignificant due to the larger confidence intervals partially driven by the reduced sample size of high-intensity caregivers. The coefficient for the family health shock indicator is consistently negative but remains insignificant throughout all periods.[13]

*Dynamic results*

Next to estimating the impact of at least one year of care provision, we aim to investigate the impact of providing longer-term or additional years of informal care. For this dynamic matching approach, we estimate the propensity of (not) providing informal care at every decision point and drop scores in case the observation is off support or out of range. In Appendix A4 we report the

[13] Ideally, we would also like to explore whether the caregiving effect is reinforced when caring for someone experiencing a health shock by using an interaction term. However, as it is highly data demanding to correctly identify the coefficient of the interaction term this is not feasible given the reduced sample size.

propensity scores for the different care-trajectories as well as an overview of the observations that are excluded from the analysis.

In the main text we do not only provide the estimated effects across both genders but also separately for male and female caregivers. We do so for two reasons. Firstly, women in our sample are more likely to provide multiple years of informal care, with 30% of female caregivers providing three or more years of consecutive care compared to 24% of men.[14] Further, it has been hypothesized that there are psychological differences in the response to caregiving between both genders (Bookwala, 2009). As we focus on individuals providing care for multiple years in a row, samples for care trajectories become small and the number of individuals that are off support is higher compared to the static model. Due to the limited sample size, we are unable to estimate the impact of different care-trajectories among individuals that provide different intensity-levels of care separately.

The first three columns of Table 2 show the estimated health effect of different care trajectories for any intensity level of care provision compared to providing no care at all. The results from this dynamic subsample are similar to the static estimates: the impact of any informal care points in the negative direction for mental health. This negative mental health effect seems persistent throughout the different care trajectories, however, is only significant when comparing individuals that provided two years of care to those who did not do so. For these individuals we find an additional negative mental health effect of -0.93 ($p<0.01$) at $t_1$ which is slightly larger when focusing on females (-1.29, $p<0.01$).

Based on these results it seems that individuals do not adapt to caregiving tasks, however, the results have to be interpreted with caution as the observed pattern could also be driven by the fact that individuals entering into caregiving spells of more than one period are experiencing a higher burden to begin with, hence the stark difference when comparing one-time to continued caregivers. When estimating the impact of an additional year of care solely among caregivers (the last two columns of Table 2) we do not find any significant health effect of an additional year of informal care except for males providing three instead of two years of care where a large mental health decline is present. As we here solely compare longer-term caregivers with each other, the sample

[14] Calculated using only caregivers starting before wave 7 to allow all caregivers to hypothetically provide three years.

sizes become very small, which makes it difficult to draw conclusions regarding the impact of duration of care provision on health.

Table 2: Dynamic Matching Estimates

| **Duration:** | **1 vs. 0 ($t_0$)** | | **2 vs. 0 ($t_1$)** | | **3 vs. 0 ($t_2$)** | | **2 vs. 1 ($t_1$)** | | **3 vs. 2 ($t_2$)** | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** |
| **All** | -0.10 | -0.05 | -0.93 | 0.16 | -0.57 | 0.32 | -0.38 | 0.46 | -0.50 | 1.18 |
| | (0.24) | (0.22) | (0.32) | (0.31) | (0.45) | (0.42) | (0.46) | (0.45) | (0.80) | (0.71) |
| **Male** | 0.04 | -0.16 | -0.56 | 0.12 | -0.37 | 0.65 | 0.01 | 0.22 | **-2.63*** | 0.94 |
| | (0.33) | (0.33) | (0.45) | (0.41) | (0.65) | (0.54) | (0.67) | (0.63) | (1.08) | (0.94) |
| **Female** | -0.15 | 0.09 | **-1.29**** | 0.17 | -0.72 | 0.23 | -0.85 | 0.66 | 0.73 | -0.85 |
| | (0.32) | (0.28) | (0.40) | (0.41) | (0.54) | (0.54) | (0.62) | (0.63) | (1.08) | (0.93) |
| **Conrol** | | 14,777 | | 13,748 | | 12,865 | | 606 | | 185 |
| **Treatment** | | 1,208 | | 593 | | 348 | | 593 | | 348 |

* p < 0.05, ** p < 0.01, *** p < 0.001. Standard error in parentheses.

# Robustness Checks

To assess the sensitivity of our estimates to choices in our empirical specification and violations of the underlying identifying assumptions we performed multiple robustness checks.[15] The observed mental health effects could be explained by an ongoing trend that started before informal care provision as individuals might anticipate future care-obligations based on observed deteriorating health states of family members. For a subset of caregivers we observe their mental and physical health for multiple periods prior to providing informal care. Appendix Figure A2.1 plots the mean MCS and PCS for all three intensity levels for up to four years prior to providing informal care. There is little evidence that the observed results are driven by a distinct negative trend in caregivers' mental health starting before the actual onset of care provision. As explained in Appendix A2, for physical health a slight downward sloping age-related health trend seems to be present.

Another concern might be the occurrence of other external shocks, such as a deceased family member, that affect mental health outcomes in the periods after caregiving started. Including information on a family member (mother, father or partner) being deceased in a given wave does not explain the observed treatment effects. The results are depicted in Table A2.1. Additionally, we conducted the simulation-based sensitivity analysis proposed by Ichino et al. (2008) to explore the robustness of our results to a violation of the conditional independence assumption (CIA). The detailed procedure is outlined in the Appendix and Figure A2.4 plots our static matching results for mental and physical health effects of high-intensity care provision and their upper- and lower-bound ATT estimates based on the simulations. The results show that even in case of a severe violation of the CIA, the initial health effect of intensive informal care provision remains negative and significant.

Further, we explore the existence of a potential downward bias underlying our results due to selective attrition. We follow De Zwart et al. (2017) by splitting our sample into two groups and re-estimating the initial treatment effects. For the first group we observe health states past $t_1$, while

[15] Appendix A2 additionally contains the results of a re-estimation of the treatment effects when separately matching by intensity of care and cropping the upper/lower 5% of propensity scores to explore the impact of the overall matching quality on our results. The results remain highly similar to our main specification.

for the second group we can only observe health immediately after providing informal care due to permanent attrition from the panel thereafter. Figure A2.5 plots the initial treatment effects for both groups and indicates that the attrition sample experiences more persistent negative mental directly before discontinuing their participation, but no physical health effects. These results indicate some evidence for selective attrition leading to a downward bias in our estimated treatment effects for mental health especially for latter periods.

While the SF12 component scores allow us to measure both mental and physical health, the interpretation of effect sizes is not straightforward. Instead we use two additional outcome measures from the mental health domain to explore whether our results are robust across measures. We consider a general subjective well-being measure, overall life-satisfaction, and a mental health screening questionnaire, the general health questionnaire (GHQ). Appendix Figure A2.7 and Table A2.5 present both the corresponding results from the static and dynamic matching. When considering alternative outcome measures our results remain generally the same, indicating an asymmetric effect on subjective-wellbeing and mental health especially among high-intensity caregivers. However, the results for GHQ scores depict a pronounced dose-response relationship, not observed when using MCS as the mental health measure. Further, these results also indicate that the decrease in mental health is economically relevant as the number of individuals with scores low enough to surpass screening thresholds increases substantially.

A concern for our dynamic matching approach stems from the fact that for the later waves we condition on a large set of covariates as all intermediate covariates at each node are included. To check whether our propensity score estimates are suffering due to the large set of covariates we are matching upon, we follow Lechner (2008) and condition on a smaller set of covariates capturing most recent information and limited information (only health-related covariates) from the previous decision nodes. The results from this alternative specification which are presented in Table A2.6 are similar to our main analysis. Additionally, we check whether our results are sensitive to more stringent regression adjustment by, next to conditioning on the full set of covariates from the previous wave, also conditioning on health-related covariates from the (if present) earlier waves. This again does not substantially alter our estimates (Table A2.7).

# Discussion & Conclusion

Providing informal care can have a negative effect on the health of the informal caregiver. Based on the current literature there is an insufficient understanding of how these effects persist over time, differ by care-intensity and duration, and whether observed mental health effects are attributable to caregiving itself or the caregivers concerns about their family-members' health. We answer these questions using a representative panel-survey by estimating the long-term and dynamic effects of caregiving on caregivers' health.

While early studies on cross-sectional data commonly report caregivers to have low physical health (Carretero et al., 2009), we only find mixed evidence for a causal relationship. On the one hand our estimates indicate that informal care leads to a small increase in caregivers physical health, however, this effect is short-lived and only found for caregivers providing less than 20 hours of weekly care. On the other hand, for individuals providing higher levels of care the physical health trajectory is downward sloping in the years after becoming a caregiver. While this is might be partially driven by a age-related trend that is not fully captured by the matching estimator, we find some indication that for younger caregivers high-intensity caregiving might be associated with small decreases in their physical health in subsequent years.

For caregivers' mental health outcomes, we find immediate and persisting negative effects of providing care. These effects are highly heterogeneous and almost exclusively incurred by high-intensity caregivers. The initial negative effects on mental health slowly decrease in size throughout the years but remain persistent up to four and five years after initial care provision depending on the specification. It hence seems that most individuals only slowly recover from or adapt to caregiving tasks. This finding stands in contrast to previous studies that, possibly due to higher selective attrition, only found direct effects (De Zwart et al., 2017) or effects up to the first three years of care provision (Schmitz & Westphal, 2015). Potentially, the limited attrition in our sample allowed for better insight into the persistence of caregiving effects over time. However, we do find evidence that selective attrition among those experiencing the strongest initial health losses could still lead to an underestimation of treatment effects at later time periods (see Appendix Figure A2.5).

This study, furthermore, also explored the health effect of providing subsequent years of care. Results indicate that the mental health effects of caregiving seem to persist throughout the care trajectory meaning that individuals do not adapt over time. The results however must be interpreted with caution as we rely on very small samples, especially when estimating the impact of an additional year of care among two groups of caregivers. Additionally, there is an increased concern for selective attrition biasing our results for longer-term caregivers as the individuals experiencing large negative health shocks might especially those to drop out of our sample.

Our large study sample allowed us to stratify our sample based on caregiver characteristics. Our results indicate that small average effects are driven by a small group of caregivers providing more than 20 hours of care per week. As mentioned in the data-section of this paper, high-intensity caregivers are underrepresented in our sample. As a result, one should be aware that these estimates might be downward biased. Among low (<10h per week) and medium (10-20h per week) intensity caregivers we find no detectable effects on mental health. This result stands in contrast to the results of Schmitz and Westphal (2015) who, using a similar method and outcome measure, find a clearer dose-response-relationship in their analysis of German caregivers.

As our dataset contains information regarding the characteristics of the care recipient and other cohabiting family members, we were able to explore whether the measured mental health effects are indeed driven by care provision instead of the caregiver's concerns for the recipient's health. Our results indicate that this so-called family effect is not the main force driving the observed mental health effects. As we rely on self-reported health measures to capture the family effect, our results might be driven by positive selection as family members need to be in sufficient health to respond to the survey. Our results are however in line with the findings of Bom et al. (2019b), who used administrative data to identify the family effect among a sample of Dutch caregivers and hence avoid the problem of limited information availability. The estimation of caregiving effects while accounting for family-members' health requires us to restrict our analysis sample to a small subset of predominantly spousal caregivers. While this is a restrictive subsample, we believe that our results still provide very insightful information. Firstly, the heterogeneous impact of informal care is largely concentrated among high intensity caregivers, a group that is mainly dominated by within-household spousal care. Secondly, especially among close relatives the family effect provides a potent alternative explanation for the measured deterioration in mental health.

While our dataset is rich, allowing us to explore the health effects of informal care provision along multiple dimensions, there are many questions we would like to answer but have to leave for future researchers to explore. Our main measure to differentiate care intensity is self-reported weekly caregiving hours. In principle, increased hours are likely to reflect a larger overall caregiving burden. However, the tasks performed by caregivers are highly disease-specific and play an important role in the experienced caregiving burden (Pearlin et al., 1990). Therefore, reported hours are an incomplete measure inadequately capturing an important source of the mental and physical strain associated with caregiving. Ideally future research would be able to have insight into the type of tasks performed next to the time spent caring. A related cause for uncertainty is the absence of information on why informal care was taken up and discontinued, a process that itself could affect especially mental health outcomes. Unfortunately, we cannot observe for all caregivers whether discontinuation is related to improved health of care recipients or their move into a permanent nursing home.

Furthermore, one might question the use of self-reported health measures and prefer, in our case unavailable, administrative information like medical claims data or hospital episode statistics to measure health effects. We would however like to argue that these self-reported health measurements are of added value as they are able to capture more subtle changes in health than information regarding health care usage. Additionally, administrative health care information might not capture all health changes as especially highly burdened caregivers might forego medical care. Foregoing medical care could be directly caused by the intensity of caregiving as well as the potential stigma associated with seeking help as a caregiver itself. In addition, our results remain unchanged when using alternative outcome measures (see Appendix Figure A2.3) and indicate that the reported effects are economically relevant from the individuals' point of view. This leaves us fairly confident that the reported mental health effects are of interest to policymakers wishing to assess the extent of spillover effects arising from the reliance on informal care to meet social care demands.

Lastly, in this paper we focus on the health effects of providing informal care irrespective of whether this care occurs alongside formal care. The USoc does not contain detailed information on the availability and utilization of social care and support services. Therefore, we cannot explore to what extent these services might serve as a complement to informal care or help mitigating the negative health effects in the medium and long run.

To conclude, our results confirm previous studies reporting negative mental health effects of informal care provision and show that the effects persist up to four or five years after initial care provision. Our estimates suggest that most UK caregivers do not experience adverse health comes after providing informal care. However, especially female, intensive and persisting caregivers show to be most strongly affected by informal caregiving. In addition, we document evidence that these effects are indeed driven by the uptake of informal care and not due to caregivers being burdened by the poor health of a family member. Our results provide useful insights for policymakers interested in weighting the advantages and disadvantages of increasingly shifting the burden of social care provision on the shoulders of informal caregivers. While informal care provides benefits to both public health care systems and care recipients themselves, targeted support, for example to high-intensity caregivers, might be necessary to offset negative effects among the most vulnerable caregivers.

# References


Age UK (2019a). Paying for homecare. https://www.ageuk.org.uk/information-advice/care/paying-for-care/paying-for-homecare/

Age UK (2019b). Paying for a care home. https://www.ageuk.org.uk/information-advice/care/paying-for-care/paying-for-a-care-home/

Barnay, T., & Juin, S. (2016). Does home care for dependent elderly people improve their mental health? *Journal of Health Economics*, *45*, 149-160. https://doi.org/10.1016/j.jhealeco.2015.10.008

Basu, A., & Meltzer, D. (2005). Implications of spillover effects within the family for medical cost-effectiveness analysis. *Journal of Health Economics*, *24*(4), 751-773.

Bauer, J.M., & Sousa-Poza, A. (2015). Impacts of Informal Caregiving on caregiver employment, health and family. *IZA Discussion Paper No. 8851*. http://hdl.handle.net/10419/110154

Bobinac, A., Van Exel, N. J. A., Rutten, F. F., & Brouwer, W. B. (2010). Caring for and caring about: disentangling the caregiver effect and the family effect. *Journal of Health Economics*, *29*(4), 549-556. https://doi.org/10.1016/j.jhealeco.2010.05.003

Bom, J., Bakx, P., Schut, F., & van Doorslaer, E. (2019a). The Impact of Informal Caregiving for Older Adults on the Health of Various Types of Caregivers. *The Gerontologist*, 59(5), e629-e642. https://doi.org/10.1093/geront/gny137

Bom, J., Bakx, P., Schut, F., & van Doorslaer, E. (2019b). Health effects of caring for and about parents and spouses. *The Journal of the Economics of Ageing*, *14*. https://doi.org/10.1016/j.jeoa.2019.100196

Bookwala, J. (2009). The impact of parent care on marital quality and well-being in adult daughters and sons. *Journal of Gerontology: Series B*, *64B*(3), 339–347. https://doi.org/10.1093/geronb/gbp018

Carretero, S., Garcés, J., Ródenas, F., & Sanjosé, V. (2009). The informal caregiver's burden of dependent people: Theory and empirical review. *Archives of Gerontology and Geriatrics*, *49*(1), 74-79. https://doi.org/10.1016/j.archger.2008.05.004

Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, *22*(1), 31-72. https://doi.org/10.1111/j.1467-6419.2007.00527.x

Carers UK (2016). State of Caring 2016. https://www.carersuk.org/for-professionals/policy/policy-library/state-of-caring-2016

Coe, N. B., Guo, J., Konetzka, R. T., & Van Houtven, C. H. (2019). What is the marginal benefit of payment-induced family care? Impact on Medicaid spending and health of care recipients. *Health Economics*, *28*(5), 678-692. https://doi.org/10.1002/hec.3873

Colombo, F., Llena-Nozal, A., Mercier, J., & Tjadens, F. (2011). Help Wanted? Providing and paying for long-term care. OECD Health Policy Studies, OECD Publishing. http://dx.doi.org/10.1787/9789264097759-en

Comas-Herrera, C., Pickard, L., Wittenberg, R., Malley, J., & King, D. (2010). The long-term care system for the elderly in England. European Network of Economic Policy Research Institutes. ENEPRI RESEARCH REPORT NO. 74 CONTRIBUTION TO WP1 OF THE ANCIEN PROJECT

European Commission (2018). Informal care in Europe: Exploring Formalisation, Availability and Quality.https://ec.europa.eu/social/main.jsp?catId=738&langId=en&pubId=8106&type=2&furtherPubs=no

Fernández, J.-L., Forder, J., Truckeschitz, B., Rokosova, M., & McDaid, D. (2009). *How can European states design efficient, equitable and sustainable funding systems for long-term care for older people?*, Policy Brief No. 11, World Health Organisation Europe, Copenhagen.

Glendinning, C. (2013). Long Term Care Reform in England: A Long and Unfinished Story in Reforms. In Ranci, C. & Pavolini, E. (Ed.) *Long-Term Care Policies in Europe.* New York: Springer.

Haley, W. E., & Pardo, K. M. (1989). Relationship of severity of dementia to caregiving stressors. *Psychology and Aging*, *4*, 389-392. https://doi.org/10.1037/0882-7974.4.4.389

Ichino, A., Mealli, F., & Nannicini, T. (2008). From temporary help jobs to permanent employment: what can we learn from matching estimators and their sensitivity?. *Journal of Applied Econometrics*, *23*(3), 305-327. https://doi.org/10.1002/jae.998

Lacey R.E., McMunn A., & Webb, E. (2019). Informal caregiving patterns and trajectories of psychological distress in the UK Household Longitudinal Study. *Psychological Medicine,* 49, 1652–1660. https://doi.org/10.1017/S0033291718002222

Lechner, M. (2009). Sequential causal models for the evaluation of labor market programs. *Journal of Business & Economic Statistics*, *27*(1), 71-83. https://doi.org/10.1198/jbes.2009.0006

Lechner, M. & Miquel, M. (2010). Identification of the Effects of Dynamic Treatments by Sequential Conditional Independence Assumptions. *Empirical Economics*, *39*(1), 111-137. https://doi.org/10.1007/s00181-009-0297-3

Leuven, E., & Sianesi, B. (2003). PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. Boston College Department of Economics, Statistical Software Components. Downloadable from http://ideas.repec.org/c/boc/bocode/s432001.html.

Lilly, M.B., Laporte, A. & Coyte, P.C. (2007). Labor market work and home care's unpaid caregivers: A systematic review of labor force participation rates, predictors of labor market withdrawal, and hours of work. *Milbank Quarterly*, *85*, pp. 641-690, https://doi.org/10.1111/j.1468-0009.2007.00504.x

de la Maisonneuve, C., & Martins, J. O. (2015). The future of health and long-term care spending. *OECD Journal: Economic Studies*, *2014*/1, 61-96. https://doi.org/10.1787/19952856

NHS (2018).When the council might pay for your care. https://www.nhs.uk/conditions/social-care-and-support-guide/money-work-and-benefits/when-the-council-might-pay-for-your-care/

OECD (2019). *Health at a Glance 2019: OECD Indicators*, OECD Publishing, Paris, https://doi.org/10.1787/4dd50c09-en.

ONS (2013). 2011 Census analysis: Unpaid care in England and Wales, 2011 and comparison with 2001. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthcaresystem/articles/2011censusanalysisunpaidcareinenglandandwales2011andcomparisonwith2001/2013-02-15#national-comparisons

ONS (2015). UK Health Accounts reference tables. - Long-term care expenditure by function and financing scheme in 2015. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthcaresystem/bulletins/ukhealthaccounts/2015

ONS (2017). Unpaid carers provide social care worth £57 billion. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthandlifeexpectancies/articles/unpaidcarersprovidesocialcareworth57billion/2017-07-10

Pearlin, L. I., Mullan, J. T., Semple, S. J., & Skaff, M. M. (1990). Caregiving and the stress process: An overview of concepts and their measures. *The Gerontologist*, *30*(5), 583-594. https://doi.org/10.1093/geront/30.5.583

Pinquart, M., & Sörensen, S. (2003a). Associations of Stressors and Uplifts of Caregiving With Caregiver Burden and Depressive Mood: A Meta-Analysis, *The Journals of Gerontology: Series B*, 58(2), 112–128. https://doi.org/10.1093/geronb/58.2.P112

Pinquart, M., & Sörensen, S. (2003b). Differences between caregivers and noncaregivers in psychological health and physical health: a meta-analysis. *Psychology and Aging*, *18*(2), 250-267. https://doi.org/10.1037/0882-7974.18.2.250

Rellstab, S., Bakx, P., García-Gómez, P., & van Doorslaer, E. (2020). The kids are alright - labour market effects of unexpected parental hospitalisations in the Netherlands. *Journal of Health Economics*, *69*. https://doi.org/10.1016/j.jhealeco.2019.102275

Robins, J. M., Hernan, M. A., & Brumback, B. (2000). Marginal Structural Models and Causal Inference in Epidemiology. *Epidemiology*, 11, 550–560. https://doi.org/10.1097/00001648-200009000-00011

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, *39*(1), 33-38. https://doi.org/10.2307/2683903

Rubin, D. B. (1973). The use of matched sampling and regression adjustment to remove bias in observational studies. *Biometrics*, *29*, 185-203. https://doi.org/10.2307/2529685

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, *66*(5), 688-701. https://doi.org/10.1037/h0037350

Rubin, D. B. (1979). Using multivariate matched sampling and regression adjustment to control bias in observational studies. *Journal of the American Statistical Association*, *74*, 318-328. https://doi.org/10.2307/2286330

Schmitz, H., & Westphal, M. (2015). Short-and medium-term effects of informal care provision on female caregivers' health. *Journal of Health Economics*, *42*, 174-185. https://doi.org/10.1016/j.jhealeco.2015.03.002

Schmitz, H., & Westphal, M. (2017). Informal care and long-term labor market outcomes. *Journal of Health Economics*, *56*, 1-18. https://doi.org/10.1016/j.jhealeco.2017.09.002

Townsend, A., Noelker, L., Deimling, G., & Bass, D. (1989). Longitudinal impact of interhouse caregiving on adult children's mental health. *Psychology and Aging*, *4*, 393-401. https://doi.org/10.1037/0882-7974.4.4.393

UK Government (2020). Carer's allowance. https://www.gov.uk/carers-allowance

University of Essex. Institute for Social and Economic Research. (2019). *Understanding Society: Waves 1-9, 2009-2018 and Harmonised BHPS: Waves 1-18, 1991-2009*. [data collection]. *12th Edition.* UK Data Service. SN: 6614, http://doi.org/10.5255/UKDA-SN-6614-13

Van Houtven, C. H., & Norton, E. C. (2004). Informal care and health care use of older adults. *Journal of Health Economics*, *23*(6), 1159-1180. https://doi.org/10.1016/j.jhealeco.2004.04.008

Van Houtven, C. H., & Norton, E. C. (2008). Informal care and Medicare expenditures: testing for heterogeneous treatment effects. *Journal of Health Economics*, *27*(1), 134-156. https://doi.org/10.1016/j.jhealeco.2007.03.002

Ware, J.E., Kosinski, M., & Keller, S.D. (1995). SF-12: How to score the SF-12 physical and mental health summary scales (2nd. ed). Boston, Massachusetts, The Health Institute, New England Medical Center

Yeandle, S. (2016). Caring for our carers: An international perspective on policy developments in the UK. *Juncture*, *23*(1), 57-62. https://doi.org/10.1111/j.2050-5876.2016.00896.x

de Zwart, P. L., Bakx, P., & van Doorslaer, E. K. A. (2017). Will you still need me, will you still feed me when I'm 64? The health impact of caregiving to one's spouse. *Health Economics*, *26*, 127-138. https://doi.org/10.1002/hec.3542

# A1 – Detailed Results

Table A1.1: Static Estimation Results Stratified by Gender and Care Intensity

**Full Sample**

| | t=0 | | t=1 | | t=2 | | t=3 | | t=4 | | t=5 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** |
| **Any Care** | **-0.421**** | **0.327**** | **-0.624***** | 0.006 | **-0.383*** | -0.248 | **-0.396*** | -0.260 | **-0.672***** | 0.111 | **-0.516*** | -0.109 |
| | (0.135) | (0.126) | (0.145) | (0.141) | (0.156) | (0.157) | (0.177) | (0.175) | (0.196) | (0.193) | (0.214) | (0.217) |
| **Low Intensity** | -0.016 | **0.314*** | **-0.423**** | 0.149 | -0.266 | -0.068 | -0.268 | -0.156 | **-0.622**** | 0.346 | **-0.462*** | 0.216 |
| <10h weekly care | (0.141) | (0.133) | (0.153) | (0.151) | (0.168) | (0.167) | (0.188) | (0.186) | (0.210) | (0.206) | (0.228) | (0.228) |
| **Medium Intensity** | -0.452 | 0.377 | -0.324 | -0.615 | -0.309 | **-0.760*** | -0.473 | -0.508 | -0.331 | -0.476 | -0.661 | -0.773 |
| 10-20h weekly care | (0.306) | (0.273) | (0.353) | (0.318) | (0.402) | (0.357) | (0.408) | (0.415) | (0.502) | (0.477) | (0.577) | (0.587) |
| **High Intensity** | **-2.591***** | 0.264 | **-2.221***** | -0.348 | **-0.984*** | **-1.040*** | **-1.374**** | -0.816 | **-1.565**** | -0.660 | -0.963 | **-1.479*** |
| >20h weekly care | (0.329) | (0.290) | (0.362) | (0.328) | (0.410) | (0.391) | (0.477) | (0.439) | (0.501) | (0.524) | (0.578) | (0.592) |
| **Control** | | **12,970** | | **12,290** | | **10,409** | | **9,590** | | **8,545** | | **8,050** |
| **Treatment** | | **6,848** | | **5,723** | | **4,802** | | **4,134** | | **3,193** | | **2,461** |
| Low | | 5,069 | | 4,264 | | 3,600 | | 3,121 | | 2,402 | | 1,852 |
| Medium | | 792 | | 684 | | 556 | | 485 | | 373 | | 289 |
| High | | 895 | | 699 | | 588 | | 473 | | 377 | | 285 |

**Females Only**

| | t=0 | | t=1 | | t=2 | | t=3 | | t=4 | | t=5 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** |
| **Any Care** | **-0.556**** | **0.432*** | **-0.792***** | -0.034 | -0.206 | **-0.419*** | **-0.469*** | **-0.702**** | **-0.775**** | -0.124 | **-0.624*** | -0.375 |
| | (0.187) | (0.169) | (0.200) | (0.190) | (0.221) | (0.213) | (0.247) | (0.235) | (0.268) | (0.259) | (0.294) | (0.286) |
| **Low Intensity** | -0.106 | **0.446*** | **-0.527*** | 0.155 | -0.115 | -0.257 | -0.341 | **-0.508*** | **-0.725*** | 0.130 | -0.442 | 0.056 |
| <10h weekly care | (0.196) | 0.179 | 0.212 | (0.204) | (0.233) | (0.227) | (0.262) | (0.249) | (0.287) | (0.278) | (0.311) | (0.303) |
| **Medium Intensity** | -0.611 | 0.314 | -0.374 | **-0.929*** | -0.549 | -0.353 | -0.288 | **-1.105*** | -1.097 | -0.430 | **-1.554*** | -1.018 |
| 10-20h weekly care | (0.421) | (0.374) | (0.466) | (0.423) | (0.514) | (0.473) | (0.545) | (0.541) | (0.676) | (0.620) | (0.763) | (0.750) |
| **High Intensity** | **-2.734***** | 0.349 | **-2.741***** | -0.396 | -0.394 | **-1.494**** | **-1.583*** | **-1.602**** | -1.245 | -1.126 | -1.321 | **-2.264**** |
| >20h weekly care | (0.433) | (0.375) | (0.470) | (0.426) | (0.522) | (0.505) | (0.629) | (0.574) | (0.641) | (0.636) | (0.778) | (0.778) |
| **Control** | | **6,839** | | **6,514** | | **5,540** | | **5,118** | | **4,540** | | **4,292** |
| **Treatment** | | **4,054** | | **3,408** | | **2,854** | | **2,464** | | **1,925** | | **1,487** |
| Low | | 2,954 | | 2,499 | | 2,098 | | 1,827 | | 1,416 | | 1100 |
| Medium | | 472 | | 415 | | 335 | | 293 | | 232 | | 176 |
| High | | 576 | | 451 | | 385 | | 308 | | 252 | | 189 |

**Males Only**

| | t=0 | | t=1 | | t=2 | | t=3 | | t=4 | | t=5 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** |
| **Any Care** | -0.164 | 0.248 | -0.301 | 0.076 | **-0.558*** | 0.058 | -0.137 | 0.420 | -0.489 | 0.542 | -0.295 | 0.326 |
| | (0.189) | (0.184) | (0.202) | (0.207) | (0.224) | (0.226) | (0.247) | (0.258) | (0.279) | (0.288) | (0.307) | (0.328) |
| **Low Intensity** | 0.149 | 0.213 | -0.206 | 0.157 | -0.445 | 0.267 | -0.048 | 0.387 | -0.507 | **0.714*** | -0.468 | 0.446 |
| <10h weekly care | (0.196) | (0.196) | (0.213) | (0.221) | (0.235) | (0.240) | (0.259) | (0.274) | (0.300) | (0.300) | (0.335) | (0.345) |
| **Medium Intensity** | -0.112 | 0.514 | -0.185 | -0.165 | 0.324 | **-1.498**** | -0.396 | 0.262 | 1.062 | -0.459 | 0.888 | -0.210 |
| 10-20h weekly care | (0.431) | (0.398) | (0.541) | (0.479) | (0.589) | (0.532) | (0.617) | (0.653) | (0.710) | (0.753) | (0.841) | (0.923) |
| **High Intensity** | **-2.197***** | 0.135 | **-1.263*** | -0.309 | **-1.881**** | -0.241 | -0.774 | 0.468 | **-2.223**** | 0.334 | -0.196 | -0.095 |
| >20h weekly care | (0.499) | (0.463) | (0.551) | (0.500) | (0.652) | (0.609) | (0.716) | (0.644) | (0.810) | (0.912) | (0.803) | (0.974) |
| **Control** | | **6,131** | | **5,776** | | **4,869** | | **4,472** | | **4,005** | | **3,758** |
| **Treatment** | | **2,794** | | **2,315** | | **1,948** | | **1,670** | | **1,268** | | **974** |
| Low | | 2,117 | | 1,765 | | 1,502 | | 1,294 | | 986 | | 752 |
| Medium | | 320 | | 269 | | 221 | | 192 | | 141 | | 113 |
| High | | 319 | | 248 | | 203 | | 165 | | 125 | | 96 |

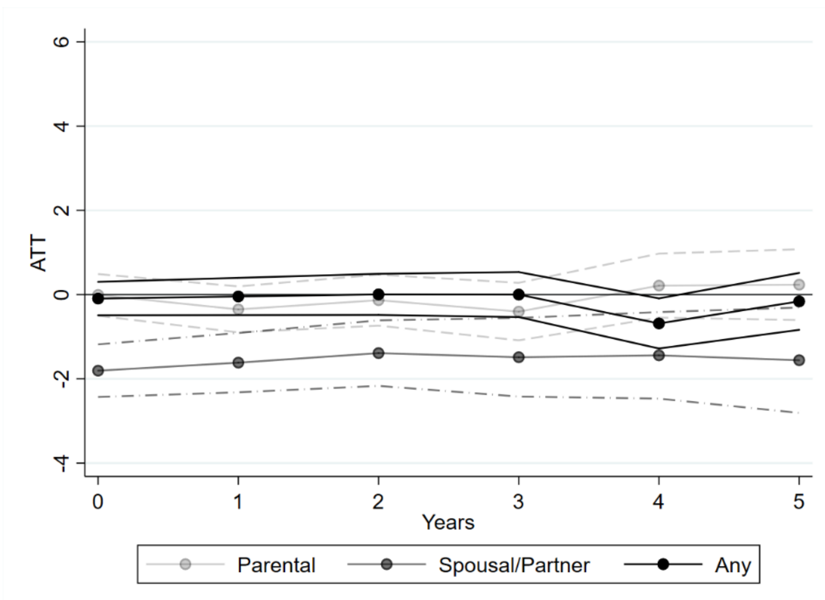
*p<0.05, **p<0.01, ***p<0.001, standard errors in parentheses

*Results by Caregiver-Recipient Relationship*

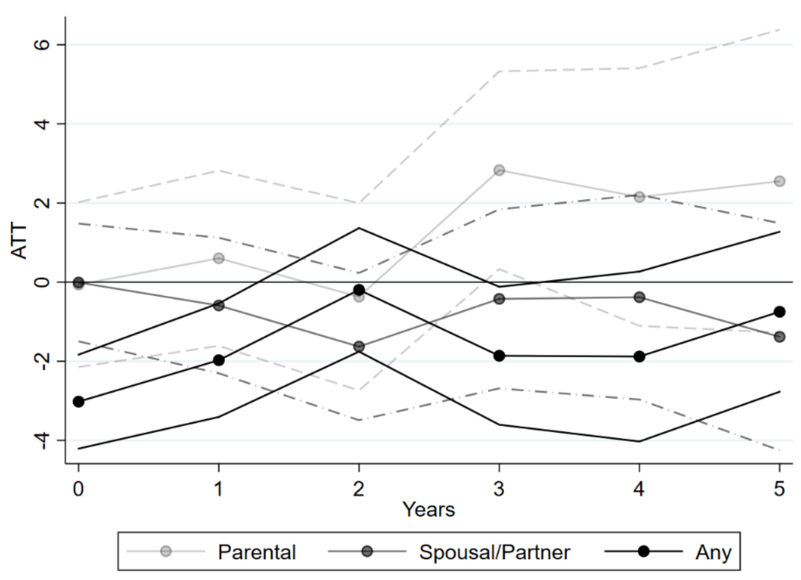
As the negative health effects are mainly in the mental domain, a potential alternative explanation is that these differences are not driven by informal care and its intensity itself but by the relationship between caregiver and care-recipient. To explore this, we identify caregivers according to their relationship to the care recipient as those providing care to a spouse or partner, one of their parents or anybody else. Most caregivers in our sample provide care within their own family, 14.6% care for a partner or spouse and 41.3% provide parental care. Figure 5 presents the results by caregiver-recipient relationship by including an interaction term capturing the caregiver-recipient-relationship for all caregivers (a) and among stratified by high-intensity care only (b).

Figure A1.1: Treatment Effect by Care-Relationship (any and high-intensity care)

a) Mental Health – Any Care



b) Mental Health – High Intensity Care




The results across all caregivers suggest that mental health effects are mainly driven by the group of spousal caregivers which is the only group experiencing strong negative health effects. The coefficient of the interaction term for spousal care is consistently negative and significant for all periods, ranging from -1.388 (p<0.001) at $t_2$ and -1.805 (p<0.001) at $t_0$. Both the treatment dummy and parental care interaction term depict insignificant coefficients close to zero, with the only exception being the treatment dummy for $t_4$ being -0.684 (p<0.05). However, across these groups there is a strong imbalance of informal care intensity. Among spousal caretakers 36% provide high intensity care while only 7% of parental caregivers provide more than 20h of care per week. When focusing on the group of high-intensity caregivers there seems to be little evidence for a consistent direct effect of the caregiver-recipient relationship.

## Table A1.2: Treatment Effects Including Care-Relationship Interaction Terms

**Any Care**

| | t=0 | | t=1 | | t=2 | | t=3 | | t=4 | | t=5 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** |
| **Any Care** | -0.093 | 0.124 | -0.044 | -0.049 | 0.006 | -0.237 | 0.001 | -0.523 | **-0.684*** | 0.028 | -0.161 | 0.026 |
| | (0.202) | (0.195) | (0.225) | (0.229) | (0.249) | (0.244) | (0.272) | (0.283) | (0.304) | (0.315) | (0.345) | (0.350) |
| **Spousal Care** | **-1.805*** ** | -0.523 | **-1.615*** ** | -0.657 | **-1.388*** ** | -0.467 | **-1.486** ** | -0.422 | **-1.442** ** | -0.290 | **-1.557*** | -0.552 |
| | (0.318) | (0.309) | (0.359) | (0.343) | (0.397) | (0.408) | (0.476) | (0.473) | (0.524) | (0.545) | (0.638) | (0.607) |
| **Parental Care** | -0.008 | 0.073 | -0.350 | 0.190 | -0.132 | 0.023 | -0.403 | **0.680*** | 0.214 | 0.097 | 0.236 | 0.007 |
| | (0.253) | (0.233) | (0.278) | (0.275) | (0.309) | (0.299) | (0.348) | (0.339) | (0.387) | (0.382) | (0.428) | (0.429) |
| **Control** | | **12,970** | | **12,290** | | **10,409** | | **9,590** | | **8,545** | | **8,050** |
| **Treatment** | | **6,848** | | **5,723** | | **4,802** | | **4,134** | | **3,193** | | **2,461** |
| Spousal Care | | 1,002 | | 813 | | 663 | | 536 | | 422 | | 316 |
| Parental Care | | 2,826 | | 2,406 | | 2,040 | | 1,776 | | 1,389 | | 1,094 |

**High Intensity Care**

| | t=0 | | t=1 | | t=2 | | t=3 | | t=4 | | t=5 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** |
| **Any Care** | **-3.020*** ** | -0.241 | **-1.972** ** | -0.240 | -0.194 | **-1.500*** | **-1.860*** | -1.460 | -1.880 | -0.296 | -0.748 | -1.100 |
| | (0.606) | (0.533) | (0.732) | (0.652) | -(1.628) | (0.722) | (0.889) | (0.773) | (1.095) | (0.984) | (1.031) | (1.011) |
| **Spousal Care** | -0.009 | 0.566 | -0.590 | -0.006 | -1.628 | 0.241 | -0.423 | 1.086 | -0.381 | 0.227 | -1.381 | 0.768 |
| | (0.759) | (0.697) | (0.874) | (0.803) | (0.949) | (0.941) | (1.154) | (1.016) | (1.319) | (1.302) | (1.464) | (1.364) |
| **Parental Care** | -0.062 | 0.344 | 0.604 | -0.403 | -0.369 | 2.043 | **2.830*** | 0.629 | 2.151 | 0.058 | 2.550 | 1.680 |
| | (1.063) | (0.838) | (1.130) | (0.994) | (1.207) | (1.114) | (1.274) | (1.164) | (1.661) | (1.509) | (1.953) | (1.475) |
| **Control** | | **12,970** | | **12,290** | | **10,409** | | **9,590** | | **8,545** | | **8,050** |
| **Treatment** | | 895 | | 699 | | 588 | | 473 | | 377 | | 285 |
| Spousal Care | | 361 | | 279 | | 232 | | 179 | | 147 | | 107 |
| Parental Care | | 199 | | 161 | | 141 | | 110 | | 91 | | 65 |

*p<0.05, ** p<0.01, ***p<0.001, standard errors in parentheses

## Table A1.3: Caregiving Effect and Family Health Shocks

| | t=0 | | t=1 | | t=2 | | t=3 | | t=4 | | t=5 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** |
| **High Intensity** | **-3.151*** ** | 0.469 | **-2.855** ** | -0.148 | -1.057 | -1.200 | -1.816 | -0.398 | -1.533 | 1.868 | -0.355 | 0.578 |
| >20h weekly care | (0.831) | (0.799) | (0.865) | (0.872) | (0.962) | (1.200) | (1.136) | (1.180) | (0.994) | (1.264) | (1.448) | (1.566) |
| **Family Health Shock** | -0.482 | 0.701 | -0.281 | 0.723 | -0.793 | 0.787 | -0.757 | -0.119 | -1.275 | 0.428 | -1.347 | 1.631 |
| | (0.761) | (0.705) | (0.696) | (0.756) | (0.806) | (0.918) | (1.000) | (1.008) | (0.899) | (1.003) | (1.049) | (1.084) |
| **Control** | | **2,878** | | **2,708** | | **2,264** | | **2,058** | | **1,879** | | **1,726** |
| **Treatment** | | 165 | | 134 | | 110 | | 83 | | 73 | | 46 |
| **Family Health Shock** | | 683 | | 626 | | 518 | | 442 | | 417 | | 374 |

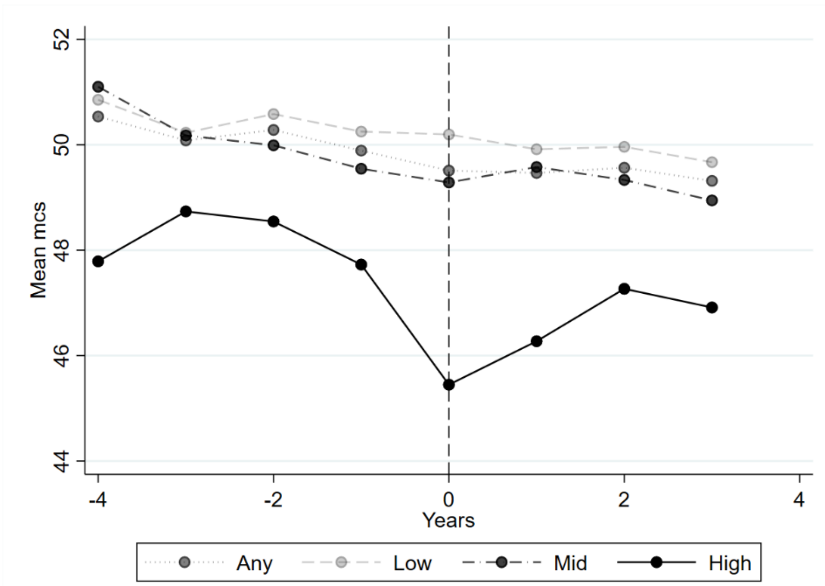
*p<0.05, **p<0.01, ***p<0.001, standard errors in parentheses

# A2 – Additional Analyses and Robustness Checks

*Pre-Treatment Outcomes & Physical Health Effects driven by Age-related Trends*

Figure A2.1: Mean MCS/PCS before Informal Care onset by Intensity

a) Mental Health




b) Physical Health

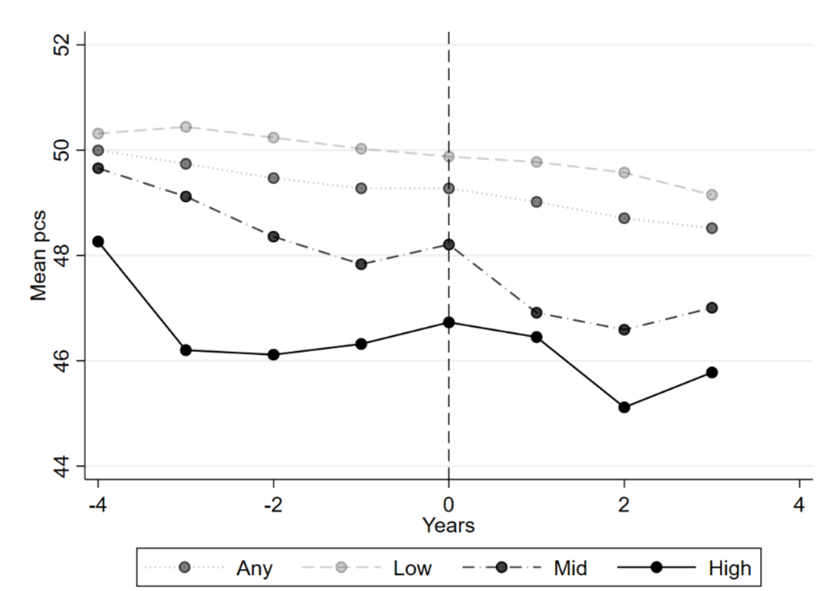


Figure A2.1 plots the mean mental and physical component scores for those caregivers whose health outcomes are observed for at least two periods prior informal care is provided. This leaves 61.4% of all caregivers (4207) and 70.1% (627) of high-intensity caregivers in the sample. Note that the plotted means are therefore based on an unbalanced panel in which all individuals are observed for at least three consecutive periods ($t_{-2}$ to $t_0$).

Figure A2.2: Age Distribution by Caregiving Status & Health Trends

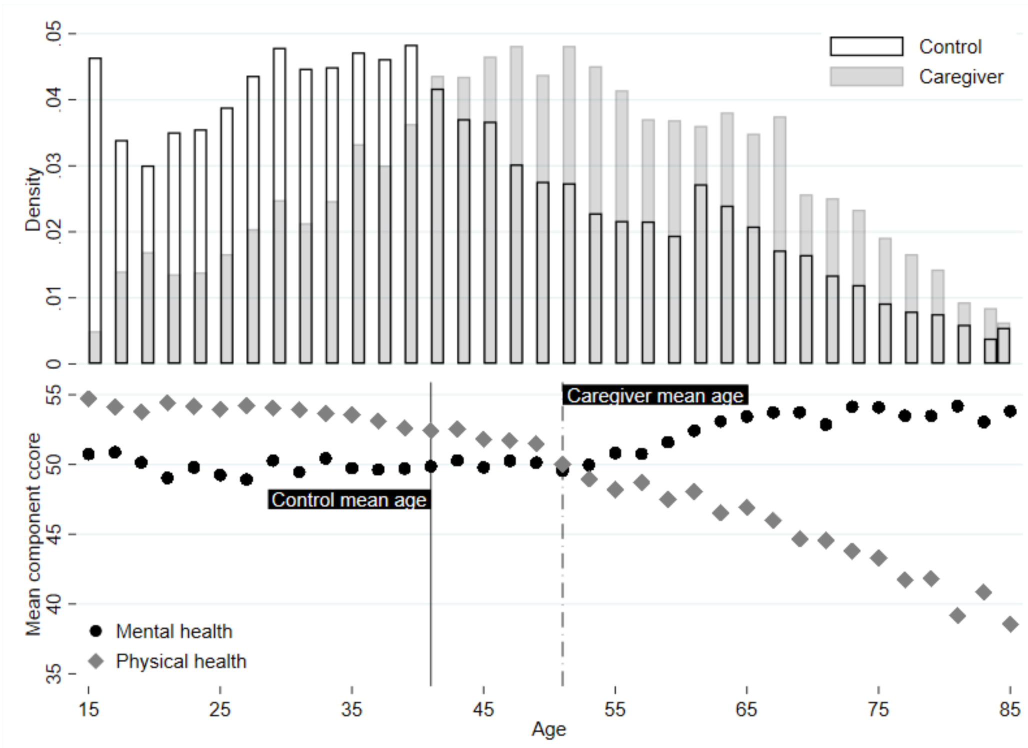



Figure A2.2 plots the age-distribution of caregivers and non-caregivers as well as the mean mental and physical component scores by age in two-year groups and using information of all individuals before any caregiving occurs and for the unmatched sample. The figure illustrates why we suspect that the physical health effects documented among (high-intensity) caregivers are partially driven by the fact that our matching approach inadequately captures ageing-related physical health trends across the lifecycle. We explore this explanation by separately re-estimating the treatment effect by caregiving intensity by age group, splitting the sample into those aged 50+ and those below 50.[16] In both groups we observe very similar mental health effects representative of the full sample

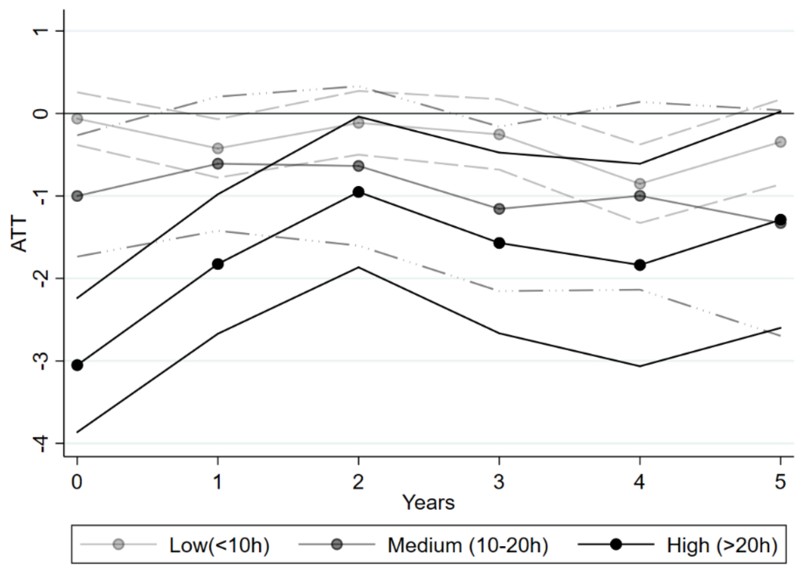
[16] We have also tried other splitting points (40, 45, 55, 60) with all results supporting the described relationship.

estimates. However, for physical health the coefficients for all intensity levels are close to zero and insignificant when estimated among older individuals. For younger caregivers on the contrary the physical health effects for medium- and high-intensity care indicate that there are some physical health effects at periods past $t_2$.[17]

*Propensity Score Matching by Intensity and Selective Propensity Score Inclusion*

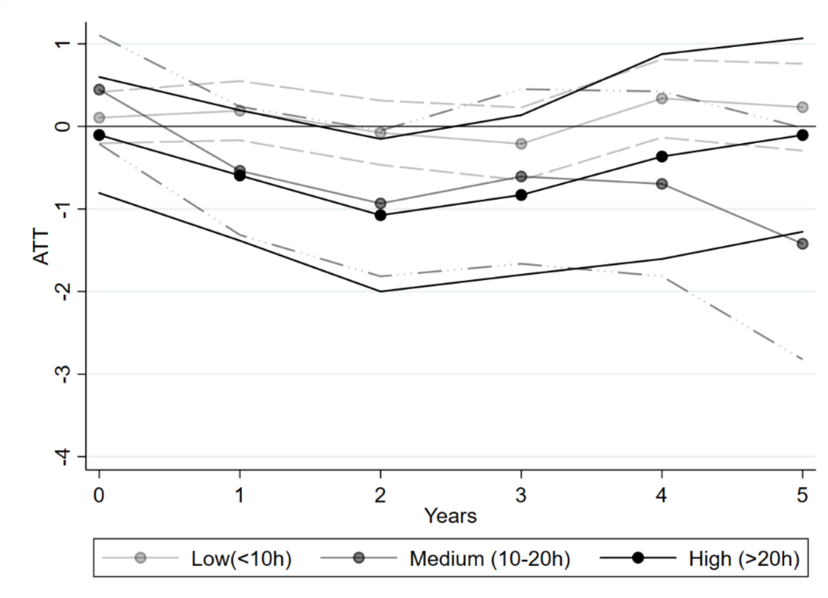
Figure A2.3: Static Results by Intensity and Selective Propensity Score Inclusion

a) Mental Health




[17] Detailed results are available upon request.

b) Physical Health




*Bereavement Effects*

We use the available information on the living status of close family members (parents or partner) to explore whether the observed treatment effects are explained by the death of a close family member. As the group of informal caregivers is older than the overall sample and more likely to have family members in poor health they are also more likely to experience a death in the family during the observation period. As in our main specification we do not account for such shocks this could lead to a wrongful attribution of the differences in (mental) health outcomes between treatment and control groups due to providing informal care and not the experienced loss of a family member. Indeed, high-intensity informal caregivers are more likely to experience the death of a close family member. While within the control group around 6.93% of all individuals experience the death of a close family member between $t_0$ and $t_5$, this rate is 11.15% among high-intensity caregivers. Table A2.1 reports the static matching results when including a binary indicator for a given family member being reported for the first time being deceased at a given $t$. However, given that the estimated treatment effects remain largely unaffected the difference in mental health outcomes between caregivers and control group are not explained by such time-varying shocks.

Table A2.1: Treatment and Bereavement Effects on Mental Health

| | t=0 | t=1 | t=2 | t=3 | t=4 | t=5 |
|---|---|---|---|---|---|---|
| **High Intensity** | **-3.043*** ** | **-2.001*** ** | **-1.036*** | **-1.449** ** | **-1.520** ** | -0.852 |
| >20h weekly care | (0.382) | (0.413) | (0.452) | (0.539) | (0.580) | (0.651) |
| **Death in the Family** | **-4.474*** ** | **-3.110** ** | **-6.150*** ** | **-6.598*** ** | **-5.043*** ** | **-5.920*** ** |
| | (1.231) | (1.094) | (1.644) | (1.430) | (1.400) | (1.287) |
| **Control** | **12,970** | **12,290** | **10,409** | **9,590** | **8,545** | **8,050** |
| **Treatment** | 895 | 699 | 588 | 473 | 377 | 285 |
| **Death in Family** | 163 | 185 | 177 | 170 | 159 | 153 |

*p<0.05, **p<0.01, ***p<0.001, standard errors in parentheses

*Simulated Violation of the Conditional Independence Assumption*

In this sensitivity analysis we focus on high-intensity caregivers and mental health outcomes to test the robustness of our main result. Our estimation strategy relies on the conditional independence assumption (CIA) stating that selection into treatment is driven only by observable variables. We use a rich set of covariates covering multiple dimensions relevant to the selection into informal caregiving and mental health in the propensity score estimation, hence we argue that the CIA is likely to hold. However, the CIA is untestable and unobserved variables might be influencing the selection into informal caregiving and mental health outcomes, thereby biasing our estimates. To assess the robustness of our estimates to such a violation we follow Ichino et al. (2008) who propose a simulation-based sensitivity analysis for matching estimators. We only roughly sketch the underlying procedure and intuition behind the procedure. A more elaborate discussion can be found in Ichino et al. (2008).

Consider the conditional independence assumption in our context:

$$Y_t^0 \perp\!\!\!\perp T \mid X_{-1}, Y_{-1} \ \forall t$$

After conditioning on a set of pre-treatment control ($X_{-1}$) and outcome variables ($Y_{-1}$) the potential outcomes in a given period $t$ in absence of informal care provision ($Y_t^0$) across treatment and control groups are the same. Now consider that this assumption is violated due to a confounder $U$. If we additionally condition on this confounder the CIA would be satisfied:

$$Y_t^0 \perp\!\!\!\perp T \mid X_{-1}, Y_{-1}, U \;\; \forall\, t$$

Ichino et al. (2008) outline a sensitivity analysis that simulates a binary $U$ in the context of a binary outcome variable. A binary $U$ is attractive as its distribution can be expressed by four parameters $p_{ij}$ where $i$ indicates the treatment status (0,1) and $j$ indicates the outcome status (0,1). For continuous outcomes they propose a transformation of the continuous outcome:

$$\hat{Y} = \begin{cases} 1, & Y > \bar{Y} \\ 0, & else \end{cases}$$

In our case we set $\bar{Y}$ equal to the sample mean of MCS across high-intensity caregivers and the overall control group. The four $p_{ij}$ that determine the distribution of $U$ are defined as:

$$\begin{aligned} p_{01} &= \Pr(U = 1 \mid T = 0, \hat{Y} = 1) \\ p_{00} &= \Pr(U = 1 \mid T = 0, \hat{Y} = 0) \\ p_{11} &= \Pr(U = 1 \mid T = 1, \hat{Y} = 1) \\ p_{10} &= \Pr(U = 1 \mid T = 1, \hat{Y} = 0) \end{aligned}$$

These $p_{ij}$ describe the distribution of U:

$$\begin{aligned} Pr(U = 1) = {} & p_{11} \Pr\big(\hat{Y} = 1 \big| T = 1\big) \Pr(T = 1) + p_{10} \Pr\big(\hat{Y} = 0 \big| T = 1\big) \Pr(T = 1) \\ & + p_{01} \Pr\big(\hat{Y} = 1 \big| T = 0\big) \Pr(T = 0) + p_{00} \Pr\big(\hat{Y} = 0 \big| T = 0\big) \Pr(T = 0) \end{aligned}$$

Ichino et al. (2008) propose to choose the values of $p_{ij}$ in such a way to deliberately control for the selection effect $s$ of $U$ on treatment uptake and the outcome effect $d$ of $U$ on the probability to observe $\hat{Y} = 1$. These effects are defined as:

$$s = p_{1\cdot} - p_{0\cdot}$$

where

$$p_{i\cdot} = \Pr(U = 1 \mid T = i) = p_{i0} * \Pr\big(\hat{Y} = 1 \big| T = i\big) + p_{i1} * \Pr\big(\hat{Y} = 1 \big| T = i\big) \quad with\; i \in \{0,1\}$$

The outcome effect is defined as

$$d = p_{01} - p_{00}$$

The sensitivity analysis is then conducted by choosing a set of $p_{ij}$ that model a confounder with specific selection and outcome effects. Given these values for $s$ and $d$ one has to assume a value for $\Pr(U = 1)$ and the relationship between $p_{11}$ and $p_{01}$ in order to be able to solve the equations for $\Pr(U)$, $d$, and $s$ for all $p_{ij}$. We follow the example given by Ichino et al. (2008) by assuming $\Pr(U = 1) = 0.5$ and $p_{11} - p_{01} = 0$ to solve for all of $p_{ij}$ given the empirically observed $\Pr(\hat{Y} = i | T = i)$ and $\Pr(T = i)$.

As an example, consider an unobserved confounder with negative selection ($s < 0$) and positive outcome effect ($d > 0$). This might be the unobserved availability of resources to purchase formal care services by the recipient itself. As informal care is largely done by family-members the availability of resources is likely to decrease the likelihood of (high-intensity) informal care provision while increasing mental health in the absence of treatment. By not accounting for such a confounder our estimation would therefore underestimate the effect of informal care on mental health. By selecting the magnitude and direction of selection and outcome effects and the corresponding $p_{ij}$ the effect of $U$ is simulated by drawing repeatedly from a Bernoulli distribution with the desired distributional properties. Robust estimates of both the ATT and the corresponding standard errors are then given by their averages across these simulations.

To calibrate the sensitivity analysis, we need a starting point of selection and outcome effects to obtain the parameters $p_{ij}$ for the distribution of a realistic confounder. Ichino et al. (2008) recommend inspecting the selection and outcome effects of important covariates in the propensity score estimation in order to find reasonable values for the simulation. Table A2.2 depicts the estimated effects ($d$ and $s$) and the parameters ($p_{ij}$) for all covariates used in the estimation of propensity scores and using the mental component scores as the outcome of interest. These are obtained by using a customized version of the user-written command for Stata by Nannicini (2007)[18] implementing the sensitivity analysis proposed by Ichino et al. (2008). Please note that

[18] Nannicini, T. (2007). Simulation-based sensitivity analysis for matching estimators. *The Stata Journal*, *7*(3), 334-350.

continuous variables were adapted to comply with the binary nature of $U$ by transforming them into categorical variables.

Table A2.2: Estimated Selection and Outcome Effects

| Variable | p11 | p10 | p01 | p00 | p1. | p0. | s | d |
|---|---|---|---|---|---|---|---|---|
| **Care Obligations** | | | | | | | | |
| Mother alive | 0.48 | 0.56 | 0.66 | 0.75 | 0.53 | 0.70 | -0.17 | -0.09 |
| Mother aged <49 | 0.06 | 0.06 | 0.12 | 0.15 | 0.06 | 0.13 | -0.07 | -0.03 |
| Mother aged 50-59 | 0.13 | 0.15 | 0.21 | 0.24 | 0.14 | 0.22 | -0.08 | -0.03 |
| Mother aged 60-69 | 0.12 | 0.17 | 0.20 | 0.22 | 0.15 | 0.21 | -0.06 | -0.02 |
| Mother aged 70-79 | 0.08 | 0.12 | 0.10 | 0.10 | 0.11 | 0.10 | 0.01 | 0.00 |
| Mother aged 80-89 | 0.08 | 0.05 | 0.03 | 0.03 | 0.06 | 0.03 | 0.03 | 0.00 |
| Mother aged >90 | 0.91 | 0.94 | 0.87 | 0.84 | 0.93 | 0.86 | 0.07 | 0.03 |
| Father alive | 0.36 | 0.48 | 0.60 | 0.68 | 0.43 | 0.63 | -0.20 | -0.08 |
| Father aged <49 | 0.03 | 0.03 | 0.08 | 0.10 | 0.03 | 0.09 | -0.06 | -0.02 |
| Father aged 50-59 | 0.10 | 0.13 | 0.17 | 0.21 | 0.12 | 0.19 | -0.07 | -0.04 |
| Father aged 60-69 | 0.11 | 0.14 | 0.20 | 0.22 | 0.13 | 0.21 | -0.08 | -0.02 |
| Father aged 70-79 | 0.05 | 0.12 | 0.11 | 0.11 | 0.09 | 0.11 | -0.02 | 0.00 |
| Father aged 80-89 | 0.05 | 0.06 | 0.03 | 0.02 | 0.05 | 0.03 | 0.02 | 0.01 |
| Father aged >90 | 0.95 | 0.96 | 0.90 | 0.89 | 0.96 | 0.90 | 0.06 | 0.01 |
| Both parents alive | 0.32 | 0.41 | 0.56 | 0.64 | 0.38 | 0.60 | -0.22 | -0.08 |
| Living siblings | 0.84 | 0.86 | 0.87 | 0.90 | 0.85 | 0.88 | -0.03 | -0.03 |
| Living partner | 0.73 | 0.75 | 0.67 | 0.59 | 0.74 | 0.64 | 0.10 | 0.08 |
| Partner aged <29 | 0.04 | 0.05 | 0.10 | 0.11 | 0.05 | 0.10 | -0.05 | -0.01 |
| Partner aged 30-39 | 0.11 | 0.15 | 0.18 | 0.19 | 0.13 | 0.19 | -0.06 | -0.01 |
| Partner aged 40-49 | 0.10 | 0.16 | 0.15 | 0.15 | 0.14 | 0.15 | -0.01 | 0.00 |
| Partner aged 50-59 | 0.13 | 0.11 | 0.09 | 0.07 | 0.12 | 0.08 | 0.04 | 0.02 |
| Partner aged 60-69 | 0.13 | 0.11 | 0.10 | 0.05 | 0.12 | 0.08 | 0.04 | 0.05 |
| Partner aged 70-79 | 0.15 | 0.12 | 0.04 | 0.02 | 0.13 | 0.03 | 0.10 | 0.02 |
| Partner aged 80-89 | 0.08 | 0.05 | 0.00 | 0.00 | 0.06 | 0.00 | 0.06 | 0.00 |
| Partner aged >90 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| **Willingness to Care** | | | | | | | | |
| Own age <29 | 0.12 | 0.12 | 0.26 | 0.33 | 0.12 | 0.29 | -0.17 | -0.07 |
| Own age 30-39 | 0.15 | 0.23 | 0.23 | 0.24 | 0.19 | 0.23 | -0.04 | -0.01 |
| Own age 40-49 | 0.14 | 0.18 | 0.18 | 0.20 | 0.16 | 0.18 | -0.02 | -0.02 |
| Own age 50-59 | 0.16 | 0.16 | 0.12 | 0.11 | 0.16 | 0.12 | 0.04 | 0.01 |
| Own age 60-69 | 0.18 | 0.15 | 0.13 | 0.07 | 0.16 | 0.11 | 0.05 | 0.06 |
| Own age 70-79 | 0.19 | 0.11 | 0.07 | 0.04 | 0.15 | 0.06 | 0.09 | 0.03 |
| Own age 80-89 | 0.06 | 0.04 | 0.02 | 0.01 | 0.05 | 0.02 | 0.03 | 0.01 |
| Own age >90 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Female | 0.59 | 0.68 | 0.50 | 0.58 | 0.64 | 0.53 | 0.11 | -0.08 |
| Education: Primary/Other lower | 0.32 | 0.33 | 0.18 | 0.17 | 0.33 | 0.18 | 0.15 | 0.01 |
| Education: Secondary | 0.41 | 0.40 | 0.43 | 0.45 | 0.41 | 0.45 | -0.04 | -0.02 |
| Educaiton: Tertiary | 0.27 | 0.26 | 0.39 | 0.38 | 0.27 | 0.39 | -0.12 | 0.01 |
| Self employed | 0.06 | 0.04 | 0.07 | 0.06 | 0.05 | 0.07 | -0.02 | 0.01 |
| Unemployed | 0.07 | 0.07 | 0.04 | 0.06 | 0.07 | 0.05 | 0.02 | -0.02 |
| Employed | 0.35 | 0.35 | 0.58 | 0.55 | 0.35 | 0.57 | -0.22 | 0.03 |
| Working full-time | 0.28 | 0.24 | 0.51 | 0.49 | 0.25 | 0.50 | -0.25 | 0.02 |
| Retired | 0.36 | 0.27 | 0.17 | 0.10 | 0.31 | 0.14 | 0.17 | 0.07 |
| In education/other | 0.04 | 0.02 | 0.09 | 0.12 | 0.03 | 0.10 | -0.07 | -0.03 |
| Homecarer | 0.09 | 0.15 | 0.04 | 0.06 | 0.13 | 0.05 | 0.08 | -0.02 |
| Disabled | 0.03 | 0.10 | 0.01 | 0.05 | 0.07 | 0.02 | 0.05 | -0.04 |
| Income Quintile 1 (Lowest) | 0.24 | 0.27 | 0.18 | 0.23 | 0.25 | 0.20 | 0.05 | -0.05 |
| Income Quintile 2 | 0.26 | 0.25 | 0.19 | 0.20 | 0.25 | 0.20 | 0.05 | -0.01 |
| Income Quintile 3 | 0.21 | 0.23 | 0.21 | 0.19 | 0.22 | 0.20 | 0.02 | 0.02 |
| Income Quintile 4 | 0.16 | 0.16 | 0.21 | 0.20 | 0.16 | 0.20 | -0.04 | 0.01 |
| Income Quintile 5 (Highest) | 0.14 | 0.10 | 0.22 | 0.19 | 0.12 | 0.21 | -0.09 | 0.03 |
| HH Income Fraciton > 0.5 | 0.51 | 0.50 | 0.50 | 0.48 | 0.50 | 0.49 | 0.01 | 0.02 |
| Married | 0.62 | 0.62 | 0.51 | 0.42 | 0.62 | 0.48 | 0.14 | 0.09 |
| Single | 0.14 | 0.12 | 0.22 | 0.29 | 0.13 | 0.25 | -0.12 | -0.07 |
| Separated/Divorced | 0.09 | 0.10 | 0.07 | 0.08 | 0.10 | 0.07 | 0.03 | -0.01 |
| Widowed | 0.04 | 0.03 | 0.04 | 0.04 | 0.03 | 0.04 | -0.01 | 0.00 |
| Partnership | 0.11 | 0.13 | 0.16 | 0.17 | 0.12 | 0.16 | -0.04 | -0.01 |
| Children in HH | 0.30 | 0.40 | 0.38 | 0.42 | 0.36 | 0.39 | -0.03 | -0.04 |
| Young children in HH | 0.28 | 0.38 | 0.34 | 0.38 | 0.33 | 0.36 | -0.03 | -0.04 |
| Region: North-East | 0.04 | 0.05 | 0.04 | 0.04 | 0.05 | 0.04 | 0.01 | 0.00 |
| Region: North-West | 0.12 | 0.13 | 0.11 | 0.12 | 0.12 | 0.11 | 0.01 | -0.01 |
| Region: Yorkshire | 0.06 | 0.08 | 0.08 | 0.08 | 0.08 | 0.08 | 0.00 | 0.00 |
| Region: East Midlands | 0.08 | 0.09 | 0.07 | 0.08 | 0.08 | 0.08 | 0.00 | -0.01 |
| Region: West Midlands | 0.05 | 0.09 | 0.07 | 0.08 | 0.07 | 0.07 | 0.00 | -0.01 |
| Region: East England | 0.11 | 0.07 | 0.10 | 0.09 | 0.08 | 0.09 | -0.01 | 0.01 |
| Region: London | 0.08 | 0.08 | 0.09 | 0.11 | 0.08 | 0.10 | -0.02 | -0.02 |
| Region: South East | 0.09 | 0.10 | 0.14 | 0.13 | 0.10 | 0.14 | -0.04 | 0.01 |
| Region: South West | 0.09 | 0.07 | 0.09 | 0.08 | 0.08 | 0.09 | -0.01 | 0.01 |
| Region: Wales | 0.11 | 0.10 | 0.06 | 0.07 | 0.11 | 0.07 | 0.04 | -0.01 |
| Region: Scotland | 0.10 | 0.07 | 0.10 | 0.08 | 0.08 | 0.09 | -0.01 | 0.02 |
| Region: Northern Ireland | 0.06 | 0.07 | 0.05 | 0.05 | 0.07 | 0.05 | 0.02 | 0.00 |
| Living in urban area | 0.72 | 0.78 | 0.74 | 0.78 | 0.75 | 0.76 | -0.01 | -0.04 |
| Big 5: Openness | 0.55 | 0.48 | 0.55 | 0.53 | 0.51 | 0.55 | -0.04 | 0.02 |
| Big 5: Conscientiousness | 0.62 | 0.49 | 0.57 | 0.45 | 0.55 | 0.52 | 0.03 | 0.12 |
| Big 5: Extraversion | 0.56 | 0.48 | 0.56 | 0.49 | 0.52 | 0.54 | -0.02 | 0.07 |
| Big 5: Agreeableness | 0.67 | 0.62 | 0.60 | 0.55 | 0.64 | 0.58 | 0.06 | 0.05 |
| Big 5: Neuroticism | 0.36 | 0.64 | 0.38 | 0.68 | 0.52 | 0.50 | 0.02 | -0.30 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Ability to Care** | | | | | | | | |
| Self-Assessed Health | 0.46 | 0.67 | 0.37 | 0.52 | 0.59 | 0.43 | 0.16 | -0.15 |
| MCS | 0.72 | 0.33 | 0.78 | 0.41 | 0.49 | 0.63 | -0.14 | 0.37 |
| PCS | 0.52 | 0.43 | 0.72 | 0.66 | 0.46 | 0.69 | -0.23 | 0.06 |
| LSI | 0.35 | 0.48 | 0.25 | 0.34 | 0.43 | 0.29 | 0.14 | -0.09 |
| Number of Functional Limitations | 0.23 | 0.36 | 0.14 | 0.23 | 0.30 | 0.18 | 0.12 | -0.09 |
| Satisfaction with Health | 0.52 | 0.30 | 0.65 | 0.44 | 0.39 | 0.57 | -0.18 | 0.21 |
| Satisfaction with Income | 0.59 | 0.38 | 0.67 | 0.49 | 0.47 | 0.60 | -0.13 | 0.18 |
| Satisfaction with Leisure | 0.70 | 0.49 | 0.69 | 0.50 | 0.58 | 0.62 | -0.04 | 0.19 |
| Life Satisfaction | 0.69 | 0.38 | 0.73 | 0.46 | 0.51 | 0.63 | -0.12 | 0.27 |
| GHQ Score | 0.80 | 0.50 | 0.83 | 0.52 | 0.62 | 0.72 | -0.10 | 0.31 |

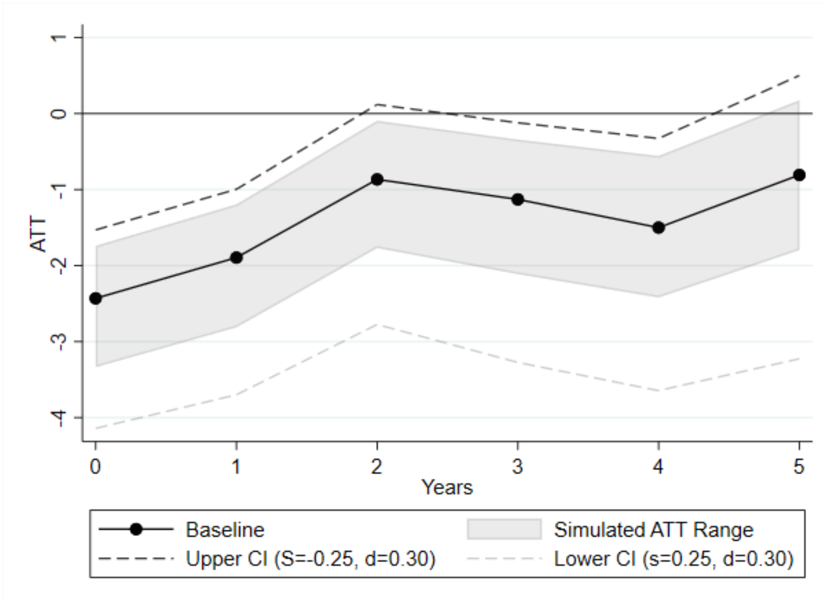
The largest outcome and selection effects in absolute terms are estimated for the set of pre-treatment health and well-being outcomes which is also precisely why we condition on pre-treatment outcomes. The second largest absolute selection and outcome effects are estimated for working full-time ($s = -0.25$) and scoring above four on the neuroticism seven-point scale ($d = -0.30$). For the sensitivity analysis we select two pairs of $s$ and $d$ to obtain upper and lower bounds for our ATT estimates. These are $s \in \{-0.25, 0.25\}$ and $d = 0.30$.[19]

Figure A2.3 plots our static matching results for mental health effects of high-intensity care provision and their upper- and lower-bound ATT estimates based on the simulations.

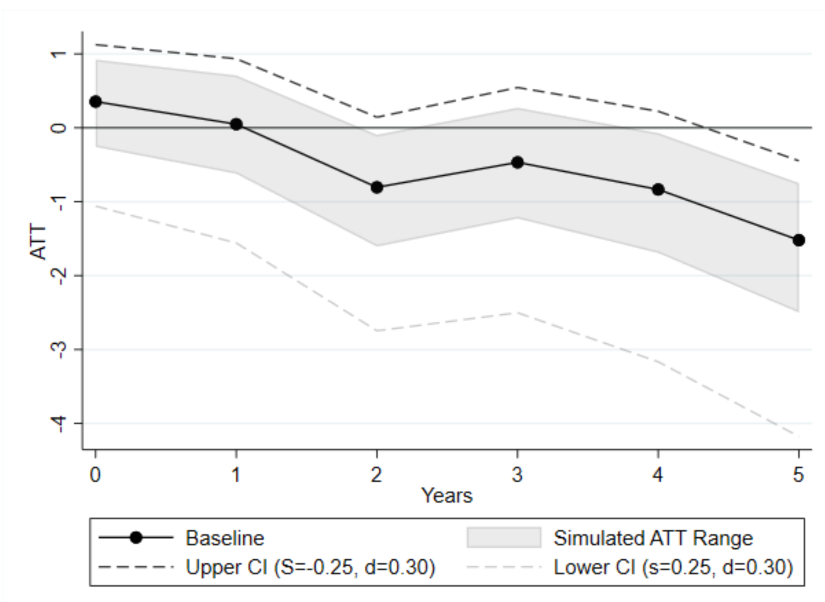
[19] The selection effects for physical health are analogous. The outcome effects differ in an expected manner with the largest effects occurring among own health outcomes and age-related covariates, representing the fact that physical health is more age-dependent than mental health. Both the maximum outcome and selection effects depict a similar range, hence we use the same simulated effects for both outcomes.

Figure A2.4: Simulated Violation of the CIA

a) Mental Health




b) Physical Health

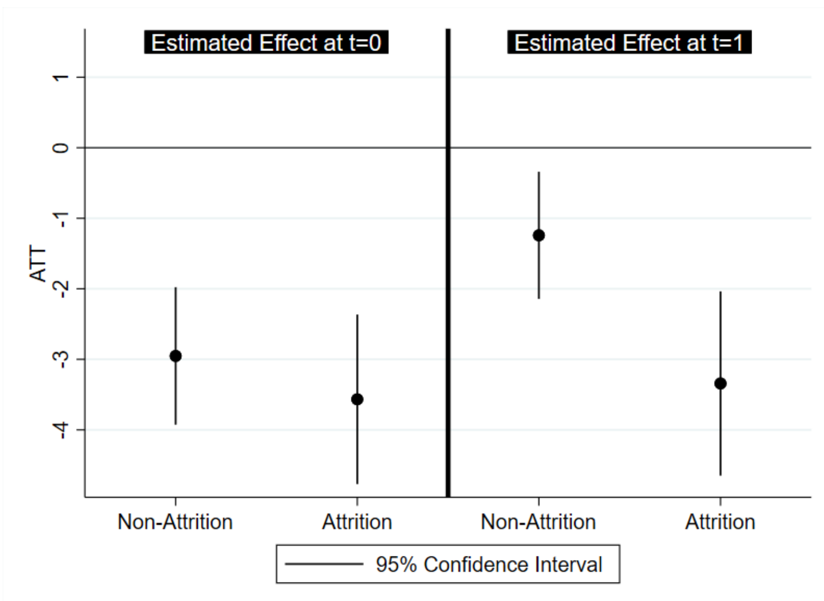



The results show that even when simulating a hypothetical confounder that combines the largest selection and outcome effects observed among the rich pool of observed variables already used in our matching procedure the immediate treatment effects would not vanish completely. These results provide additional confidence that the estimated negative mental health effects are not simply a chance finding or attributable to a conceivable violation of the CIA. It is crucial to note that this sensitivity analysis has an important drawback as it does operate under the assumption that the modelled effects are time-invariant. However, given that there are other concerns about the treatment effects in the longer-term, such as selective attrition driven by caregiving burden, we argue that to test our main finding this simulation is still insightful.

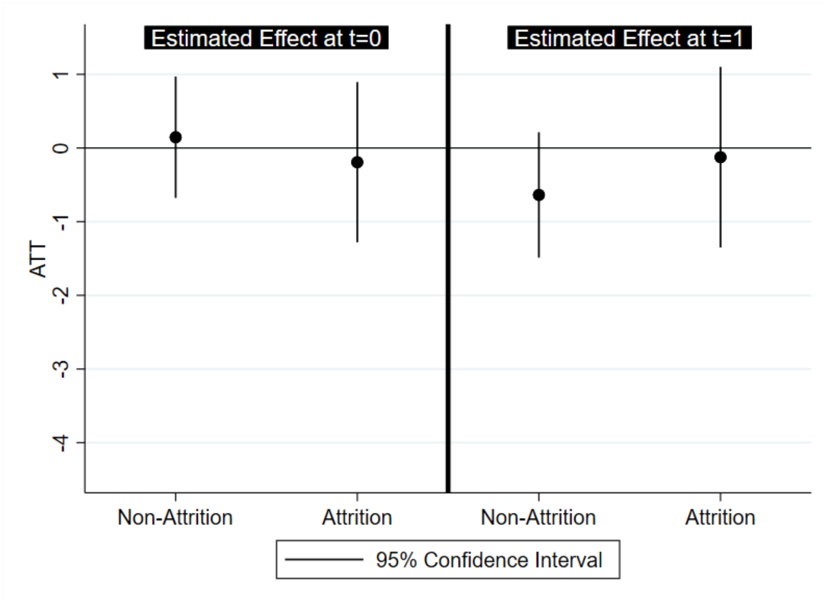
*Selective Attrition*

Figure A2.5: Initial Treatment Effect (High Intensity) by Attrition

a) Mental Health



b) Physical Health




*Alternative Outcomes Measures*

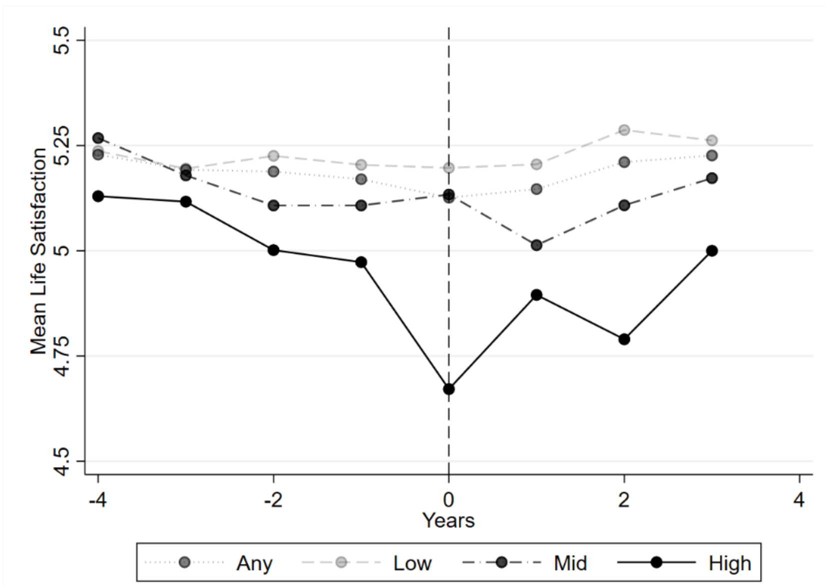
We explore the existence and magnitude of treatment effects on mental health using two alternatives but closely related outcome measures to assess the robustness of our results and their economic relevance. Firstly, we explore the effect of providing informal care on individual's life satisfaction. Life satisfaction is measured in the USoc by asking individuals about their overall level of satisfaction with their life in general and offering a discrete seven-point scale ranging from *"completely dissatisfied"* (1) to *"completely satisfied"* (7).[20] The second alternative outcome measure is the General Health Questionnaire (GHQ). The GHQ is a screening device to identify individuals at high risk of developing a non-psychotic minor psychiatric disorder such as anxiety or depression in general population surveys or outside of clinical environments (Goldberg et al., 1997).[21] Respondents' answers to the 12 items are transformed to a single score on a 0 (best) to 12 (worst) scale. To ease the visual interpretation of results in line with the other outcomes measures we have inverted the scale to range from 12 (best) to 0 (worst) so that negative coefficients indicate

[20] For simplicity, we treat life satisfaction as a continuous variable in order to be able to conduct the analysis without having to switch from our regular matching adjusted regression framework to an ordered response model.
[21] Goldberg, D. P., Gater, R., Sartorius, N., Ustun, T. B., Piccinelli, M., Gureje, O., & Rutter, C. (1997). The validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Psychological Medicine*, *27*(1), 191-197.

negative mental health effects. Figure A2.5 plots mean life satisfaction and inversed GHQ scores prior and after informal care onset. Importantly, the mental health focused GHQ scores do not depict strong negative trends prior to informal care provision for high-intensity caregivers, however with the start of caregiving scores among this group fall drastically.

Figure A2.6: Mean Life Satisfaction and Inversed GHQ Scores

a) Life Satisfaction



b) Inversed GHQ Scores

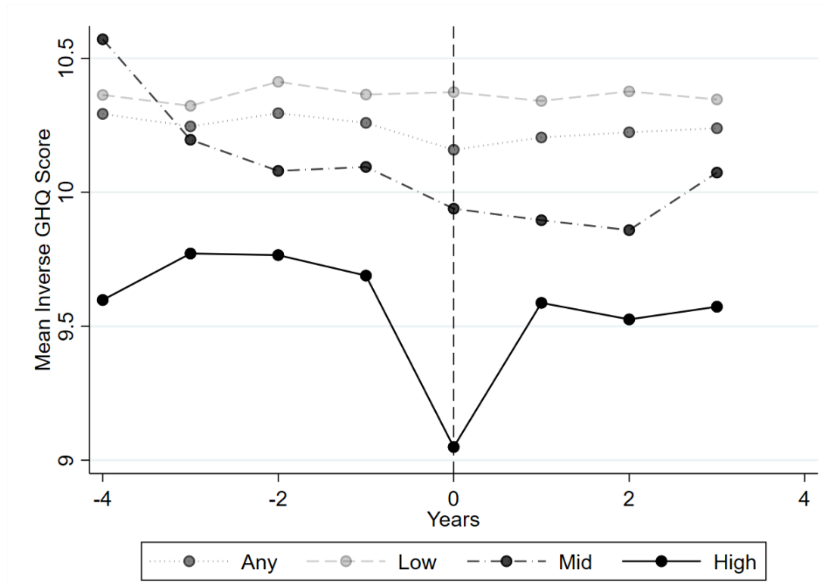
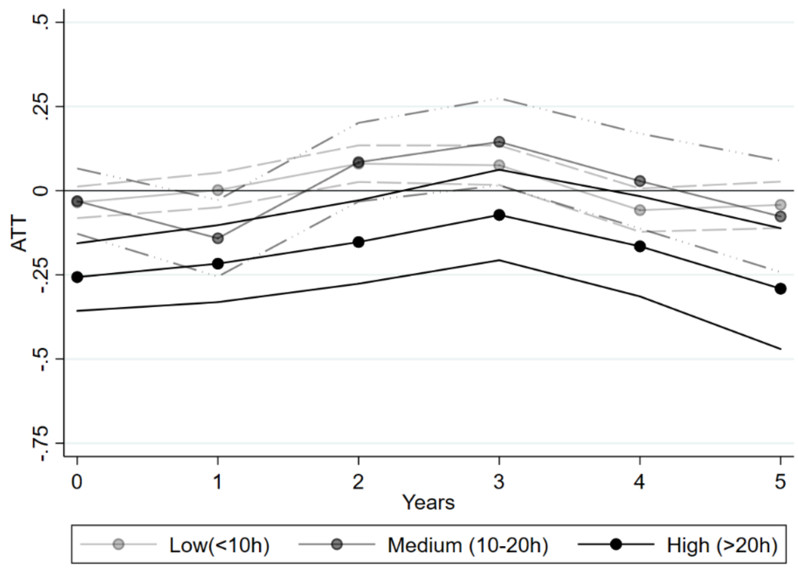


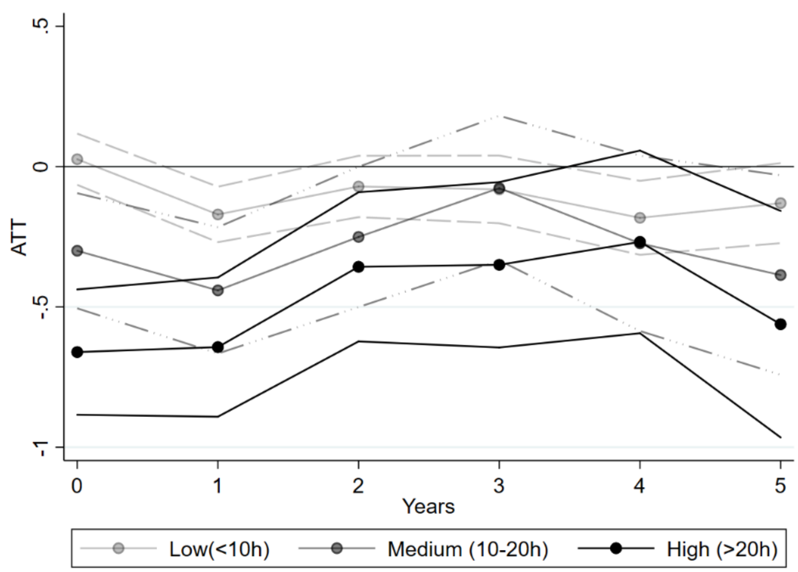

Figure A2.7 panel a) depicts the estimated treatment effects across the care-intensity levels. The overall structure of results across time and care intensity is generally in line with the overall results based on the SF-12 MCS. Figure A2.6 panel b) depicts the caregiving effects across intensity levels on reported GHQ scores. Interestingly our results using the GHQ as an outcome measure indicate a clearer dose-response relationship between informal care intensity in hours per week and mental health effects than our results using the MCS to measure mental health. For low intensity caregiving the coefficient for $t_0$ is zero but all subsequent coefficients are negative with those for $t_2$ and $t_4$ being significant with -0.170 ($p<0.001$) and -0.184 ($p<0.01$). The mental health effects increase with higher intensity. Among medium intensity caregivers the estimated initial effects are -0.301 ($p<0.01$) and -0.441 ($p<0.001$) at $t_0$ and $t_1$ while remaining similar in size and negative for later periods. The strongest and most persistent effects are found for high intensity caregivers with strong initial negative health effects of -0.663 ($p<0.001$) and -0.643 ($p<0.001$) at $t_0$ and $t_1$ that decrease to -0.355 ($p<0.01$) by $t_2$ but remain negative and persistent. This previously non-existent dose-response relationship might be explained by the fact that the GHQ is a more precise mental health measure as its 12 questions are solely focused on the mental and not the physical domain.

Figure A2.7: Treatment Effects by Intensity – Life Satisfaction and GHQ

a) Life Satisfaction




b) Inversed GHQ Score




To assess the practical implications of these results we convert the GHQ scores into a 'caseness' dummy. Following this definition, individuals scoring a 4 and above are identified as being a "case", meaning that these individuals experience high mental strain and are at risk of developing a mental disease. This definition does not indicate the definitive presence of a minor psychiatric disease, however, individuals scoring a 4 and above on the GHQ survey in a primary care environment should be referred to a mental health specialist for further investigation due to concern for their long-term mental health (Jackson, 2007)[22].

Across non-caregivers 16.31% of individuals cross this threshold at the pre-treatment period while the share is identical among low-intensity caregivers. Among medium-intensity caregivers the share is already considerably higher before care-provision (21.97%) and even more so for the high-intensity group (26.15%). In the period of first informal care provision these shares increase with caregiving intensity; for low intensity caregivers to 17.10%, for medium intensity to 23.99% and to 30.50% for high intensity caregivers.

Table 4.2 depicts the dynamic results for the alternative outcome measures. In accordance to the results using the MCS as outcome measure, we find negative effects of care provision on the GHQ among female caregivers for the second year of care. These negative effects remain present the third year as well but are not significant anymore. We do not find any effect on life satisfaction.

Table A2.5. Estimated effects of care trajectories using alternative mental health outcomes

| **Duration:** | **1 vs. 0 ($t_0$)** | | **2 vs. 0 ($t_1$)** | | **3 vs. 0 ($t_2$)** | | **2 vs. 1 ($t_1$)** | | **3 vs. 2 ($t_2$)** | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **LS** | **GHQ** | **LS** | **GHQ** | **LS** | **GHQ** | **LS** | **GHQ** | **LS** | **GHQ** |
| **All** | 0.02 | -0.04 | -0.08 | **-0.38**** | 0.06 | -0.02 | -0.04 | -0.11 | -0.05 | 0.26 |
| | (0.04) | (0.08) | (0.06) | (0.12) | (0.07) | (0.13) | (0.09) | (0.16) | (0.13) | (0.29) |
| **Male** | 0.07 | 0.17 | -0.05 | -0.18 | 0.11 | **0.36*** | -0.02 | 0.18 | -0.12 | 0.11 |
| | (0.06) | (0.10) | (0.10) | (0.17) | (0.09) | (0.15) | (0.13) | (0.24) | (0.19) | (0.25) |
| **Female** | -0.01 | -0.16 | -0.13 | **-0.55**** | -0.04 | -0.29 | -0.17 | -0.39 | -0.23 | 0.03 |
| | (0.05) | (0.11) | (0.07) | (0.15) | (0.09) | (0.18) | (0.11) | (0.21) | (0.16) | (0.41) |
| **Conrol** | | 14,777 | | 13,748 | | 12,865 | | 606 | | 185 |
| **Treatment** | | 1,208 | | 593 | | 348 | | 593 | | 348 |

* p < 0.05, ** p < 0.01, *** p < 0.001. Standard error in parentheses.

[22] Jackson, C. (2007). The general health questionnaire. *Occupational Medicine*, *57*(1), 79-79.

*Robustness checks dynamic sample*

Table A2.6 – Robustness check: Regression adjustment at all nodes

| Duration: | 1 vs. 0 ($t_0$) | | 2 vs. 0 ($t_1$) | | 3 vs. 0 ($t_2$) | | 2 vs. 1 ($t_1$) | | 3 vs. 2 ($t_2$) | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** |
| **All** | -0.10 | -0.05 | **-0.90**** | 0.14 | -0.66 | 0.16 | -0.28 | 0.40 | -0.88 | 0.72 |
| | (0.24) | (0.22) | (0.31) | (0.29) | (0.42) | (0.38) | (0.45) | (0.44) | (0.72) | (0.58) |
| **Male** | 0.04 | -0.16 | -0.60 | 0.18 | -0.59 | 0.60 | 0.21 | 0.16 | **-2.28*** | 1.13 |
| | (0.33) | (0.33) | (0.43) | (0.40) | (0.59) | (0.47) | (0.69) | (0.60) | (1.10) | (0.87) |
| **Female** | -0.15 | 0.09 | **-1.18**** | -0.02 | -0.66 | -0.20 | -0.57 | 0.40 | -0.01 | -0.96 |
| | (0.32) | (0.28) | (0.39) | (0.40) | (0.49) | (0.50) | (0.61) | (0.63) | (0.94) | (0.79) |
| **Conrol** | | 14,777 | | 13,748 | | 12,865 | | 606 | | 185 |
| **Treatment** | | 1,208 | | 593 | | 348 | | 593 | | 348 |

* p < 0.05, ** p < 0.01, *** p < 0.001. Standard error in parentheses.

Table A2.7 - Robustness check: Limiting covariates in propensity score estimations

| Duration: | 1 vs. 0 ($t_0$) | | 2 vs. 0 ($t_1$) | | 3 vs. 0 ($t_2$) | | 2 vs. 1 ($t_1$) | | 3 vs. 2 ($t_2$) | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** | **MCS** | **PCS** |
| **All** | -0.10 | -0.05 | **-0.85**** | 0.11 | -0.53 | 0.21 | -0.31 | 0.48 | -0.30 | 0.98 |
| | (0.24) | (0.22) | (0.32) | (0.31) | (0.44) | (0.40) | (0.46) | (0.46) | (0.78) | (0.66) |
| **Male** | 0.04 | -0.16 | -0.54 | 0.07 | -0.32 | 0.51 | -0.08 | 0.25 | -1.59 | 0.88 |
| | (0.33) | (0.33) | (0.45) | (0.41) | (0.63) | (0.51) | (0.67) | (0.63) | (1.10) | (0.93) |
| **Female** | -0.15 | 0.09 | **-1.19**** | 0.11 | -0.78 | 0.29 | -0.68 | 0.64 | 0.36 | -0.33 |
| | (0.32) | (0.28) | (0.40) | (0.41) | (0.52) | (0.52) | (0.61) | (0.63) | (0.99) | (0.84) |
| **Conrol** | | 14,777 | | 13,748 | | 12,865 | | 606 | | 185 |
| **Treatment** | | 1,208 | | 593 | | 348 | | 593 | | 348 |

* p < 0.05, ** p < 0.01, *** p < 0.001. Standard error in parentheses.

# A3 – Static Matching

Table A3.1: Propensity Score Estimation Results

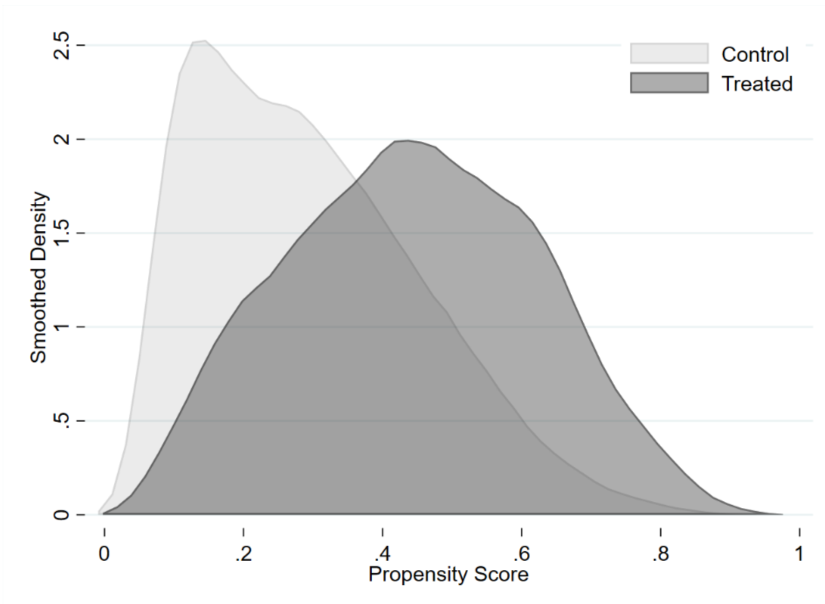
| | Coefficient | Standard Error |
|---|---|---|
| **Care Obligations** | | |
| Mother alive | -0.121 | (0.166) |
| Age of mother | **0.010*** ** | (0.002) |
| Father alive | -0.222 | (0.162) |
| Age of father | **0.010*** ** | (0.002) |
| Both parents alive | **-0.734*** ** | (0.062) |
| Living siblings | **-0.110*** ** | (0.031) |
| Living partner | **-0.482*** ** | (0.109) |
| Age of partner | **0.016*** ** | (0.001) |
| | | |
| **Willingness to Care** | | |
| Age | **0.013*** ** | (0.002) |
| Female | **0.121*** ** | (0.023) |
| Education: Tertiary (*Ref: Secondary*) | **-0.160*** ** | (0.023) |
| Education: Primary/Other lower | **-0.192*** ** | (0.029) |
| Job Status:Self-Employed (*Ref: Employed*) | **0.086* ** | (0.038) |
| Job Status: Unemployed | 0.069 | (0.054) |
| Job Status: Retired | 0.077 | (0.045) |
| Job Status: Homecarer | **0.168** ** | (0.050) |
| Job Status: Disabled | 0.131 | (0.073) |
| Job Status: Student/Other | **-0.174** ** | (0.058) |
| Working Full-Time | **-0.183*** ** | (0.031) |
| Income (logarithmic) | **0.111*** ** | (0.020) |
| HH Income Fraction | 0.071 | (0.042) |
| Marital Status: Partnership (*Ref: Married/Widowed*) | 0.028 | (0.033) |
| Marital Status: Single | **0.321*** ** | (0.069) |
| Marital Status: Sperated/Divorced | **0.234*** ** | (0.062) |
| Children in Household | -0.015 | (0.054) |
| Children < 14 in Household | 0.042 | (0.055) |
| Region: North-East (*Ref: London*) | **0.181** ** | (0.059) |
| Region: North-West | 0.014 | (0.044) |
| Region: Yorkshire | 0.027 | (0.049) |
| Region: East-Midlands | **0.145** ** | (0.048) |
| Region: West-Midlands | **0.153** ** | (0.047) |
| Region: East England | 0.061 | (0.046) |
| Region: South-East | -0.010 | (0.042) |

| | | |
|---|---|---|
| Region: South-West | **0.112*** | (0.046) |
| Region: Wales | **0.164**** | (0.048) |
| Region: Scotland | 0.046 | (0.047) |
| Region: Northern Ireland | **0.279***** | (0.055) |
| Living in Urban Area | 0.037 | (0.023) |
| Big-5: Openness | 0.011 | (0.008) |
| Big 5: Conscientiousness | **0.026**** | (0.010) |
| Big 5: Extroversion | **0.026**** | (0.008) |
| Big 5: Agreeableness | **0.029**** | (0.010) |
| Big 5: Neuroticism | -0.004 | (0.008) |
| **Ability to Care** | | |
| Self-Assessed Health | **-0.083***** | (0.014) |
| SF12 - Mental Score | **-0.019***** | (0.002) |
| SF12 - Physical Score | **-0.013***** | (0.002) |
| Chronic-Illness/Disability | **0.057*** | (0.026) |
| Number of Functional Limitations | **-0.128***** | (0.012) |
| Satisfaction with Health | **-0.063***** | (0.008) |
| Satisfaction with Income | **-0.024**** | (0.008) |
| Satisfaction with Leisure Time | -0.005 | (0.008) |
| Satisfaction with Life | 0.003 | (0.010) |
| GHQ Score | **0.023***** | (0.005) |
| **Observations** | | **19,822** |
| **Log-Likelihood** | | **-11,240.66** |
| **Chi^2** | | **2,676.72** |
| **Prob > Chi^2** | | **0.000** |
| **Pseudo R^2** | | **0.1204** |

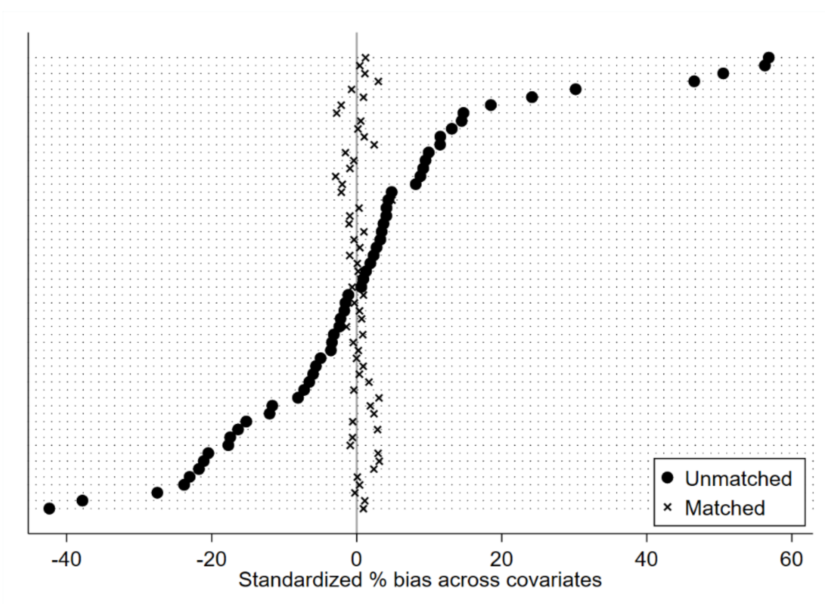
*p<0.05, **p<0.01, ***p<0.001, standard errors in parentheses

Figure A3.1: Propensity Score Distribution and Bias Reduction – Static Matching

a) Propensity Score Distribution



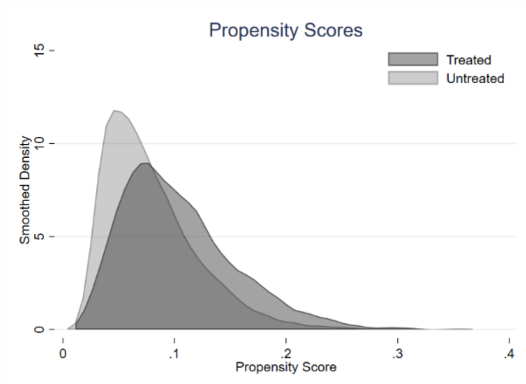

b) Standardized Bias Reduction after Matching



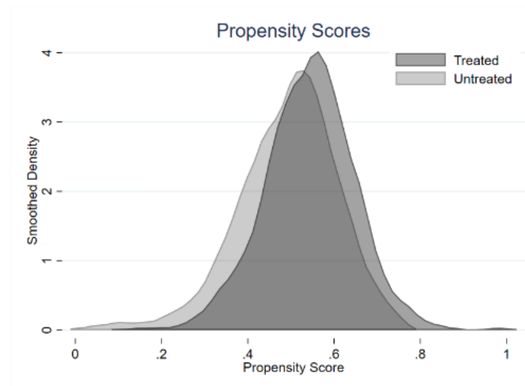
# A4 – Dynamic Sequential Matching

Figure A4.1: Estimated Propensity Scores

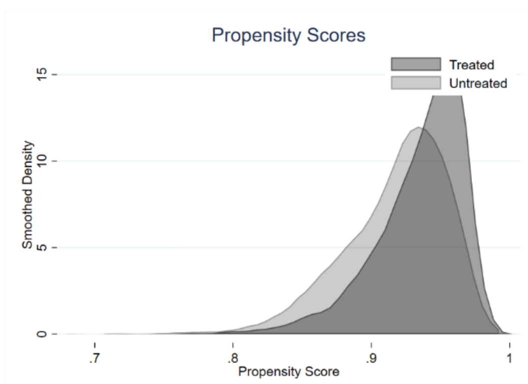
a) Propensity to start IC (Pr $(D_0=1|X_{-1}, Y_{-1})$)



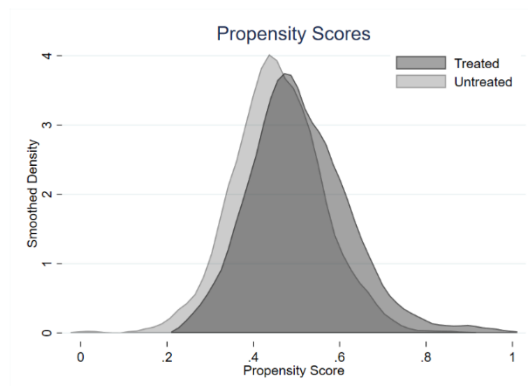

b) Propensity of IC after IC in year 1 (Pr$(D_1=1 \mid D_0=1, X_{-1}, Y_{-1}, X_0, Y_0)$)




c) Propensity of no IC after no IC in year 1 (Pr$(D_1=0 \mid D_0=0, X_{-1}, Y_{-1}, X_0, Y_0)$)

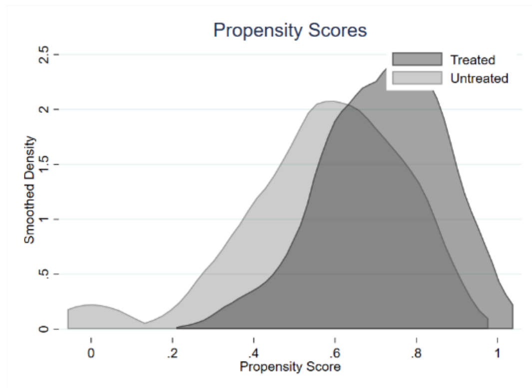


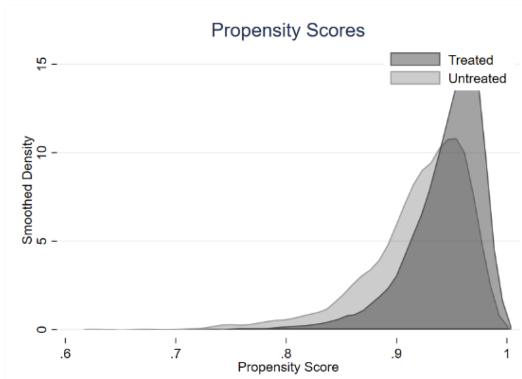

d) Propensity of no IC after IC in year 1 (Pr$(D_1=0 \mid D_0=1, X_{-1}, Y_{-1}, X_0, Y_0)$)



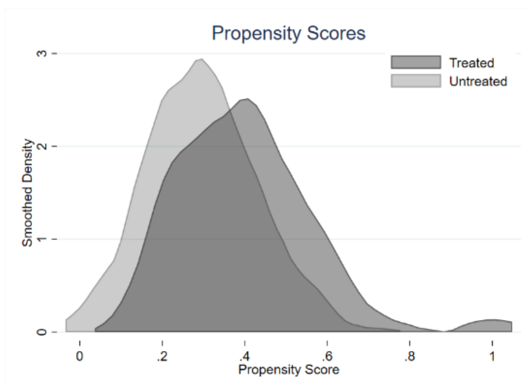

e) Propensity of IC after 2 years of IC (Pr$(D_2=1 \mid D_1=1, D_0=1, X_{-1}, Y_{-1}, X_0, Y_0, X_1, Y_1)$)

f) Propensity of no IC after 2 years of no IC (Pr$(D_2=0 \mid D_1=0, D_0=0, X_{-1}, Y_{-1}, X_0, Y_0, X_1, Y_1)$)




g) Propensity of no IC after 2 years of IC ($Pr(D_2=0 \mid D_1=1, D_0=1, X_{-1}, Y_{-1}, X_0, Y_0, X_1, Y_1)$




h) Propensity of no IC after IC in t=0 and no IC in t=1 ($Pr(D_2=0 \mid D_1=0, D_0=1, X_{-1}, Y_{-1}, X_0, Y_0, X_1, Y_1)$

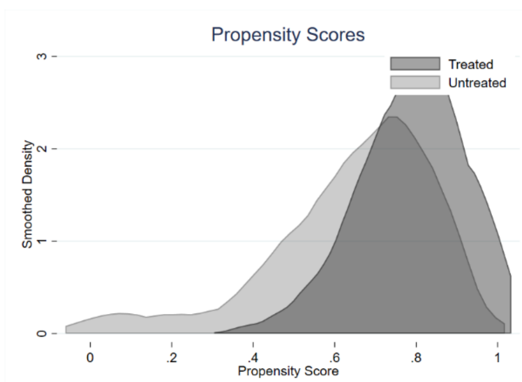



Table A4.1: Overview excluded propensity scores

| **Informal Care Trajectory** | **Off support** |
|---|---|
| Initial informal care ($Pr(D_0=1 \mid X_{-1}, Y_{-1})$) | 2 |
| 2nd year of care provision ($Pr(D_1=1 \mid D_0=1, X_{-1}, Y_{-1}, X_0, Y_0)$ | 5 |
| 2nd years of no care provision ($Pr(D_1=0 \mid D_0=0, X_{-1}, Y_{-1}, X_0, Y_0)$ | 86 |
| Informal care at $t_0$ but not $t_1$ ($Pr(D_1=0 \mid D_0=1, X_{-1}, Y_{-1}, X_0, Y_0)$ | 7 |
| 3rd year of care provision ($Pr(D_2=1 \mid D_1=1, D_0=1, X_{-1}, Y_{-1}, X_0, Y_0, X_1, Y_1)$ | 49 |
| 3rd year of no care provision ($Pr(D_2=0 \mid D_1=0, D_0=0, X_{-1}, Y_{-1}, X_0, Y_0, X_1, Y_1)$ | 162 |
| Informal care at $t_0$ and $t_1$ but not $t_2$ ($Pr(D_2=0 \mid D_1=1, D_0=1, X_{-1}, Y_{-1}, X_0, Y_0, X_1, Y_1)$ | 117 |
| Informal care at $t_0$ but not $t_1$ and $t_2$ ($Pr(D_2=0 \mid D_1=0, D_0=1, X_{-1}, Y_{-1}, X_0, Y_0, X_1, Y_1)$ | 26 |

Additionally, 1774 propensity scores at the first node are excluded as a result of being out of range (<5% or >95%)