TI 2018-050/V
Tinbergen Institute Discussion Paper

Health effects of caring for and about parents and spouses

Judith Bom[1]
Pieter Bakx[1]
Erik Schut[1]
Eddy (E.K.A.) van Doorslaer[1]

[1] Erasmus School of Health Policy and Management, Erasmus University Rotterdam

Health effects of caring for and about parents and spouses

Judith Bom[a], Pieter Bakx[a], Frederik Schut[a] and Eddy van Doorslaer[a,b,c]

a Erasmus School of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands
b Erasmus School of Economics, Erasmus University Rotterdam, Rotterdam, The Netherlands
c Tinbergen Institute, Amsterdam, The Netherlands

Abstract

Informal caregiving is a potentially attractive alternative to formal care but may entail health costs for the caregiver. We examine the mental and physical health impact of providing informal care and disentangle the caregiving effect – the effect of caring *for* someone in need – from the family effect – the effect of caring *about* someone in need. We account for potential endogeneity in the caregiving decision and control for previous health status using Arellano-Bond difference GMM models. We use four waves (2010-2013) of panel data from the *Dutch Study on Transitions in Employment, Ability and Motivation (STREAM)*.

We find that caregiving harms the mental health of caregivers; this effect is mainly present for spousal caregivers. A negative health shock of a family member also has a direct negative effect on mental health, providing evidence of a family effect. These findings imply that most studies may have overestimated the negative health effects of caregiving by not accounting adequately for the family effect. As the caregiving effect differs strongly between various types of caregivers, policies to counteract this effect should specifically target subgroups of caregivers that carry the largest burden of informal caregiving.

Keywords: long-term care; informal care; caregiver effect; family effect; mental health
JEL classification: J14; I10; J18

1. Introduction

In most Western countries, the demand for long-term care (LTC) is expected to keep rising in the decades to come. In the Netherlands the demand for LTC is projected to grow at an average annual rate of 1,6% between 2014 and 2030 (Eggink et al., 2016). Part of this growing demand is likely to be met by informal care, i.e. by unpaid care provided by relatives and friends. Such care is often preferred over formal care by both the care receiver and the government. Enhancing reliance on social networks for providing informal care is even an explicit policy goal of the recent Dutch Social Support Act (2015).

Potentially because of this enhanced stimulation of informal care provision, the number of caregivers has risen throughout the past years. In 2016, 23% of the Dutch adult population was estimated to provide informal care, equaling more than 3 million persons. Two-thirds of this group is female. Caregiving duties are furthermore concentrated among the older inhabitants, 60% of the caregivers is aged 45-65 another 20% is aged over 65 (Gezondheidsmonitor, 2016).

While the costs of informal caregiving are typically low for the recipient, they may be substantial for the caregiver and society. The stress and physical strain involved in informal caregiving risks hurting the health of caregivers (e.g. Pinquart and Sörensen, 2003). Yet, a decline in caregivers' health might not only be the result of informal care provision, it could also be driven by experiencing illness of a family member (Amirkhanyan and Wolf, 2006; Bobinac et al., 2010). This paper aims to improve our understanding of the health effects experienced by informal caregivers by disentangling both effects.

2. Earlier work on health effects of informal caregiving

The health effects of informal caregiving have been studied quite extensively, with most evidence demonstrating that caregivers experience worse physical and mental health compared to non-caregivers (e.g. Pinquart and Sörensen, 2003; Pinquart and Sörensen, 2007; Vitaliano et al., 2003). Many of these studies, however, used non-representative samples or focused on the provision of care for a specific disease (Hirst, 2004). While the recent literature has moved towards estimating the impact of informal caregiving using larger, representative datasets, two main challenges of adequate empirical identification of these effects have emerged.

A first challenge relates to the potential endogeneity between the decision to provide informal care and one's own health. Persons with lower health expectations or a lower propensity to work might be more inclined to take up the caregiver role for their parents (Schulz, 1990). In addition, omitted-variables may lead to bias when unobserved variables like personality affect both the propensity of providing care and the health of the caregiver.

A second challenge, which has received limited attention thus far, is dealing with the notion that two distinct effects may be present in situations of informal caregiving: the family effect and the caregiving effect. The

family effect refers to the impact of *caring about* a person and is different from the caregiving effect of *caring for* a person in need. Irrespective of care provision, experiencing a health decline of a loved one can have a negative effect on one's own health or well-being (Amirkhanyan and Wolf, 2006; Bobinac et al., 2010).

It is important to make a distinction between the caregiving effect and the family effect, as addressing them adequately would require different government interventions. Respite care, for instance, could alleviate the caregiving effect, but offers no solution to the family effect of worrying about an ill family member. As the caregiving and the family effect often occur simultaneously, it is difficult to disentangle them. Yet, not controlling for the family effect might lead to overestimation of the caregiving effect. The existence of a family effect also means that a health shock of a family member is not a valid instrumental variable (IV) to estimate the causal effect of informal caregiving on health. In the presence of the family effect, this IV violates the exclusion restriction, which requires that the instrument is uncorrelated with the error term. In other words, using health of a family member as an IV for informal caregiving could overestimate the effect of caregiving on health by attributing the entire difference in health between individuals with and without ill family members to caregiving.

A number of studies have attempted to address endogeneity problems when evaluating the health effects of informal caregiving by using statistical matching, fixed effects (FE) methods or IV approaches (see table 1 for an overview). Three studies (Brenna and Di Novi, 2016; De Zwart et al., 2017; Schmitz and Westphal, 2015) addressed endogeneity of caregiving by statistically matching caregivers and non-caregivers on observable characteristics. All three papers found negative effects of caregiving on mental health but the method might not guarantee matching on unobservables.

Van den Berg et al. (2014) used FE models on Australian caregivers. They estimated significant negative effects of caregiving on subjective wellbeing. Using FE they avoided all time-constant heterogeneity, however they did not consider selection into caregiving based on time-variant elements such as previous health.

The last group of studies used IVs to estimate the impact of informal care provision. All found significant negative effects of informal caregiving on the caregiver's health. Two out of three studies made use of an instrumental variable relating to death of one's parents. Coe and Van Houtven (2009) used death of one's mother as IV for selection out of caregiving in US data and control for lagged health. Heger (2016) used the transition from having two parents to only one parent as IV for caregiving. The risk of these instruments lies in their potential violation of the exclusion restriction.

Table 1: Overview of literature regarding health effects of informal caregiving accounting for endogeneity of caregiving

Authors	Data source	Care recipient	Sample (sample size)	Health measure	Focus on family effect	Estimator	Instrumental variable	Results
Coe and Van Houtven (2009)	HRS, 1992-2004 (7 waves)	Parent	Males and females aged 50 to 64, having only a mother alive (1,467 current caregivers) Distinction married/single	Mental health (CES-D8); Physical health (self-reported health on a 5-point, diagnosed heart condition & blood pressure).	-	IV, Arellano-Bond	(1) Death of a parent (2) Sibling characteristics	(1) Continued caregiving: ↑ CES-D-8 (married males and females) ↑ Heart condition (single males) likelihood of reporting excellent or very good health (↓ married females, ↑ married males) (2) Initial caregiving: ↑ CES-D-8 (married females)
Do et al. (2015)	Korean LSA, 2006-2010 (3 waves)	Parent (in-law)	Women with living parent(in law), aged 45+, (2,528 daughters-in-law & 4,108 daughters)	Pain affecting daily activities; Fair or poor self-rated health; Any outpatient care use; OOP spending for outpatient care; Any prescription drug use; OOP spending prescription drugs.	Claim to avoid family effect by focusing on physical health and care for parents-in-law	IV-2SLS, IV-probit	ADL limitations of the mother(-in-law) and father(-in-law)	↑ Pain affecting daily activities, Health Self-rated as poor, OOP outpatient care (daughters & daughters-in-law) ↑ Any outpatient care use, Any prescription drug use (daughters)
Heger (2016)	SHARE, 2004-2013 (4 waves)	Parent	Women and men aged 50-70 (3,669 women & 2,752 men)	EURO-D depression scale, indicator whether someone suffers from 4 or more depressive symptoms.	Estimate family effect by adding health of parent as variable to model	IV-FE	Indicator of whether one parent is alive	↑ Euro-D, 4+ depressive symptoms (females) Family effect small
Van den Berg et al. (2014)	HILDA, 2001-2011 (11 waves)	Partner, parent (in-law) or relative	Aged 16+ (23,285 individuals)	Individual subjective wellbeing	-	FE- ordered logit	NA	↓ Well-being
Schmitz and Westphal (2015)	GSOEP, 2002-2010	Unknown	Women aged 18+, (31,177 person-year observations at t = 0)	SF-12v2 MCS & PCS	-	Matched regression	NA	Short term: ↓ SF12 Mental Component Scale
Brenna and Di Novi (2016)	SHARE, 2004-2007 (2 waves)	Parent	Women aged 50-75 (N=3,936)	Euro-D depression scale	-	Matched regression	NA	↑ Euro-D (females, Southern European countries)
De Zwart et al. (2017)	SHARE, 2004, 2006, 2010, 2013	Partner	Males and females aged 50+ (N=10,472)	Prescription drugs usage; number of doctor visits in the past 12 months; EURO-D depression scale; self-perceived health.	-	Matched regression	NA	Short term: ↑ Euro-D, ↓ self-reported health; ↑ prescription drug use (females), ↑ doctor visits (females)

The IV employed by Coe and Van Houtven could namely directly affect adult children's health via grief or relief from worries. One of their robustness tests confirms that mother's death in some cases positively affects the mental health of non-caregiving daughters, potentially because these daughters are relieved from the family effect. As a result, the negative effect of continued caregiving may be overestimated.

A similar risk is present for the instrument employed by Heger (2016): the indicator of whether exactly one parent is alive. Although she controlled for the family effect via an indicator of poor health of a family member, and for grief via indicators of whether each of the parents are still alive, the family effect may still cause an overestimation of the caregiving effect. The loss of one parent while the other is still alive is likely to cause grief itself. Furthermore, the instrument used (the transition from having two parents to one) is unlikely to meet the monotonicity assumption, i.e. the instrument affects the treatment status in the same direction for all units. In this case, monotonicity would mean that the death of the first parent should lead to the provision of informal care or have no effect. There may be 'defiers', i.e. those who stop providing care in response to the instrument, for example because the deceased parent required informal care or because the now single parent moves to an elderly home.

Also Do et al. (2015) have considered the family effect. They used poor parental health as an IV to estimate the impact of caregiving on self-reported health and prescription drug use. The authors argue that they managed to avoid picking up the family effect by only focusing on (i) physical health effects and (ii) females providing care to their parents-in-law. It is however, unclear whether self-reported health and prescription drug use only measure physical and not mental health. Moreover, mental and physical health could affect each other, thereby making isolation of the family from the caregiving effect difficult as stress related to the family effect may induce physical health problems.

We contribute to the literature in two ways. First, we deal with both the endogeneity of caregiving and the family effect when estimating the health effect of informal caregiving. We account for potential endogeneity and specifically control for previous health status using Arellano-Bond difference GMM models.

Second, we estimate the impact of informal caregiving and the family effect for various groups of caregivers. While most papers typically focus on either parental or spousal caregiving, we examine whether the health effects of caregiving differ when caring for a spouse instead of a parent. There may be various reasons why the impact of caregiving differs between these care types. As spousal caregivers tend to be older, they often have fewer physical and psychological resources to deal with stress related to caregiving (Pinquart and Sörensen, 2003). Furthermore, differences in the levels of care provision are seen as large drivers of differences in stress between spousal and parental caregivers (Pinquart and Sorensen, 2011).

3. Methods

To deal with the potential endogeneity of caregiving and own health, we start from the economic intuition behind the caregiving decision as presented by De Zwart et al. (2017). According to their model, several elements affect the caregiving decision. The first relates to personal ability, both reflected in wage and health. Persons with a higher ability might prefer paid work to providing informal care. Availability of other types of informal and formal care might affect the caregiving decision as well. Household income and size could for example reflect someone's possibilities to purchase formal care or to transfer caregiving duties to family members. Lastly, non-monetary factors such as cultural factors might affect the caregiving decision. Schmitz and Westphal (2015) made a similar summary of influencing factors by referring to *care obligations*, *willingness to provide care* and *ability to provide care.*

We account for these factors that shape the caregiving decision in three ways. First, we control for both observed and unobserved time-invariant personal characteristics, like personality traits and education. We do so by taking differences.

Second, we deal with selection based on health by controlling for pre-treatment health status.[1] Including pre-treatment health status in an individual FE model is not possible as the lagged dependent variables correlate with the fixed effects in the error term and would give rise to dynamic panel bias. This bias could affect our estimates of the lagged dependent variable, as well as the coefficients of our other independent variables, especially when the dataset contains few waves but many observations (Nickell, 1981).

In order to control for time-invariant unobservables while including a lagged variable in our model, we make use of the Arellano-Bond (1991) (A-B) estimation technique. Replacing the fixed effects by first differences to rule out time-invariant elements does not yet solve the problem as H_{it-1} remains correlated with $\Delta\varepsilon_{it}$ via ε_{it-1}. However, by using first differences instead of fixed effects (e.g. mean-centered) transformations, deeper lags of health become orthogonal to the error term. In the A-B models, we use these deeper lags of health to instrument for the lagged health difference. We can use H_{it-2} as an instrument for ΔH_{it-1} because H_{it-2} is correlated with ΔH_{it-1} but not with $\Delta\varepsilon_{it}$ as long as the error-terms are not serially correlated. Following Arellano and Bond (1991), we do not only include the second lag of health, but all available deeper lags of health as instruments. As we have four waves of data, we can include the second and third lag.

Third, we control for remaining observed time-varying factors related to the caregiving decision and own health by including these as covariates into our models. Undoubtedly, the most prominent are the income and employment status, and the presence of other caregivers. Furthermore, we include measures

[1] Note that controlling for lagged health also helps to mitigate of unobserved characteristics (cf. Lechner, 2009).

to capture the family effect in the model. Accounting for these elements, we assume to control for all components potentially affecting both own health and the propensity of caregiving in the past year.

The resulting dynamic panel data model is estimated using a difference Generalized Method of Moments (GMM) regression through the Stata command *xtabond2* (Roodman, 2006).[2] The model is specified as follows:

$$\Delta H_{it} = \beta_1 \Delta H_{it-1} + \beta_2 \Delta IC_{it} + \beta_3 \Delta HF_{it} + \beta_4 \Delta X_{it} + \Delta \varepsilon_{it} \qquad (1)$$

ΔH_{it} refers to a change in health of respondent *i* at time *t.* This health outcome is dependent on the change in one's health status the previous year ΔH_{it-1}; on ΔIC_{it}, indicating the change of the informal care provision status; on ΔHF_{it}, referring to the a change in the health state of the individual's partner or close family member, as well as a vector of changes in individual time-varying characteristics, ΔX_{it}. All time-invariant individual characteristics are factored out by differencing. Our main parameters of interest are β_2, indicating the caregiving effect, and β_3, indicating the family effect. By including both variables into the model, we disentangle the caregiving effect from the family effect. As said ΔH_{it-1} is instrumented by H_{it-2} and H_{it-3}.

4. Data

We use the *Study on Transitions in Employment, Ability and Motivation* (STREAM) panel survey, which includes four annual waves of data ranging from 2010 to 2013. The objective of the survey was to collect information on determinants of transitions into and out of employment and of work productivity among persons aged 45-64 years. This is also the age group providing most informal care in the Netherlands (Gezondheidsmonitor, 2016). The STREAM sample is stratified at baseline based on age and work status and is drawn from an existing internet panel (Ybema et al., 2014). In the first wave, 15,118 persons responded to the survey. In later waves, this original sample was invited to participate again without replacement. Attrition is low: in total almost two-thirds (64%) of the sample responded to all four surveys.

The panel data are linked at the individual level to administrative data for all registered inhabitants in the Netherlands obtained via Statistics Netherlands.[3] These administrative registers include: (i) personal demographics including information about family structure from the municipal register; (ii) information on prescribed drugs covered by health insurance from the Dutch Healthcare Insurance Board (CvZ); (iii) health care expenses covered by the Dutch Health Insurance Act obtained from insurance claim data; (iv) household income information from the Dutch Tax Administration; and (v) data on eligibility for formal LTC assessed by the national needs assessment agency (CIZ).

[2] We present robust twostep estimates.
[3] We use non-public microdata which, when adhering to various conditions, can be accessed via a secured remote access connection.

Sample

From the 15,118 first wave respondents, we exclude 1,873 individuals as they (i) did not approve linkage to the register data; (ii) could not be identified in the register data; or (iii) completed the survey in the same wave twice.[4] We then select a subsample that is able to provide informal care to either their parent or partner. We therefore exclude all 2,363 respondents without a living parent or spouse at baseline. A respondent is considered to have a partner when he or she, according to the administrative records, is married or in a registered partnership. As the difference GMM regression requires at least three waves of data, we solely include individuals who responded at least three times to the survey. The inclusion criteria are listed in table 2. Our sample at baseline consists of 4,400 males and 3,528 females; across all waves we have 17,055 male and 13,693 female observations.

Table 2: Sample size and sample selection criteria

Inclusion criteria	Overall sample (% of total)
Total respondents at T_1 (2010)	15,118
Agreement to be linked to administrative data	13,672 (90.4)
Identified in administrative data	13,398 (88.6)
Did not submit survey twice in same wave	13,218 (87.4)
Having at least one parent alive and/or having a spouse at T_1	10,855 (71.8)
Fully completed ≥3 surveys	7,928 (52.6)
Total number of respondents included T_1	*7,928*

Health measurement

This panel dataset allows us to use four (complementary) self-reported health outcomes. The first two measures are derived from the SF-12 health survey, which contains 12 questions regarding physical and mental health during the past four weeks. From this survey, we derive two subscales: the Physical Component Scale (PCS) and the Mental Component Scale (MCS) (Ware et al., 1995).[5] Both scales range from 0 to 100, a higher score equals a better health status.

In addition, we use two measures that capture specific aspects of health. We expect caregiving strain to have particularly large effects on fatigue and depression. Informal caregiving often leads to caregiver fatigue because caregivers may prioritize the patient's needs over their own (Schulz et al., 1990). The SF-36 vitality subscale (0-100) assesses fatigue based on responses to four items[6], where a higher score relates to lower fatigue/higher vitality (Ware and Sherbourne, 1992). Although the vitality subscale was developed as part of a broader health measure, the subscale is used in isolation in various patient populations (Harel et al., 2012; Hewlett et al., 2011).

Finally, the stress involved in caregiving (Pinquart and Sörensen, 2003) as well as the stress involved in illness of family members (Amirkhanyan and Wolf, 2006) can lead to an increase in depressive

[4] We consider this sample a random subsample, no large discrepancies in observable characteristics between the subsample and total sample were detected.

[5]The Physical Component Scale consists of sub-scales: Physical functioning (2 questions), Role-Physical (2 questions), Bodily Pain and General Health. The Mental Component Scale consists of sub-scales: Vitality, Social Functioning, Role-Emotional (2 questions) and Mental Health (2 questions).

[6] The past four weeks: (1) Did you feel full of life? (2) Did you have a lot of energy? (3) Did you feel worn out? (4) Did you feel tired?

symptoms. To measure depression, we use the CES-D-10 scale, consisting of 10 questions (Radloff, 1977). Scores can range from 0-30, a higher score relates to increased presence of depressive symptoms. The health scores were reported as missing in case the respondent failed to answer any (MCS and PCS) or >2 questions (CES-D-10 and vitality scale).

Measurement of informal caregiving

The main treatment variable is a binary variable indicating whether someone provided informal care (IC) in the past year. Respondents were asked: 'Did you in the past 12 months spend part of your time on any of the following activities?' When they answer 'Giving Informal Care' affirmatively, they are considered informal caregivers.[7] In the last two waves of the survey, respondents were also asked to indicate *to whom* they provided care. To analyze differences in the type of care given, we distinguish spousal from parental caregiving in subgroup analyses. Based on the 2012 and 2013 observations, we impute the type of care in the first two waves assuming that the care recipient (parent or spouse) remains the same throughout the years. In table 3 an overview of the number of informal caregivers is given, specified by care recipient. As our sample is limited to respondents aged 45-65 we solely capture a part the caregiving-population, especially spousal caregivers tend to be older and hence underrepresented in our data. Our results might therefore underestimate the full effect, as older caregivers might be more prone to the negative health effects of caregiving (Pinquart and Sörensen, 2003).

Table 3: Number of informal caregivers, specified by care recipient

	Male				Female			
	2010	2011	2012	2013	2010	2011	2012	2013
Providing informal care to spouse	94 [a]	112 [a]	140	174	114 [a]	128 [a]	134	131
Providing informal care to close family member (parent)	238[a]	292[a]	394	433	625[a]	686[a]	804	732
Total informal caregivers	653	699	697	867	1,156	1,202	1,267	1,168

a: Imputed based on care recipient in 2012-2013

Measurement of the family effect

In order to control for and estimate the family effect, a variable indicating severe illness of a family member is included in the model. This variable is self-reported and indicates whether a spouse or close family member has become severely ill within the past year. Although illness of a family member and informal caregiving often occur simultaneously, this is not necessarily the case.

Covariates

As explained in section 3, we take first-differences and control for lagged health. To deal with any remaining (time-variant) characteristics that influence both the decision to provide care and the respondent's health, we include the following individual-level covariates: age, age-squared, self-reported financial difficulties[8], percentile group standardized household income, marital status, having

[7] Informal care (in Dutch: Mantelzorg) refers to providing non-professional care for a person in need in your own close environment, it does not include looking after healthy family members.

[8] This variable equals 1 when the respondent indicated that their household is currently very short or a bit short on money.

children living at home, employment status and whether or not the father or mother is alive. Finally, we include wave dummies to capture time trends affecting all respondents, including for instance any trends in formal LTC use.[9]

5. Results

Descriptive statistics

Table 4 presents descriptive statistics for our sample at baseline. The data is stratified based on whether the respondent ever provided informal care throughout the four waves of STREAM. In table A1 in the appendix the descriptive statistics for the sample stratified by gender and care-tasks can be found. Caregivers are more often female, and show to have a lower health status at baseline. Furthermore, they are less often employed, married or living with children at home. As expected, informal caregivers more often have ill family members.

Table 4: Descriptive statistics of sample at baseline

	Never informal caregiver		Ever informal caregiver		
	Mean	SD	Mean	SD	
Health outcomes					
SF-12 physical component scale	49.27	9.75	48.82	9.98	*
SF-12 mental component scale	52.64	8.26	51.24	9.46	***
SF-36 vitality scale	66.31	19.11	63.85	19.79	***
CES-D-10 depression index	5.02	4.82	5.69	5.30	***
Health family					
Severe illness spouse/family member	0.14	0.34	0.27	0.45	***
Severe illness spouse	0.02	0.15	0.07	0.25	***
Severe illness close family	0.12	0.32	0.22	0.41	***
Personal characteristics					
Age	53.74	5.52	53.94	5.20	
Age Squared	2,918.66	596.09	2,936.92	562.26	
Gender	0.35	0.48	0.57	0.49	***
Married/registered partnership	0.87	0.34	0.83	0.38	***
Children living at home	0.47	0.50	0.44	0.50	***
Employed	0.89	0.32	0.85	0.35	***
Perc. group household income	68.20	22.71	67.92	23.33	
Financial difficulties	0.19	0.39	0.20	0.40	*
Father alive	0.53	0.50	0.66	0.47	***
Mother alive	0.31	0.46	0.35	0.48	***
Number of observations in 2010	4,654		3,274		

SF-12 PCS and MCS and SF-36 Vitality range from 0-100 (lowest - highest level of health). CES-D-10 ranges from 0-30. a score ≥10 is considered a sign of depression. ***p < 0.01; **p < 0.05; *p < 0.1 indicate differences between never and ever informal care sample

Health effects

The estimation results of the A-B models presented in table 5 suggest that informal caregiving only has a negative effect on mental health as measured by the MCS. The effect is small compared to the mean MCS score (only about 1%). Compared to the average individual yearly change of 0.2 points on the MCS scale, the change in mental health seems however noteworthy.

[9] There are no differences across regions or between households in formal care availability that we need to account for. In general, co-payments are low and income-related and there are virtually no waiting lists for formal care use.

We also observe significant family effects on mental health: a severe illness occurring to a family member leads to a significant decrease in the mental health score of about the same size as the caregiving effect and to a significant increase in depressed feelings. Not many other covariates show significant contributions to health changes: a change in marital status negatively correlates with changes in depressed feelings, whereas a change in experiencing financial difficulties is associated with enhanced feelings of depression and drops in vitality scores.

Table 5: Arellano-Bond difference GMM regressions

A-B				
	SF-12 PCS	SF-12 MCS	SF-36 vitality	CES-D-10 depression
Informal care	-0.07 (0.19)	-0.45** (0.22)	-0.49 (0.35)	0.13 (0.10)
Severe illness spouse/family	0.14 (0.15)	-0.43*** (0.17)	-0.29 (0.27)	0.16** (0.08)
Lagged health	0.10*** (0.02)	0.11*** (0.02)	0.09*** (0.02)	0.08* (0.04)
Employed	0.46 (0.29)	-0.23 (0.31)	-1.74*** (0.54)	-0.00 (0.16)
Financial difficulties	-0.22 (0.21)	-0.18 (0.25)	-1.14*** (0.40)	0.32*** (0.12)
Perc. group household income	-0.01 (0.01)	0.01* (0.00)	0.01 (0.01)	-0.00 (0.00)
Age	0.43 (0.53)	-0.68 (0.54)	-0.20 (0.97)	0.30 (0.27)
Age Squared	-0.00 (0.00)	0.01 (0.00)	0.00 (0.00)	0.00 (0.00)
Married/Registered partnership	-0.58 (0.54)	1.18 (0.77)	1.26 (1.11)	-0.65* (0.35)
Children living at home	-0.07 (0.30)	-0.30 (0.33)	-0.48 (0.56)	0.14 (0.16)
Mother Alive	0.10 (0.39)	-0.98** (0.43)	-0.84 (0.70)	0.06 (0.21)
Father Alive	0.54 (0.40)	-0.21 (0.51)	-0.22 (0.74)	0.20 (0.22)
Hansen J-test (p-value)	0.50 (0.78)	0.96 (0.61)	3.2 (0.21)	2.93 (0.23)
Number of instruments	3	3	3	3
N	13,626	13,626	14,824	14,835

Robust standard errors are in parentheses. All regressions include wave dummies. ***p < 0.01; **p < 0.05; *p < 0.1.

Our A-B models rely on deeper lags of health to instrument the lagged first-difference in health. The relevance of these instruments is demonstrated in the first stage results (table 6): the lagged excluded levels of health are strongly correlated with the lagged difference in health.[10] This is confirmed by the Kleibergen-Paap Wald *rk F*-statistics for the excluded instruments, indicating that the instruments are not weak. The Hansen (1982) J-test statistics for overidentifying restrictions provide an indication of the validity of our instruments: for all models, we cannot reject the null-hypothesis that all instruments are valid, i.e. not correlated with the error term, under the assumption that at least one instrument is valid.[11]

[10] The instrument matrix is not collapsed, hence we have separate instruments for each time period.

[11] The difference GMM model furthermore relies on the assumption of no serial correlation among the errors, which can be assessed using the test proposed by Arellano and Bond (1991). This test focuses on finding autocorrelation among the differenced error terms. We however cannot test for second order autocorrelation, as it requires five waves of data. In the absence of any formal test for this assumption, we rely on Coe and Van Houtven (2009) and Roy and Schurer (2013) who did not find any second order serial correlation of the residuals for mental health. We also estimated our model using a deeper lag of health (H_{it-3}), which would solve the problem in case any second order correlation were present. Use of this deeper lag of health hardly affected our estimates, supporting the robustness of our assumption.

Table 6: First stage statistics of A-B estimations

1st Stage				
	Lag SF-12 PCS	Lag SF-12 MCS	Lag SF-36 vitality	Lag CES-D-10 depression
2.Lag Health indicator	-0.25*** (0.01)	-0.46*** (0.01)	-0.30*** (0.01)	-0.33*** (0.01)
2.Lag Health indicator	-0.34*** (0.01)	-0.48*** (0.02)	-0.47*** (0.01)	-0.52*** (0.01)
3.Lag Health indicator	0.13*** (0.01)	0.11*** (0.01)	0.31*** (0.01)	0.32*** (0.02)
Kleibergen-Paap Wald *rk F*-statistic	446.29	726.89	755.6	622.3
N	13,626	13,626	14,824	14,835

Robust standard errors are in parentheses. These regressions also include: informal care, severe illness of spouse/family member, age, age^2, financial difficulties, children at home, married, employed, financial difficulties, standardized household income, mother alive, father alive, wave dummies. ***p < 0.01; **p < 0.05; *p < 0.1.

Heterogeneous effects: subsample analyses

In order to assess whether the caregiving effect and the family effect differ for various types of caregiving, we carry out various subsample analyses. First, we analyze whether the provision of more informal care also leads to larger mental health damage. There indeed appears to be a dose-response relationship: for individuals that start providing at least 8 hours of care per week (31% of the caregivers provide this amount of care), the impact of informal care on mental health and vitality nearly doubles (table 7). The caregiving effect is hence larger for individuals that provide more hours of informal care per week.

Table 7: AB estimates for higher intensity informal caregivers

	SF-12 PCS	SF-12 MCS	SF-36 vitality	CES-D-10 depression
Intensive informal care (≥8 hours)	0.06 (0.27)	-0.71** (0.29)	-1.11** (0.48)	0.10 (0.14)
Illness family member	0.14 (0.15)	-0.44** (0.17)	-0.28 (0.27)	0.17** (0.08)
Hansen J-test	0.5 (0.78)	0.96 (0.62)	3.2 (0.20)	2.91 (0.23)
Number of instruments	3	3	3	3
N	13,626	13,626	14,824	14,835

Robust standard errors are in parentheses. Included controls: lagged health, age, age^2, financial difficulties, married, children at home, employed, standardized household income, mother alive, father alive, wave dummies. ***p < 0.01; **p < 0.05; *p < 0.1.

The effects also seem to differ by gender. When stratifying the sample by gender, both the caregiving effect and the family effect only affect the mental health of females. Males, by contrast, experience a physical health decline in response to informal caregiving (table 8). The caregiving effect significantly differs between both groups at the 0.01 level for all health outcomes except vitality.

Table 8: AB estimates males and females

	Male				Female			
	SF-12 PCS	SF-12 MCS	SF-36 vitality	CES-D-10 depression	SF-12 PCS	SF-12 MCS	SF-36 vitality	CES-D-10 depression
Informal care	-0.56**	-0.06	-0.22	-0.09	0.34	-0.81**	-0.79	0.34**
	(0.27)	(0.29)	(0.48)	(0.14)	(0.27)	(0.32)	(0.49)	(0.14)
Severe illness spouse/family	0.11	-0.37	-0.45	0.15	0.19	-0.51**	-0.14	0.19
	(0.19)	(0.23)	(0.35)	(0.11)	(0.23)	(0.25)	(0.41)	(0.12)
Hansen J-test	0.67	0.18	3.50	0.92	0.18	0.97	0.93	2.8
	(0.71)	(0.91)	(0.18)	(0.63)	(0.91)	(0.62)	(0.63)	(0.25)
# of instruments	3	3	3	3	3	3	3	3
N	7,588	7,588	8,228	8,227	6,038	6,038	6,596	6,608

Robust standard errors are in parentheses. Included controls: lagged health, age, age^2, financial difficulties, married, children at home, employed, standardized household income, mother alive, father alive, wave dummies. ***p < 0.01; **p < 0.05; *p < 0.1.

To assess whether the caregiving effect and the family effect differ when providing care to either parents or spouses, we carry out our analyses for a sample of spousal caregivers and a sample of parental caregivers. We present the results for parental caregivers in table 9 and spousal caregivers in table 10.

Considering spousal and parental caregivers, we note that the caregiving effect is especially present among spousal caregivers and affects both physical heath, vitality and depression scores. Caregiving to a spouse for example leads to significant positive effects on depression scores. This effect relates to a change of more than 10% of the average CES-D-10 score, and is substantial considering the average yearly individual change in CES-D-10 scores of about 0.05. Illness of a spouse also leads to drops in mental health and vitality and increases in depressive signs. These effects are absent for parental caregivers. The results stratified by gender can be found in tables A2 and A3 in the appendix. The caregiving effect is again mostly present for females. Considering female spousal caregivers we especially note large negative health effects related to vitality and depression. These caregiving effects represent changes at the size of respectively 5 and 15 percent of the mean average scores.

Table 9: AB- regressions informal care to parent

Parental caregivers				
	SF-12 PCS	SF-12 MCS	SF-36 vitality	CES-D-10 depression
Informal care to parent	0.04 (0.25)	-0.44 (0.30)	0.01 (0.46)	0.04 (0.14)
Close family member severely ill	0.18 (0.16)	-0.23 (0.18)	-0.02 (0.29)	0.11 (0.08)
Hansen J-test	0.50 (0.78)	0.98 (0.61)	3.23 (0.20)	2.95 (0.23)
Number of instruments	3	3	3	3
N	13,626	13,626	14,824	14,835

Robust standard errors are in parentheses. Included controls: Lagged health, age, age^2, financial difficulties, married, children at home, employed, financial difficulties, standardized household income, mother alive, father alive, wave dummies. ***p < 0.01; **p < 0.05; *p < 0.1.

Table 10: AB- regressions informal care to spouse

Spousal caregivers				
	SF-12 PCS	SF-12 MCS	SF-36 vitality	CES-D-10 depression
Informal care to spouse	-1.16** (0.53)	-0.61 (0.63)	-1.57* (0.91)	0.60** (0.27)
Spouse severely ill	-0.16 (0.33)	-1.05** (0.44)	-1.50** (0.65)	0.43** (0.19)
Hansen J-test	0.50 (0.78)	0.98 (0.61)	3.16 (0.21)	2.98 (0.23)
Number of instruments	3	3	3	3
N	13,626	13,626	14,824	14,835

Robust standard errors are in parentheses. Included controls: Lagged health, age, age^2, financial difficulties, children at home, employed, financial difficulties, standardized household income, mother alive, father alive, wave dummies. ***p < 0.01; **p < 0.05; *p < 0.1.

Robustness analyses

We performed four tests of robustness of our model estimates, the results of these tests are presented in the appendix. We first compare OLS and FE estimation results (Tables A4 and A5). As OLS (not controlling for individual fixed effects) and FE (suffering from dynamic panel bias), lead to biases in different directions, the estimates of both models represent upper and lower bounds for our autoregressive coefficient (Bond, 2002). Our A-B estimate of lagged health lies between the OLS and

FE estimations, strengthening our confidence in the difference GMM model. The estimated caregiving effect and family effect in both models do not differ substantially from the A-B estimates.

Secondly, in all our models, we rely on self-reported information regarding health shocks of family members. This variable could be prone to justification bias: caregivers may justify their decision to provide care (and possibly their withdrawal from the labor market) by overstating the need of their family member (Bound, 1991). To rule this out, we rerun all models using a variable indicating the occurrence of a negative health shock of a family member (i.e. parent or spouse) as obtained from administrative data. For this test, we use a variable indicating whether a parent or partner has become eligible for formal LTC in the current year. The estimated family effects do not differ much between both indicators (table A6). We use severe illness instead of LTC eligibility in the main analyses, as this variable covers a broader set of health problems.

Third, our model rests on the assumption that our first-difference model along with carefully selected covariates eliminates all potential biases. A threat is that our indicator of the family effect does not capture all family related health shocks the individual deals with, thereby overestimating our caregiving estimate. To test the robustness of our estimates we include a selection of variables regarding the health of family members from administrative data. These variables have been selected by a machine learning algorithm (Belloni et al., 2012) from a large set of detailed information on use of medication, medical care expenditures and long-term care eligibility. The caregiving effect on mental health persists after adding these additional health variables (table A7). In our main models, we rely on the health shock question as it serves as a more intuitive proxy of the family effect.

Finally, we test the robustness of our results using a different methodology. Instead of the A-B method, we use a bias-correction estimation method to correct for the dynamic panel bias. For this test we make use of the Stata command *xtlsdvc* (Bruno, 2005). The estimated caregiving and family effect closely match our initial results (table A8). We prefer to rely on difference-GMM methods in our main models as it is better suited towards models where the number of observations is large (Bruno, 2005).

6. Discussion and conclusion

Illness and frailty may have health consequences for individuals who might *care for* and *care about* spouses and parents in bad health. We label these two mechanisms the "caregiving effect" and the "family effect", respectively and using a Dutch panel survey of respondents aged 45-65, we find evidence for both. We find negative effects of informal caregiving on the mental health of caregivers. We also find a direct family effect: illness of a family member has a negative impact on own mental health. Our estimates of the caregiving effect on mental health are smaller than the findings of previous studies such as De Zwart et al. (2017) and Schmitz & Westphal (2015). The difference in effect size could be the result of separating the family effect from the caregiving effect, yet could also be shaped

by the Dutch context. The Dutch public long-term care insurance scheme is very comprehensive (Colombo et al., 2011), potentially alleviating pressure to provide informal care.

Our findings contribute to the literature on informal caregiving in a number of ways. First, they highlight the importance of controlling for the family effect when studying informal caregiving. Ignoring the family effect and using health of a family member as instrument for informal caregiving is problematic because the direct negative effect of the health of a family member on the caregiver's health could violate the exclusion restriction of the instrument. This violation could lead to overestimations of the caregiving effect.

Second, our study allowed for some interesting subgroup analyses. These indicate that the negative mental health impact of caregiving differs for certain groups of caregivers. Female caregivers seem to experience larger caregiving effects on mental health than male caregivers. Explanations for this finding could be that females are often the primary caregiver and more likely to experience social pressure to become a caregiver (Pinquart & Sörensen. 2003). The caregiving effect is also larger for spousal than for parental caregivers. An explanation for this distinction between spousal and parental caregivers could lie in the intensity of caregiving; spousal caregivers often provide more informal care than parental caregivers.

The findings of significant negative caregiving and family effects on mental health indicate that policy makers who seek to mitigate the negative spillovers from illness of an elderly person should focus on relieving the burden of caregiving activities but should not neglect the other family members. Furthermore, they highlight the importance of considering the family effect in order to prevent overestimation of the caregiving effect. Finally, the findings show that the impact of caregiving is not the same for all subgroups of caregivers. Especially female and spousal caregivers experience larger negative mental health effects of caregiving. Policy could specifically aim to support these groups of caregivers with targeted interventions.

Acknowledgements

The authors gratefully acknowledge comments on earlier versions of this article by Rudy Douven, Agnès Gramain and Johan Bonekamp. We, furthermore, would like to thank the participants of various seminars and conferences for useful comments and discussion. Additionally, we thank TNO and Statistics Netherlands for providing access to survey and administrative data.

References

Amirkhanyan, A.A., Wolf, D.A., 2006. Parent Care and the Stress Process: Findings From Panel Data. J. Gerontol. 61, S248–S255. https://doi.org/10.1093/geronb/61.5.S248

Arellano, M., Bond, S., 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Rev. Econ. Stud. 58, 277–297. https://doi.org/10.2307/2297968

Belloni, A., Chen, D., Chernozhukov, V., Hansen, C., 2012. Sparse Models and Methods for Optimal Instruments with an Application to Eminent Domain 80, 2369–2429. https://doi.org/10.3982/ECTA9626

Bobinac, A., Van Exel, N.J.A., Rutten, F.F.H., Brouwer, W.B.F., 2010. Caring for and caring about: Disentangling the caregiver effect and the family effect. J. Health Econ. 29, 549–556. https://doi.org/10.1016/j.jhealeco.2010.05.003

Bond, S.R., 2002. Dynamic panel data models: a guide to micro data methods and practice. Port. Econ. J. 1, 141–162. https://doi.org/10.1007/s10258-002-0009-9

Bound, J., 1991. Self-Reported Versus Objective Measures of Health in Retirement Models. J. Hum. Resour. 26, 106–138.

Brenna, E., Di Novi, C., 2016. Is caring for older parents detrimental to women's mental health? The role of the European North-South gradient. Rev. Econ. Househ. 14, 745–778. https://doi.org/10.1007/s11150-015-9296-7

Bruno, G.S.F., 2005. Approximating the bias of the LSDV estimator for dynamic unbalanced panel data models. Econ. Lett. 87, 361–366. https://doi.org/10.1016/j.econlet.2005.01.005

Coe, N.B., van Houtven, C.H., 2009. Caring for mom and neglecting yourself? The health effects of caring for an elderly parent. Health Econ. 18, 991–1010. https://doi.org/10.1002/hec

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.

de Zwart, P.L., Bakx, P., van Doorslaer, E.K.A., 2017. Will you still feed me when I'm 64 ? The health impact of caregiving. Health Econ. 26, 127–138. https://doi.org/10.1002/hec.3542

Do, Y.K., Norton, E.C.., Stearns, S.., Van Houtven, C.H., 2015. Informal care and caregiver's health. Health Econ. 24, 224–237. https://doi.org/10.1002/hec

Eggink, E., Ras, M., Woittiez, I., 2016. Dutch long-term care use in an ageing population. J. Econ. Ageing 9, 63–70. https://doi.org/10.1016/j.jeoa.2016.08.001

Gezondheidsmonitor, 2016. Own calculations using: Health monitor adults. Public Health Services, Statistics Netherlands, and National Institute for Public Health and the Environment.

Hansen, L., 1982. Large Sample Properties of Generalized Method of Moments Estimators. Econometrica 50, 1029–1054.

Harel, D., Thombs, B.D., Hudson, M., Baron, M., Steele, R., 2012. Measuring fatigue in SSc: A comparison of the short form-36 vitality subscale and functional assessment of chronic illness therapy-fatigue scale. Rheumatol. (United Kingdom) 51, 2177–2185. https://doi.org/10.1093/rheumatology/kes206

Heger, D., 2016. The mental health of children providing care to their elderly parent. Health Econ. 26, 1617–1629. https://doi.org/10.1002/hec

Hewlett, S., Dures, E., Almeida, C., 2011. Measures of fatigue: Bristol Rheumatoid Arthritis Fatigue Multi-Dimensional Questionnaire (BRAF MDQ), Bristol Rheumatoid Arthritis Fatigue

Numerical Rating Scales (BRAF NRS) for Severity, Effect, and Coping, Chalder Fatigue Questionnaire (CFQ), Checklist . Arthritis Care Res. 63. https://doi.org/10.1002/acr.20579

Hirst, M., 2004. Health Inequalities and Informal Care: End of Project Report. Social Policy Research Unit, University of York, York, pp. 1–53.

Lechner, M., 2009. Long-run labour market and health effects of individual sports activities. J. Health Econ. 28, 839–854. https://doi.org/10.1016/j.jhealeco.2009.05.003

Nickell, S., 1981. Biases in dynamic panel models with fixed effects. Econometrica 49, 1417–1426. https://doi.org/10.2307/1911408

Pinquart, M., Sorensen, S., 2011. Spouses, Adult Children, and Children-in-Law as Caregivers of Older Adults: A Meta-Analytic Comparison. Phychology and Aging 26, 1–14. https://doi.org/doi:10.1037/a0021863

Pinquart, M., Sörensen, S., 2007. Correlates of Physical Health of Informal Caregivers : A Meta-Analysis. J. Gerontol. 62, 126–137.

Pinquart, M., Sörensen, S., 2003. Differences between caregivers and noncaregivers in psychological health and physical health: A meta-analysis. Psychol. Aging 18, 250–267. https://doi.org/10.1037/0882-7974.18.2.250

Radloff, L.S., 1977. The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. Appl. Psychol. Meas. 1.

Roodman, D., 2006. How to Do xtabond2: An Introduction to "Difference" and "System" GMM in Stata. Cent. Glob. Dev. 1–48. https://doi.org/The Stata Journal

Roy, J., Schurer, S., 2013. Getting stuck in the blues: Persistence of mental health problems in Australia. Heal. Econ. (United Kingdom) 22, 1139–1157. https://doi.org/10.1002/hec.2967

Schmitz, H., Westphal, M., 2015. Short- and medium-term effects of informal care provision on female caregivers' health. J. Health Econ. 42, 174–185. https://doi.org/10.1016/j.jhealeco.2015.03.002

Schulz, R., 1990. Theoretical perspectives on caregiving: concepts, variables, and methods, in: Biegel DE, B.A. (Ed.), Aging and Caregiving: Theory, Research and Policy. Sage Publications, New York, pp. 27–52.

Schulz, R., O'Brien, A.T., Bookwala, J., Fleissner, K., 1995. Psychiatric and Physical Morbidity Effects of Dementia Caregiving: Prevalence, Correlates, and Causes. Gerontologist 35, 771–791.

Schulz, R., Visintainer, P., Williamson, G.M., 1990. Psychiatric and Physical Morbidity Effects of Caregiving. J. Gerontol. 45, 181–191.

Social Support Act - Wet Maatschappelijke Ondersteuning, 2015. http://wetten.overheid.nl/BWBR0035362/2018-01-01.

Van Den Berg, B., Fiebig, D.G., Hall, J., 2014. Well-being losses due to care-giving. J. Health Econ. 35, 123–131. https://doi.org/10.1016/j.jhealeco.2014.01.008

Vitaliano, P.P., Scanlan, J.M., Zhang, J., 2003. Is Caregiving Hazardous to One's Physical Health? A Meta-Analysis. Psychol. Bull. 129, 946–972. https://doi.org/10.1037/0033-2909.129.6.946

Ware, J.E., Kosinski, M., Keller, S.D., 1995. SF-12: How to score the SF-12 physical and mental health summary scales. Health Institute, New England Medical Center.

Ware, J.E., Sherbourne, C.D., 1992. The MOS 36-Item Short-Form Health Survey (SF-36): I. Conceptual Framework and Item Selection. Med. Care 30, 473–483.

Ybema, J.F., Geuskens, G.A., Van Den Heuvel, S.G., De Wind, A., Leijten, F.R.M., Joling, C.,

Blatter, B.M., Burdorf, A., Van Der Beek, A.J., Bongers, P.M., 2014. Study on Transitions in Employment, Ability and Motivation (STREAM): The Design of a Four-year Longitudinal Cohort Study among 15,118 Persons Aged 45 to 64 Years. Br. J. Med. Med. Res. 4, 1383–1399.

Appendix

Table A1: Descriptive statistics of sample at baseline

	Male					Female				
	Never Informal caregiver		Ever Informal caregiver			Never Informal caregiver		Ever Informal caregiver		
Health outcomes	Mean	SD	Mean	SD		Mean	SD	Mean	SD	
SF-12 physical component scale	50,35	8,39	50,35	8,52		47,28	11,59	47,67	10,81	
SF-12 mental component scale	53,15	7,69	51,72	9,26	***	51,70	9,15	50,88	9,59	**
SF-36 vitality scale	68,24	17,87	66,38	19,58	***	62,79	20,73	61,96	19,74	
CES-D-10 depression scale	4,60	4,49	5,17	5,15	***	5,78	5,28	6,08	5,37	
Health family										
Severe illness spouse/family	0,14	0,34	0,25	0,43	***	0,14	0,35	0,29	0,45	***
Severe illness spouse	0,02	0,15	0,08	0,27	***	0,02	0,15	0,06	0,23	***
Severe illness close family	0,11	0,32	0,18	0,38	***	0,12	0,32	0,25	0,43	***
Personal characteristics										
Age	54,05	5,49	54,54	5,07	***	53,18	5,54	53,50	5,24	
Age Squared	2951,4	594,2	2999,9	551,3	***	2858,9	595,1	2890,2	565,9	
Married/registered partnership	0,89	0,31	0,88	0,33	***	0,82	0,38	0,79	0,41	***
Children living at home	0,49	0,50	0,46	0,50	***	0,44	0,50	0,42	0,49	
Employed	0,93	0,26	0,91	0,29	***	0,81	0,39	0,81	0,39	
Perc. group household income	69,08	21,83	69,16	22,57		66,58	24,14	67,00	23,84	
Financial difficulties	0,17	0,38	0,20	0,40	***	0,22	0,41	0,20	0,40	
Father alive	0,53	0,50	0,62	0,49		0,53	0,50	0,69	0,46	***
Mother alive	0,31	0,46	0,33	0,47	**	0,31	0,46	0,37	0,48	***
Number of observations in 2010	3006		1394			1648		1880		

SF-12 PCS and MCS and SF-36 Vitality range from 0-100 (lowest - highest level of health). CES-D-10 ranges from 0-30, a score ≥10 is considered a sign of depression. ***p < 0.01; **p < 0.05; *p < 0.1 indicate differences between never and ever IC sample

Table A2: AB estimates parental caregivers

A-B Parents	Males				Females			
	SF-12 PCS	SF-12 MCS	SF-36 vitality	CES-D-10 depression	SF-12 PCS	SF-12 MCS	SF-36 vitality	CES-D-10 depression
Informal care	-0.61*	0.18	0.24	-0.21	0.52	-0.88**	-0.09	0.22
	(0.34)	(0.39)	(0.61)	(0.19)	(0.36)	(0.43)	(0.65)	(0.20)
Severe illness close family	0.15	-0.21	-0.31	0.17	0.22	-0.26	0.23	0.07
	(0.20)	(0.24)	(0.39)	(0.12)	(0.24)	(0.25)	(0.43)	(0.12)
Hansen J-test	0.65	0.19	3.49	0.92	0.17	1.02	0.99	3.00
	(0.72)	(0.91)	(0.18)	(0.63)	(0.92)	(0.60)	(0.61)	(0.22)
Number of instruments	3	3	3	3	3	3	3	3
N	7,588	7,588	8,228	8,227	6,038	6,038	6,596	6,608

Robust standard errors are in parentheses. Included controls: Lagged health, age, age^2, financial difficulties, married, children at home, employed, financial difficulties, standardized household income, wave dummies. ***p < 0.01; **p < 0.05; *p < 0.1

Table A3: AB estimates spousal caregivers

A-B Spouses	Males				Females			
	SF-12 PCS	SF-12 MCS	SF-36 vitality	CES-D-10 depression	SF-12 PCS	SF-12 MCS	SF-36 vitality	CES-D-10 depression
Informal care	-1.32*	-0.49	0.54	0.33	-0.99	-0.73	-3.98***	0.93***
	(0.71)	(0.75)	(1.20)	(0.41)	(0.79)	(1.04)	(1.36)	(0.35)
Severe illness spouse	-0.31	-0.73	-1.14	0.17	0.03	-1.42**	-1.85*	0.68**
	(0.44)	(0.55)	(0.83)	(0.26)	(0.50)	(0.69)	(1.00)	(0.28)
Hansen J-test	0,69	0.17	3.32	0.94	0.18	1.12	1.20	3.03
	(0.71)	(0.92)	(0.19)	(0.62)	(0.91)	(0.57)	(0.55)	(0.22)
Number of instruments	3	3	3	3	3	3	3	3
N	7,588	7,588	8,228	8,227	6,038	6,038	6,596	6,608

Robust standard errors are in parentheses. Included controls: Age, age^2, financial difficulties, children at home, employed, financial difficulties, standardized household income, wave dummies. ***p < 0.01; **p < 0.05; *p < 0.1

Table A4: OLS estimations of main model

OLS				
	SF-12 PCS	SF-12 MCS	SF-36 vitality	CES-D-10 depression
Informal care	-0.24**	-0.31***	-0.57***	0.17***
	(0.11)	(0.11)	(0.21)	(0.06)
Severe illness spouse/family	-0.22*	-0.80***	-0.90***	0.34***
	(0.12)	(0.13)	(0.23)	(0.07)
Lagged health	0.71***	0.53***	0.69***	0.63***
	(0.01)	(0.01)	(0.01)	(0.01)
Age	-1.38***	-1.01***	-2.55***	0.65***
	(0.18)	(0.18)	(0.35)	(0.10)
Age squared	0.01***	0.01***	0.02***	-0.01***
	(0.00)	(0.00)	(0.00)	(0.00)
Married/Registered partnership	0.13	0.76***	1.21***	-0.47***
	(0.14)	(0.16)	(0.28)	(0.08)
Children living at home	0.42***	0.05	0.50**	-0.47
	(0.11)	(0.11)	(0.21)	(0.08)
Employed	2.63***	1.49***	2.77***	-1.05***
	(0.15)	(0.14)	(0.27)	(0.08)
Financial difficulties	-0.63***	-1.05***	-2.34***	0.78***
	(0.13)	(0.14)	(0.25)	(0.07)
Perc. group household income	0.01***	0.01**	0.01***	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)
Mother Alive	0.17*	0.08	0.24	-0.10*
	(0.10)	(0.10)	(0.19)	(0.05)
Father Alive	0.29*	-0.02	0.19	-0.06
	(0.11)	(0.11)	(0.21)	(0.06)
N	21.539	21.539	22.789	22.800

Robust standard errors are in parentheses. All regressions include wave dummies. ***p < 0.01; **p < 0.05; *p < 0.1.

Table A5: Fixed effects estimations of main model

Fixed Effects				
	SF-12 PCS	SF-12 MCS	SF-36 vitality	CES-D-10 depression
Informal care	-0.07	-0.34*	-0.47	0.15*
	(0.17)	(0.18)	(0.30)	(0.08)
Severe illness spouse/family	0.15	-0.35**	-0.23	0.14*
	(0.13)	(0.15)	(0.24)	(0.07)
Lagged health	-0.25***	-0.26***	-0.24***	-0.27***
	(0.01)	(0.01)	(0.01)	(0.01)
Age	0.37	-0.47	-0.45	0.37
	(0.54)	(0.56)	(1.01)	(0.28)
Age squared	-0.00	0.01	0.01	-0.00
	(0.00)	(0.01)	(0.01)	(0.00)
Married/Registered partnership	-0.55	1.73**	1.85**	-1.11***
	(0.50)	(0.60)	(0.93)	(0.31)
Children living at home	0.19	-0.46*	-0.65	0.10
	(0.26)	(0.28)	(0.48)	(0.13)
Employed	0.39	0.22	-1.30***	-0.15
	(0.25)	(0.26)	(0.48)	(0.13)
Financial difficulties	-0.26	-0.42***	-1.46***	0.43***
	(0.18)	(0.21)	(0.35)	(0.10)
Perc. group household income	-0.00*	0.00	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.00)
Mother Alive	-0.32	-0.48	-0.49	0.09
	(0.35)	(0.34)	(0.60)	(0.18)
Father Alive	0.60	-0.12	0.32	0.07
	(0.38)	(0.45)	(0.72)	(0.20)
N	21.539	21.539	22.789	22.800

Robust standard errors are in parentheses. All regressions include wave dummies. ***p < 0.01; **p < 0.05; *p < 0.1.

Table A6: Robustness check, using LTC eligibility as health shock

	Both genders			
	SF-12 PCS	SF-12 MCS	SF-36 vitality	CES-D-10 depression
Informal care	-0.01	-0.46**	-0.50	0.11
	(0.19)	(0.22)	(0.35)	(0.10)
LTC-eligibility	-0.34	0.33	0.13	0.26**
	(0.22)	(0.28)	(0.43)	(0.13)
Hansen J-test	0.5	1.09	2.80	3.12
	(0.78)	(0.58)	(0.25)	(0.21)
Number of instruments	3	3	3	3
N	13,485	13,485	14,657	14,669

Robust standard errors are in parentheses. Included controls: Lagged health, age, age^2, financial difficulties, married, children at home, employed, financial difficulties, standardized household income and wave dummies. ***p < 0.01; **p < 0.05; *p < 0.1

Table A7: Robustness check, including additional covariates regarding health of family members

Lasso	Both genders			
	SF-12 PCS	SF-12 MCS	SF-36 vitality	CES-D-10 depression
Informal care	-0.11	-0.39*	-0.45	0.12
	(0.20)	(0.22)	(0.36)	(0.10)
Lagged health	0.09***	0.12***	0.09***	0.09***
	(0.03)	(0.02)	(0.02)	(0.02)
Age squared	-0.00	-0.00*	-0.00**	0.00
	(0.00)	(0.00)	(0.43)	(0.00)
Employed	0.47	-0.29	-1.71***	0.02
	(0.29)	(0.32)	(0.55)	(0.16)
Poor	-0.28	-0.18	-1.08***	0.29**
	(0.22)	(0.26)	(0.42)	(0.12)
Perc. group household income	-0.01	0.02*	0.01	-0.00
	(0.01)	(0.01)	(0.01)	(0.00)
Self-reported illness family	0.10	-0.48***	-0.46*	0.18**
	(0.15)	(0.17)	(0.28)	(0.08)
Children at home	0.03		-0.79	
	(0.31)		(0.58)	
Spouse – prescription for drugs used in diabetes	2.81**			
	(1.20)			
Mother - prescription for drugs for obstructive airway diseases	0.36			
	(0.39)			
Father - prescription for endocrine therapy	0.06			
	(0.81)			
Married		0.87	0.78	-0.44
		(0.81)	(1.16)	(0.37)
Mother prescription for psycho-analeptics		-0.41	0.07	-0.26
		(0.41)	(0.66)	(0.19)
Spouse prescription other products for alimentary tract and metabolism		0.29		
		(0.83)		
Mother - prescription for digestives. including enzymes		1.03		
		(1.65)		
Father – health expenses abroad		0.00*		
		(0.00)		
Spouse - prescription for psycho-analeptics			-1.51	
			(0.94)	
Spouse – Eligible for LTC			-0.88	0.55
			(1.62)	(0.51)
Mother - prescription for muscle relaxants				0.32
				(0.80)
Father - prescription for antiseptics and disinfectants				-1.86**
				(0.91)
Hansen J-test	0.23	0.47	2.64	2.83
	(0.89)	(0.79)	(0.10)	(0.24)
Number of instruments	3	3	3	3
N	12,791	12,791	13,922	13,931

Robust standard errors are in parentheses. All regressions include wave dummies. ***p < 0.01; **p < 0.05; *p < 0.1.

Table A8: Robustness check, main model using bias correction

Bias correction (75 bootstraps)	Both genders			
	SF-12 PCS	SF-12 MCS	SF-36 vitality	CES-D-10 depression
Informal care	-0.14	-0.32*	-0.52*	0.14*
	(0.18)	(0.19)	(0.30)	(0.08)
Severe illness spouse/family	0.13	-0.37***	-0.24	0.16**
	(0.13)	(0.14)	(0.22)	(0.06)
Lagged health	0.09***	0.06***	0.07***	0.06***
	(0.01)	(0.01)	(0.01)	(0.01)
Employed	0.39*	0.09	-1.54***	-0.08
	(0.23)	(0.25)	(0.46)	(0.11)
Financial difficulties	-0.22	-0.34	-1.38***	0.38***
	(0.21)	(0.22)	(0.34)	(0.08)
Perc. group household income	-0.01	0.00	0.00	0.00
	(0.00)	(0.01)	(0.01)	(0.00)
Age	0.63	-0.97*	-0.92	0.46*
	(0.48)	(0.52)	(0.81)	(0.26)
Age squared	-0.00	0.01	0.00	-0.00
	(0.00)	(0.00)	(0.01)	(0.00)
Married/Registered partnership	-0.59	1.51***	1.68**	-0.88***
	(0.44)	(0.48)	(0.85)	(0.21)
Children living at home	0.14	-0.44	-0.68	0.14
	(0.26)	(0.29)	(0.53)	(0.12)
Mother Alive	-0.21	-0.69*	-0.67	0.12
	(0.34)	(0.37)	(0.65)	(0.14)
Father Alive	0.69**	-0.02	0.24	-0.00
	(0.33)	(0.35)	(0.72)	(0.18)
N	21,539	21,539	22,789	22,800

Standard errors are in parentheses. All regressions include wave dummies. ***p < 0.01; **p < 0.05; *p < 0.1.