

1     **The impact of contracting formal care benefits on informal care**  
2             **provision and caregiver well-being: evidence from Japan** \*

3  
4                     Rong FU<sup>†</sup>, Dung LE<sup>‡</sup>, Yoko IBUKA<sup>§</sup>

---

\* This paper uses the confidential data provided by the Ministry of Health, Labour and Welfare (MHLW) (Approval number: 2020-0727-3). The data can be obtained by submitting an application to the MHLW (<https://www.mhlw.go.jp/toukei/sonota/chousahyo.html>). The authors were financially supported by the Open Research Area for the Social Sciences program (ORA) of the Japan Society for the Promotion of Science (JPJSJRP20181403). Institutional Review Board approval was not required since the authors used secondary data.

<sup>†</sup> Corresponding author. School of Commerce, Waseda University, 1-6-1 Nishiwaseda, Shinjuku-ku, Tokyo, 169-8050, Japan. Declarations of interest: none. E-mail: [nataliefu@waseda.jp](mailto:nataliefu@waseda.jp)

<sup>‡</sup> Faculty of economics, Keio University, 2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan. Declarations of interest: none. E-mail: [ldd@keio.jp](mailto:ldd@keio.jp)

<sup>§</sup> Faculty of economics, Keio University, 2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan. Declarations of interest: none. E-mail: [ibuka@econ.keio.ac.jp](mailto:ibuka@econ.keio.ac.jp)

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18

## Abstract

This study investigated how changes in formal care availability affect informal caregiver well-being, adding to the heated discussion on what constitutes an appropriate level of formal care that is both sustainable and adequate. We leveraged a reform of Japan’s public long-term care insurance, which contracted formal care benefits among a group of recipients. Using a difference-in-difference approach with fixed effects to a nationally representative sample of informal caregivers, we examined the dynamic and aggregated impacts of the reform. We found that informal caregivers took on care duties when formal care benefits were reduced, which was detrimental to their well-being. The negative impact was reflected in decreased mobility and stability, and increased stress from financial strains and domestic work. However, the impact was not always detrimental. Caregivers who are female, particularly those who work, are vulnerable to the reduced formal care, whereas male caregivers who do not work or have high incomes benefit from it. Policymakers should take the diverse well-being impacts into account for future reforms.

**Keywords:** formal long-term care, informal caregivers, well-being, difference-in-difference, event study

**JEL codes:** I14, I18, I31, I38

# 1. Introduction

The global growth in the older population is a primary concern that accompanies a rapid demographic transition, resulting in a vastly increased need for long-term care (LTC) (Agree and Glaser, 2009). Evidence suggests that family members account for the majority of caregivers of older people in developed and aging societies, including Japan (Colombo et al., 2011; Ministry of Health, Labour and Welfare (MHLW), 2014).

However, informal care power has been declining for decades and is expected to continue declining because of smaller family sizes, lower fertility rates, increased mobility of younger generations, and a higher proportion of labor force participation among women, who are considered primary family caregivers. As a result, many countries established formal LTC systems. Likewise, in 2000, Japan was one of the first countries to implement a public insurance scheme for LTC, which was deemed to be successful in supporting the elderly's independent life with dignity and relieving family caregivers of care duties (Tamiya et al., 2011).

On the other hand, growing care needs and dwindling informal care supply put upward pressure on the fiscal sustainability of public LTC systems, especially in light of the global economy's stagnation. The annual growth rate of public LTC was 4.6% from 2005 to 2015 among OECD countries, more than double the contemporary GDP growth rate (OECD, 2017). According to projections, public resources devoted to LTC as a proportion of GDP might treble or more by 2060 (De La Maisonneuve and Martins, 2013). Striking a balance between providing adequate formal care and ensuring the fiscal sustainability of such coverage is one of the most difficult tasks for governments.

To that end, governments have been imposing reforms on LTC systems, many of which are geared toward a lower level of welfare protection (Ranci and Pavolini, 2013). In this regard, Japan implemented a significant reform in the insurance system in 2006. The reform attempted to maintain the system's fiscal sustainability by contracting formal care availability for a group of care recipients with moderate care needs. However, contracting formal care sparked public concern not only about potential decreases in care recipients' well-being, but also in caregivers' well-being, given that informal

1 care is substitutable to formal care, particularly among recipients with moderate care needs (Bonsang,  
2 2009; Bremer et al., 2017). Accordingly, there are heated discussions about what constitutes an  
3 appropriate level of formal care that is both sustainable and adequate. Particularly, understanding how  
4 changes in formal care availability affect informal caregiver well-being is critical from a policy  
5 perspective.

6 To address this objective, we leveraged the 2006 reform of Japan's public LTC insurance, which  
7 contracted formal care availability. Using a difference-in-difference approach with fixed effects, we  
8 investigated the dynamic and aggregated impact of contracting formal care benefits on the intensity of  
9 informal care provision and caregivers' well-being in terms of physical and mental health. We found  
10 that informal caregivers take on care duties when formal care is reduced, which is detrimental to their  
11 well-being. The negative impact is reflected in decreased mobility and stability, as well as increased  
12 stress from financial strains and domestic work. We also found that the impact of contracting formal  
13 care is not always detrimental. Caregivers who are female, particularly those who work, are vulnerable  
14 to the negative effects of a lack of formal care, whereas male caregivers who do not work or have high  
15 incomes benefit from it in terms of reduced stress levels.

16 We advise considering the diverse well-being impact of providing informal care for future reforms  
17 on LTC systems. A proper hybrid of formal and informal care is expected to be fiscally sustainable  
18 while also protecting the well-being of both individuals who find informal caregiving hazardous and  
19 those who find it beneficial.

## 21 **2. Literature**

### 22 **2.1. Conceptual frameworks**

23 Several conceptual frameworks have been developed to better understand the relationship between  
24 informal care load and caregiver well-being, which can aid in evaluating how formal care availability  
25 affects caregivers' well-being. A widely used framework is the caregiver stress model (Pearlin et al.,  
26 1990), which proposes two main stressors stemming from providing informal care, namely primary and

1 secondary stressors. Primary stressors refer to strains or problems that are directly connected to the care  
2 situation, such as excessive care intensity (i.e., long care hours). Secondary stressors are the difficulties  
3 associated with performing multiple roles in the family (e.g., as a spouse or parent) or society (e.g., a  
4 formal job) in addition to the caregiving responsibility. Primary stressors are considered the sources of  
5 secondary stressors.

6 According to the stress model, formal care availability is considered a part of the caregiving setting  
7 that affects caregivers' well-being by modifying the two stressors. Formal care availability, as an  
8 external resource to cope with caregiving situations, can buffer informal caregiving intensity and relieve  
9 time constraints due to caregiving, allowing caregivers to perform other roles in family and society.  
10 Therefore, contracting formal care options may negatively impact caregivers' well-being because of an  
11 increase in care intensity (primary stressors), which may in turn increase the likelihood of family and  
12 working-caregiving conflicts, and financial problems (secondary stressors). As conflicts and problems  
13 escalate, informal caregivers may suffer physically and mentally (Lazarus and Folkman, 1984; McEwen,  
14 1998). Furthermore, research suggests that women are more likely than men to be subjected to  
15 caregiving stressors, and that they perceive and cope with these stressors differently than men (Sharma  
16 et al., 2016). As a result, female caregivers tend to report higher care burdens and lower well-being than  
17 male caregivers (Pinquart and Sörensen, 2006).

18 However, the view that caregiving jeopardizes caregivers' well-being may impede the full  
19 understanding of the impact of caregiving. A distinct perspective highlights the advantages of providing  
20 informal care. Specifically, caregivers can acquire control in the sense that they are capable of providing  
21 care and preventing unexpected and unfavorable events when they are free to choose to provide care  
22 based on their knowledge, skills, and resources, rather than finding caregiving to be burdensome (Roth  
23 et al., 2015). In line with this view, a psychological theory suggests that individuals may experience  
24 less stress if they believe that an unfavorable event occurs as a result of their own actions, as opposed  
25 to those who believe that they have no control over the occurrence of that unfavorable event (Rotter,

1 1966). The advantages of providing care also stem from caregivers' altruism and volunteerism (Brown  
2 and Brown, 2014), as well as their personal growth (Marino et al., 2017).

3 Overall, providing informal care may positively or negatively impact caregivers' well-being.  
4 Measurable well-being outcomes frequently represent a mixture of the two competing impacts. The  
5 impact that predominates over the other may be related to the intensity of care, the demographic and  
6 socioeconomic characteristics of the caregivers, and the conditions of the care recipients.

## 7 8 **2.2. Empirical evidence**

9 Several empirical studies, mostly in European countries, examined the association between formal  
10 care availability and caregivers' well-being (Ruppanner and Bostean, 2014). A general finding is that  
11 formal care usage is positively associated with informal caregivers' well-being. Specifically, previous  
12 studies found that formal care usage alleviates the adverse impacts of informal caregiving on caregivers'  
13 well-being in terms of reduced happiness, increased loneliness, worse self-rated health, and a higher  
14 frailty index, and that the magnitude of the alleviation increases with the generosity of formal care  
15 programs (Uccheddu et al., 2019; Verbakel, 2014; Wagner and Brandt, 2015).

16 According to previous studies, the positive association between formal care usage and informal  
17 caregivers' well-being may derive from a decline in informal care provision through the outsourcing of  
18 care responsibilities to professionals (Suanet et al., 2012; Verbakel et al., 2018). The outsourcing can  
19 be at either an extensive level (i.e., probability of providing care) or intensive level (i.e., care intensity).  
20 It can also be at a qualitative level; for instance, informal caregivers tend to undertake less demanding  
21 care tasks (e.g., paperwork) when more demanding ones (e.g., lifting and transferring) are handled by  
22 professionals (Brandt et al., 2009).

23 Although the above-mentioned studies provide valuable information on the positive association  
24 between formal care availability and informal caregivers' well-being, they did not investigate a causal  
25 relationship between the two. Thus, they do not demonstrate how an increase or decrease in formal care  
26 usage benefits or degrades the well-being of informal caregivers. To date, several studies have

1 investigated this association under a causal framework, and the findings have been mixed. Using panel  
2 data, Gonçalves et al. (2021) investigated the impact of formal home care usage on spousal health by a  
3 baseline matching between couples who sought formal care and those who did not. They found that  
4 formal care exerted neither physical nor mental health impacts after controlling for a set of confounders.  
5 Dong et al. (2019) utilized the instrumental variable method to examine how a downward shift from  
6 nursing homes to formal home care affects the well-being of spouse caregivers. They found that the  
7 reduction worsens physical health but improves mental health, which is explained by increased care  
8 obligations and satisfaction, respectively. Some studies use the so-called regime method that categorizes  
9 countries with comparable ideology and political traditions into one group and compares the gradient  
10 in formal care usage across groups to identify the causal relationship. Di Novi et al. (2015) found  
11 disparities between North and South Europe in the impact of informal caregiving on caregivers' well-  
12 being. They discovered that such disparities are caused by gradients in formal care settings, which are  
13 more generous in the north than in the south. Accounting for the formal care gradients, Brenna and Di  
14 Novi (2016) also found that providing informal care affects the mental health of daughter caregivers  
15 only in the south.

16 We contribute to the literature in four aspects. First, we construct a quasi-experimental setting for  
17 causal inference in this context by using an exogenous reduction in formal care benefits among a group  
18 of care recipients. Second, our dataset captures various aspects of well-being, including specific  
19 physical problems and causes of depressive feelings. This unique information helps us gain a deeper  
20 understanding of the reasons underlying the adverse effect of caregiving burden on well-being. Third,  
21 we explore both the dynamic and aggregated impacts of reducing formal care benefits, as well as their  
22 various heterogeneous impacts, providing us with a much more comprehensive picture of the impacts.  
23 Fourth, our findings provide new evidence from Japan, the world's most aged country, which may be  
24 valuable for policymakers in the future as they construct social welfare programs for seniors. One study  
25 in Japan investigated this causal relationship using the policy reform used in our analysis (Miyawaki et  
26 al., 2020). However, our research extends that study by exploring dynamic impacts of the reform,

1 examining a broader variety of well-being outcomes, and providing a more extensive breakdown of the  
2 outcomes. We also adjusted for multiple sets of fixed effects and possible geographic linear trends that  
3 had not been considered previously and carefully selected the data period to avoid confusing impacts  
4 from other reforms that were not previously excluded. Thus, we expected substantially improved  
5 accuracy in the results.

## 6 7 **3. Background**

### 8 **3.1. The long-term care insurance**

9 A public long-term care insurance (LTCI) was introduced in Japan in 2000 with the goal of  
10 socializing LTC for dependent older people and reducing the load on family caregivers. All citizens  
11 aged 40 and older (40+) are obliged to participate in the LTCI, and the monthly premium is determined  
12 by their age and income. Citizens aged 65 or older and those 40+ with specific types of disease are  
13 eligible for LTCI benefits (i.e., recipients) as long as they receive a care-need certificate issued by local  
14 authorities after taking an assessment of their physical and mental status (Fu and Noguchi, 2019).

15 When receiving formal care, recipients copay 10% of the costs until a monthly upper limit is  
16 reached.<sup>1</sup> The upper limit is determined by the recipient's care-need level, which is divided into six  
17 categories, from the mildest support-required level (SL) to care-required level (CL)1, CL2, CL3, CL4,  
18 and to the most severe CL5.<sup>2</sup> The upper limit of monthly care costs is highest for CL5 and decreases as  
19 the care-need level decreases. It is noteworthy that neither the assessment of care needs nor upper limit  
20 requires a mean test.

21 The demand for formal care rises in line with the care-need level, particularly among recipients  
22 classified as CL3 or above, posing a significant fiscal challenge for the government (Campbell et al.,  
23 2010). Public expenditures on LTCI doubled in the first five years after the system was introduced, from  
24 0.7% of GDP in 2000 to 1.3% in 2005 (MHLW, 2019). To ensure fiscal sustainability, the government

---

<sup>1</sup> Recipients have to pay 100% out-of-pocket if their monthly service use exceeds the upper limit.

<sup>2</sup> The assessment exam determines the care-need level.



1 implemented a significant reform in April 2006, which reduced the benefits for recipients with mild  
2 care needs—those in the SL and CL1 categories. The reform included adjustments to both institutional  
3 care and home care, with the latter being the focus of this paper and detailed below.<sup>3</sup>  
4

### 5 **3.2. The 2006 reform**

6 The process of the 2006 reform was twofold. First, the reform refined the care-need levels. It  
7 renamed category SL to SL1 and added a new category, SL2. Therefore, the total number of categories  
8 for care-need levels grew to seven, from the mildest, SL1 and SL2, to the most severe, CL5. Recipients  
9 who were previously classified as SL were reassigned to SL1.<sup>4</sup> Recipients previously classified as CL1  
10 were reclassified as SL2 if their health was not predicted to deteriorate in the following six months;  
11 otherwise, they remained in CL1.

12 Second, the reform restricted the type and quantity of formal care provided to SL1 and SL2  
13 recipients. The types of available services were limited and focused on preventing recipients' health  
14 conditions from deteriorating. Services such as housekeeping and transportation assistance to and from  
15 hospitals were no longer available for SL1 and SL2 recipients. The maximum monthly amount of care  
16 (i.e., the upper limit) was also reduced. The reduction was most noticeable for SL1 recipients, who  
17 could use services worth up to 61,500 JPY per month (17.2% of the upper limits for CL5) before the  
18 reform, but only 50,030 JPY thereafter (13.9% in terms of the upper limits for CL5) (Fu et al., 2017). It  
19 is noteworthy that the reform had no effect on the type or quantity of care provided to CL2–CL5  
20 recipients; thus, we used them as the control group, as fully described in Section 4.2.

21 In sum, the 2006 reform contracted formal care benefits for recipients with mild care needs. The  
22 contraction may induce informal caregiving uptake among caregivers of recipients who were affected.  
23 There is evidence that the reform has reduced caregivers' labor force participation (Fu et al., 2017), but  
24 little is known about its impact on caregivers' well-being.

---

<sup>3</sup> We did not explore how the institutional care reform would affect the well-being of informal caregivers because changes in institutional care were less likely to affect informal caregivers who cared for recipients at home.

<sup>4</sup> Only those who had a serious health deterioration during the reform period were moved to SL2.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21

## 4. Methods

### 4.1. Data

We used the Comprehensive Survey of Living Conditions (CSLC), a large-scale, nationally representative, repeated cross-sectional survey in Japan on non-institutionalized populations, conducted every three years by the MHLW. The CSLC includes four questionnaires assessing household, health, income/savings, and LTC. The dataset is widely used (Fu et al., 2017; Shigeoka, 2014; Tamiya et al., 2011). The household and health questionnaires were administered to over 600,000 respondents from over 300,000 households in each wave and assessed each household member’s demographic, socioeconomic, and health status. The income/savings and LTC questionnaires *complementarily* involved around 100,000 and 6,000 respondents, respectively, supplementing the household and health questionnaires with information on individual income and LTC. Since we used information on LTC, we had to exclude income information from the income/savings questionnaire.

We identified each pair of care recipient and his/her co-residing primary informal caregiver using the household and long-term care questionnaires.<sup>5</sup> We then matched each identified caregiver to his/her own demographic, socioeconomic, and health status information from the household and health questionnaires. We generated the sample from 2001 because the earliest available LTC questionnaire was in 2001, one year after the introduction of LTCI. We omitted waves after 2013 to avoid confounders from the 2015 coinsurance rate reform.<sup>6</sup> Thus, the sample of our study consists of 2001, 2004, 2007, 2010, and 2013 surveys, and consisted of caregivers who cared for a recipient in the same household.<sup>7</sup>

---

<sup>5</sup> The procedure through which we identified an informal caregiver is as follows. We posed a question to each recipient, namely “Please report the household-member identification number of your main caregiver living with you,” and cross-referred the number to that of all household members of this recipient to identify the primary caregiver.

<sup>6</sup> The 2015 reform increased the coinsurance rate from 10% to 20% for recipients with incomes equivalent to the current labor force. This reform may affect formal care usage among recipients with relatively high care-need levels, particularly those classified as CL3 and above.

<sup>7</sup> There are two advantages to focusing on co-residing caregivers. First, care provision by co-residing caregivers may be more time-elastic than that by non-co-residing caregivers, rendering them more vulnerable to reductions in formal care benefits. As a result, we may observe more evident impacts of the reform. Second, the impact of the reform may tend to be homogenous by excluding caregivers who live apart from the recipients.

## 1 **4.2 Measurements**

### 2 **4.2.1 Treatment and control groups**

3 We leveraged the exogenous variation in care-need levels induced by the 2006 reform to identify  
4 the impact of contracting formal care benefits on caregivers' well-being. The treatment group included  
5 informal caregivers of recipients classified as SL before the reform and as SL1 afterward. These  
6 recipients (and consequently their caregivers) experienced contractions in the type and quantity of  
7 available formal care services. Informal caregivers of recipients classified as CL2 or CL3 before and  
8 after the reform comprised the control group. The reform had no impact on these recipients and  
9 caregivers.

10 It is worth noting that caregivers of CL1 recipients throughout the study period and those of SL2  
11 recipients after the reform were not included in either the treatment or control groups because we could  
12 not identify if CL1 recipients were reassigned to SL2 (potentially "treatment") or remained in CL1  
13 (potentially "control"). In addition, caregivers of CL4 and CL5 recipients were not included in the  
14 control group. Although these caregivers were not affected by the reform, their characteristics may  
15 differ from those in the treatment group in a systematic way because they cared for recipients with the  
16 most severe conditions. As a robustness check, we included these caregivers in the control group and  
17 compared the estimates with the main results (see Section 5.5).

### 18 19 **4.2.2 Outcome variables**

20 To measure the intensity of informal care provision, we constructed a binary variable of "care  
21 intensity" by allocating a value of one if caregivers cared for recipients for more than two hours per day  
22 and a value of zero if they only provided care as needed.<sup>8</sup> The variable indicated whether informal care  
23 was provided on a regular basis or only when it was needed.

---

<sup>8</sup> "For how long your primary family caregiver provides care per day?" was the question we used to measure care intensity. There were five possible responses: "full day," "half day," "2 to 3 hours," "only when needed," or "other." If recipients selected one of the first three responses, we assigned one to the outcome variable; if they selected the fourth response, we assigned zero.

1 To assess caregivers' well-being, five health-related dummies were created: "Any symptom,"  
2 "Outpatient visits," "Limitations in activities of daily living (ADL limitations)," "Self-reported poor  
3 health," and "Being stressed." Each dummy was assigned a value of one if the caregivers reported the  
4 corresponding condition and a value of zero if they did not.<sup>9</sup> Thus, these dummies represent poor health  
5 status.

6 We also developed a set of dummies for the breakdowns of specific causes of "Any symptom" and  
7 "Stress," respectively, to take a deeper look at the physical and mental health of informal caregivers.  
8 Each dummy took a value of one if the caregivers reported the corresponding condition and a value of  
9 zero if they did not. We pinned down eleven breakdowns of "Any symptom" associated with organ  
10 systems or physical parts: "systemic symptoms," "respiratory," "musculoskeletal system," "limb,"  
11 "injury," "eyes and ears," "chest," "digestive system," "gum," "skin," and "urinary tract."<sup>10</sup> We  
12 classified the causes of "Stress" into nine groups to maintain consistency across survey waves: "family  
13 relationships," "social networks," "reason for living," "having no free time," "financial strain," "own  
14 health and long-term care issues," "family health and long-term care issues," "childcare and domestic  
15 work," and "marriage, love, and sexual life." Appendix Table A1 shows the structure of the outcomes.

### 17 **4.2.3 Heterogenous analyses**

---

<sup>9</sup> "Do you have any of the symptoms mentioned below?" was the question for "any symptom." If caregivers answered "yes" to any of the symptoms, we assigned a value of one to the outcome. "Do you visit a doctor on a frequent basis?" was the question for "outpatient visits." If caregivers answered yes, we assigned a value of one to the outcome. "Do you have difficulties in the following activities of daily living?" was the question for "difficulties in ADLs." If caregivers had difficulties in any of the ADLs, we assigned a value of one to the outcome. "How is your current health status?" was the question for "self-report poor health." There were five possible options: "very good," "good," "moderate," "bad," or "very bad." We assigned a value of one to the outcome if the caregivers answered "bad" or "very bad." "Do you feel stressed?" was the question for "Stress." We assigned a value of one to the outcome if caregivers answered yes.

<sup>10</sup> "Systemic symptoms" included having fever, feeling dizzy, having a headache, and so on. Nasal congestion and coughing were examples of "respiratory system" symptoms. Backpain and arthralgia were symptoms of the "musculoskeletal system." Feeling numb or chills in the hands or feet were examples of "limb" symptoms. The "injury" consisted of a shattered bone and a gash. Blurred vision and tinnitus were examples of "eye and ear" symptoms. Palpitation and chest pain were examples of "chest" symptoms. Heartburn, diarrhea, and stomachache were examples of "digestive system" symptoms. Toothache and other dental issues were included under the "gum" condition. The "skin" problem comprised a rash or itchy skin, while the "urinary tract" problem included urinary problems.

1 According to the literature, male and female caregivers may have quite different attitudes toward  
2 providing informal care (Pinquart and Sörensen, 2006). They may perceive the stress and benefits of  
3 providing care differently, and, thus, react differently to the reform in terms of changes in well-being  
4 outcomes. Therefore, apart from the full sample estimation, we conducted heterogenous estimations by  
5 gender to investigate the gender disparities in the impact of the reform.

6 To better understand reasons behind the gender disparity in the impact on well-being, we further  
7 performed additional estimations based on female and male caregivers' socioeconomic and  
8 demographic statuses, such as working status (working or not working), income level, as proxied by  
9 per-capita monthly household expenditures (expenditure below the sample median, hereinafter Low  
10 HH-Exp or above the median, hereinafter High HH-Exp),<sup>11</sup> and caregiver-recipient relationships  
11 (spouse, parent, parent-in-law).

#### 13 **4.2.4 Covariates**

14 We further controlled for caregivers' age (in years), marital status (married as reference, single,  
15 widowed, and divorced), health insurance status (national health insurance as reference, employee  
16 health insurance, advance elderly health insurance, and other), pension status (seven types of pensions  
17 with basic pension as reference), gender and age (in years) of recipients, and number of family members.

#### 19 **4.3 Empirical Strategy**

20 Based on the definitions of the treatment and control groups, we conducted an event study based  
21 on difference-in-difference approach with fixed effects (DD-FE) to isolate the impacts of contracting

---

<sup>11</sup> We used monthly household expenditure as a proxy measure of income, as income information from the independent income survey is not available in our sample.

1 formal care on informal caregivers' well-being.<sup>12</sup> We used a linear probability model because all the  
 2 outcomes are binary,<sup>13</sup>

$$3 \quad Y_{it} = \sum_t \alpha_t \text{Treat}_{it} \times \text{Year}_t + \mathbf{X}_{it} \boldsymbol{\beta} + \boldsymbol{\lambda}_c + \boldsymbol{\lambda}_t + \boldsymbol{\lambda}_p + \boldsymbol{\Lambda}_p + \tau_{it},$$

$$4 \quad \text{where } t = 2001, 2007, 2010, 2013, \quad (1)$$

5 for respondent  $i$  in year  $t$ .  $Y_{it}$  indicates the outcomes of care intensity and well-being, as previously  
 6 defined.  $\text{Treat}_{it}$  represents the treatment status and  $\text{Year}_t$  is a set of year dummies with  $\text{Year}_{2004}$   
 7 omitted as the reference. Thus,  $\alpha_t$  is the set of coefficients of interest. Since 2004 is the reference year,  
 8 we expected the pre-reform coefficient  $\alpha_{2001}$  to be statistically indifferent from zero, suggesting that  
 9 there was no pre-trend across treatment status prior to the reform. The “zero”  $\alpha_{2001}$  also implied that  
 10 the reform had no anticipation effect. Coefficients of years post the reform— $\alpha_{2007}$ ,  $\alpha_{2010}$ , and  
 11  $\alpha_{2013}$ —captured the impact of the reform. According to the theoretical framework, the sign of the post-  
 12 reform coefficients might be positive, negative, or zero. They could be *positive* if the adverse well-being  
 13 impact induced by the formal care contraction outweighs the positive impact generated by feelings of  
 14 control and psychological reward; they could be *negative* if the opposite is true; and they can be zero if  
 15 the two impacts cancel one another out. Furthermore, the set of post-reform coefficients allowed us to  
 16 track the dynamic impact of the reform. If the magnitude and significance of the coefficients remain  
 17 constant or increase over time, we might conclude that the reform had a long-term impact on caregiver  
 18 well-being. Instead, if the coefficients become insignificant or their magnitudes revert to zero over time,  
 19 we might conclude that the impact of the reform was transitory.  $\mathbf{X}_{it}$  represents the set of covariates.  
 20  $\boldsymbol{\lambda}_c$  and  $\boldsymbol{\lambda}_t$  re the care-level FE and year FE, respectively, based on which we defined the  
 21 treatment/control group and before/after period.  $\boldsymbol{\lambda}_p$  is the prefecture's FE. In addition to the FEs,  $\boldsymbol{\Lambda}_p$

---

<sup>12</sup> Because our data are repeated cross-sectional, we did not employ propensity score matching (PSM) before conducting the DD estimation as Miyawaki et al. (2020) did. PSM may balance observed treatment-control differences when utilizing cross-sectional or panel data. However, it may not function as expected when using repeated cross-sectional data. Furthermore, because treatment status was assigned based on recipients' care-need levels, the unconfoundedness assumption of PSM is unlikely to hold in this study.

<sup>13</sup> We used a linear probability model rather than logit or probit models also for better interpretability on interaction terms (Puhani, 2012).

1 included either a prefectural linear trend by treatment status  $c\lambda_p$ , a prefectural-year trend  $t\lambda_p$ , or both.  
 2 The linear trends adjusted for geographical variations in the outcome variables over treatment status  
 3 and time. As shown in Table 1, we considered four models, including four possible combinations of  
 4 linear trends.  $\tau_{it}$  is the error term.

5 [Table 1]

6 We also performed an aggregated DD-FE to measure the overall impact of the reform. Specifically,  
 7 we replaced the set of interaction terms in equation (1) with a single interaction between the treatment  
 8 indicator  $\text{Treat}_{it}$  and a dummy  $\text{Post}_t$  indicating years post the reform,

$$9 \quad Y_{it} = \gamma \text{Treat}_{it} \times \text{Post}_t + \mathbf{X}_{it} \boldsymbol{\delta} + \lambda_c + \lambda_t + \lambda_p + \Lambda_p + \tau_{it}. \quad (2)$$

10 The overall impact of the reform over the data period is presented by  $\gamma$ . The magnitude and  
 11 statistical significance of  $\gamma$  are determined by the performance of each post-reform  $\alpha_t$  in equation (1).  
 12 The aggregated DD-FE approach has the advantage of requiring a smaller sample size than the event  
 13 study. Therefore, we estimated equation (2) when conducting heterogeneous analyses by caregiver  
 14 gender, as well as socioeconomic and demographic statuses for each gender.<sup>14</sup> To test the robustness of  
 15 the results, we applied an alternative specification of Equation (2),

$$16 \quad Y_{it} = \rho \text{Robust}_{it} \times \text{Post}_t + \mathbf{X}_{it} \boldsymbol{\sigma} + \lambda_c + \lambda_t + \lambda_p + \Lambda_p + \tau_{it}, \quad (3)$$

17 where  $\text{Treat}_{it}$  was replaced with  $\text{Robust}_{it}$ , in which caregivers of CL2–CL5 recipients served as the  
 18 control group. We expected  $\rho$  to be close to  $\gamma$  in terms of sign, magnitude, and statistical inference.  
 19 For all estimations, standard errors were clustered at the municipality level.

20 When applying the DD-FE approach, a central concern was whether the common trend assumption  
 21 holds. In addition to the  $\alpha_{2001}$  from the event study as an indication of the pre-trend, we also conducted  
 22 a placebo test based on Equation (2) by concentrating on the period before to 2006 and assuming a  
 23 placebo reform in 2003,

---

<sup>14</sup> We were unable to estimate the impact for male caregivers who cared for parents-in-law in the heterogeneous estimations because the limited sample size (23 persons) did not allow us to establish appropriate statistical inference.

$$Y_{it} = \theta \text{Treat}_{it} \times \text{Placebo}_t + \mathbf{X}_{it} \boldsymbol{\pi} + \boldsymbol{\lambda}_c + \boldsymbol{\lambda}_t + \boldsymbol{\lambda}_p + \boldsymbol{\Lambda}_p + \tau_{it}. \quad (4)$$

Placebo<sub>t</sub> took the value of zero for the period before 2003 and one for the period between 2003 and 2006; thus,  $\theta$  shall be non-significant if the common trend holds. The other variables were identically defined as in Equation (2).<sup>15</sup> Specifically, we determined the primary model of Equation (1) and (2) to present and discuss based on how well they performed in the placebo estimation. In the primary model, the common trend should hold for both full-sample and heterogeneous estimations.

Given our data are repeated cross-sectional, another major concern was whether the composition of the treatment and control groups altered before and after the reform. Appendix Table A2 verifies that the proportions of recipients categorized as SL and SL1 (treatment) and CL2-CL3 (control) remained stable before and after the reform. Appendix Table A3 also demonstrates little before-and-after difference in caregiver characteristics across treatment status, except for an increased proportion of female caregivers caring for their spouse in the treatment group, offset by a decrease in those caring for their parents-in-law. In sum, the confounding effect of changes in caregiver composition would be modest in our study. The third key concern was the multiple-testing issue, which may arise as we examined the impacts of reform on a wide range of outcomes. Therefore, in addition to the clustered standard errors, we also adjusted standard errors in accordance with Simes (1986) for testing multiple hypotheses.

## 5. Results

### 5.1. Basic statistics

[Table 2]

Table 2 shows the basic statistics for outcomes and caregiver characteristics that were used for heterogeneous analyses, and the other covariates were adjusted in the estimations. Columns (1) and (2) are for the control and treatment groups before the reform; columns (3) and (4) correspond to the post-

---

<sup>15</sup> We also conducted a placebo test, in which caregivers of CL2 and CL3 recipients were given a placebo treatment. The control group consisted of caregivers of CL4 and CL5 recipients. Appendix Figure B2 summarizes the results. Except for intensive care, we found little difference in well-being between the placebo treatment and the control after the reform.



1 reform period. Column (5) shows the t-statistics of treatment-control differences before the reform, and  
2 column (6) depicts the t-statistics of after-before differences for the treatment group.

3 Column (5) shows a systematic difference in caregivers' well-being between treatment and control  
4 groups prior to the reform. The treatment group provided less intensive care and had fewer unfavorable  
5 health outcomes than the control group. They also reported lower incidence of specific symptoms and  
6 stress. Column (6), however, shows an overall decline in well-being outcomes for the treatment group  
7 following the reform (except for feeling stressed). Furthermore, caregivers in the treatment group had  
8 different characteristics than caregivers in the control group prior to the reform. When compared to the  
9 control group, the treatment group was overrepresented among males, those caring for their parents or  
10 parents-in-law, and those working, younger, more likely to be single or divorced, less likely to enroll in  
11 employee health insurance, and less likely to receive any sort of pension. The basic statistics for the  
12 robustness estimation, which employed caregivers of CL2–CL5 recipients as the control, are included  
13 in Appendix Table A4.

14 The basic statistics by caregiver gender and socioeconomic status are presented in Appendix  
15 Figures A1-A3 and Appendix Table A5. Overall, female caregivers provided as intensive care as male  
16 caregivers, but they were more likely to report symptoms and stress. Caregivers who did not work, had  
17 a higher income, and cared for spouses, on the other hand, provided more intense care and were more  
18 likely to have poor health outcomes, regardless of gender.

## 19 20 **5.2. Placebo tests for common trend and model selection**

21 The treatment-control disparities in caregiver characteristics highlighted the importance of testing  
22 the common trend assumption (i.e., whether the disparities were time-invariant). Therefore, before  
23 providing the results of DD-FE event study and aggregated DD-FE model, we first analyze the placebo  
24 impacts on the main outcomes across the four models based on Equation (4).

25 [Figure 1]

1 All models demonstrated a common trend between the treatment and control groups before the  
2 reform (Figure 1). Appendix Figure B1 further depicts the heterogeneous placebo impacts by caregiver  
3 characteristics, where the four models varied in their ability to satisfy the common trend assumption.  
4 Specifically, Model 3 was the one where the common trend assumption held in both the main and  
5 heterogeneous estimations. Thus, we adopted Model 3 as the primary model, which controlled for year  
6 FE, care level FE, prefecture FE, and prefecture-year trend. For reference, we presented the main results  
7 from all models in Appendix C, where we confirmed that the sign, magnitude, and statistical inferences  
8 of the impacts were consistent across the models.

### 9 10 **5.3. Main results**

#### 11 **5.3.1. DD-FE Event Study**

12 Figure 2 depicts the dynamic impacts of the 2006 reform on informal care intensity and caregivers'  
13 well-being based on Model 3 of the DD-FE event study. Consistent with the placebo tests, the pre-  
14 reform coefficients in 2011 were insignificant, with magnitudes close to zero. This echoed the  
15 conclusion that the treatment-control differences were time-invariant and further suggested that the  
16 reform had no anticipation effect.

17 [Figure 2]

18 In terms of post-reform coefficients, we discovered that caregivers were roughly 16 percentage  
19 points (pp) more likely to provide intensive care immediately after the reform, and that this increase  
20 persisted over time. Such a notably increased care load may have adverse impacts on the caregivers'  
21 well-being. Indeed, right after the reform, caregivers reported a significant increase in symptoms (about  
22 8 pp), with the impact appearing to reduce significantly over time. Meanwhile, although the reform had  
23 no immediate impact on the likelihood of regular outpatient visits, the extent of the impact increased  
24 with time, and caregivers visited hospitals significantly more frequently later in the study period (about  
25 7 pp). The reform appeared to have no statistically significant impacts on the other well-being outcomes.

26 [Figure 3]

1 We further investigated which specific symptoms were responsible for the significant increase in  
2 symptoms. Figure 3 depicts the dynamic impacts of the 2006 reform on each symptom breakdown.  
3 Many of the symptoms increased shortly after the reform, with the top three being musculoskeletal  
4 system (10.3 pp), systemic symptoms (7.4 pp), and chest conditions (4.8 pp). Caregivers appeared to  
5 have more difficulties in functions related to mobility and stability. The deteriorations were somewhat  
6 attenuating, yet generally steady throughout time.

7 [Figure 4]

8 We also examined the dynamic impacts of the reform on each specific reason of stress. Although  
9 caregivers' overall stress levels did not alter much following the reform, Figure 4 reveals that they  
10 tended to be more concerned with financial strain (4.2 pp) and domestic work (3.3 pp) while being less  
11 concerned with reason for life (4.3 pp) and family relations (7.7 pp). The heightened concerns reflected  
12 the fact that the objective reasons for stress—the work-related opportunity costs of caregiving and the  
13 difficulties of allocating time between caregiving and other household tasks—have worsened since the  
14 reform. The lessened concerns suggested that the subjective causes of stress—low self-esteem and  
15 interpersonal tension—have improved after the reform. The changes in specific stresses were less  
16 persistent than changes in specific symptoms and faded with time, suggesting that the mental health  
17 consequences of greater care load seemed to be transitory.

### 19 **5.3.2. Aggregated DD-FE**

20 We also assessed the aggregated impact across our research period based on Equation (2).  
21 Appendix Table D1 confirms that the aggregated impacts were quite comparable with the dynamic  
22 impacts in terms of the sign, magnitude, and statistical inferences. Furthermore, Appendix Table D2  
23 shows that the impacts on main well-being outcomes and specific symptoms survived the adjustment  
24 to multiple-hypothesis testing, whereas the impacts on specific reasons of stress were less resilient to  
25 the adjustment. Our aggregated DD-FE findings were qualitatively consistent with those of Miyawaki  
26 et al. (2020). We did not, however, detect significant deteriorations in self-reported health as they did,

1 and our findings were not as statistically significant as theirs. The differences may be attributed to the  
2 fact that PSM altered the sample for DD in their study.

## 3 4 **5.4. The heterogeneous impacts of the reform**

### 5 **5.4.1. Gender disparities**

6 [Table 3]

7 Since the literature highlighted differences in attitudes towards caregiving by gender, we examined  
8 the gender disparities in the impact on caregivers' well-being. According to Panel A of Table 3, although  
9 both genders provided more intensive care following the reform, the impacts of the reform appeared to  
10 be gender specific. Female caregivers reported a significant decline in physical health. They were more  
11 likely to have symptoms, regular outpatient visits, and ADL limitations. Male caregivers, by contrast,  
12 were significantly less likely to be stressed. The gender disparities were more pronounced in Panels B  
13 and C. Female caregivers reported more symptoms due to significant deterioration in their  
14 musculoskeletal, systemic, chest, respiratory, and skin conditions. They were also more stressed because  
15 of financial strain and domestic work. Male caregivers, on the other hand, reported no decline in  
16 physical health<sup>16</sup> and were less concerned about their own and their family's health and LTC.<sup>17</sup>

17 The gender-specific results raised the question, "what drove men and women to react to increased  
18 care burdens in opposite ways?" To answer the question, we further examined the gender disparities by  
19 caregivers' socioeconomic status (working status and income level) as well as demographic status  
20 (caregiver-recipient relationship).

### 21 22 **5.4.2 Gender disparities by caregivers' characteristics**

23 [Figure 5]

24 [Figure 6]

---

<sup>16</sup> Indeed, male caregivers became 8.5pp less likely to have skin issues after the reform.

<sup>17</sup> Appendix Table E1 reveals that the female results survived after adjusting for multiple-hypothesis testing, but the male results fail the test, indicating a caution when interpreting the findings of male caregivers.

1 [Figure 7]

2 Figures 5-7 show the gender-specific effects of the reform across caregivers' working status,  
3 income level, and caregiver-recipient relationship, respectively. We verified that the considerable  
4 decline in well-being for female caregivers was driven by those who worked, had a relatively high  
5 income (High HH-Exp), and cared for spouses or parents-in-law. According to Appendix Figures F1-  
6 F3, these female caregivers were more likely to report musculoskeletal, systemic, chest conditions, as  
7 well as financial stress. Meanwhile, male caregivers who did not work (most likely retired), had a high  
8 income, and cared for parents were those who had significant mental health improvements. Appendix  
9 Figures F1-F3 reveals that these male caregivers were less concerned about their own and their family's  
10 health and LTC issues.

## 11

### 12 **5.5. Robustness checks**

13 Based on Equation (3), we conducted a robustness check, in which not only caregivers of CL2–  
14 CL3 recipients but also those of CL4–CL5 recipients were included in the control group. Appendix  
15 Figures G1-G3 show the robustness of the main results based on the aggregated DD-FE. The estimates  
16 from the robustness check were close to those of the main results in terms of sign, magnitude, and  
17 statistical inference.

## 18

### 19 **6. Discussion and Conclusion**

20 In this study, we examined the impact of contracting formal care benefits on informal care  
21 provision and the well-being of informal caregivers. We leveraged the 2006 reform of LTCI in Japan  
22 and applied a difference-in-difference approach to isolate both the dynamic and overall impacts of the  
23 reform. We conducted the estimations using not only the full sample of the CSLC survey, but also the  
24 subsamples across caregivers' gender and socioeconomic statuses. For two well-being outcomes—self-  
25 reported symptoms and feeling stressed—we investigated the impact on each of the breakdowns.

1 We confirmed the following: (1) contracting formal care benefits increased the intensity of  
2 informal care immediately and persistently; (2) it had a negative impact on informal caregivers' physical  
3 health, with the greatest impact on self-reported symptoms and outpatient visits; (3) deterioration in  
4 musculoskeletal, systemic, and chest conditions were the top three reasons for the symptom increases;  
5 (4) despite little change in overall mental health, caregivers were more concerned about their financial  
6 situation and ability to undertake domestic tasks; and (5) the reform had a relatively long-term impact  
7 on physical conditions but a transitory impact on mental health. We additionally confirmed the  
8 following by examining the gender-specific impacts of the reform: (6) female caregivers' well-being  
9 deteriorated substantially in terms of both physical and mental health, with the deterioration being  
10 concentrated for those who worked, had a high income, or cared for parents-in-law; (7) male caregivers'  
11 well-being improved in the form of enhanced mental health, with the benefits being more prominent  
12 for those who did not work or had a high income.

13 In line with the literature, findings (1) through (4) indicate that among informal caregivers, the  
14 primary and secondary stressors of caring generally outweigh the benefits, indicating that contracting  
15 formal care is hazardous to their well-being (Bauer and Sousa-Poza, 2015). Increased care intensity  
16 places physical strain on caregivers (the primary stressors), particularly on their musculoskeletal  
17 systems, which may lead to mobility and stability difficulties. Furthermore, without adequate formal  
18 care support, informal caregivers lose a crucial external resource for coping with caregiving situations  
19 and are thus unable to balance their roles as caregivers and other roles in family and society (second  
20 stressors). Indeed, informal caregivers were more concerned about their financial status, which reflects  
21 a high work-related opportunity cost associated with providing more informal care (Heitmueller and  
22 Inglis, 2007). They are also concerned about the difficulties they experience in dealing with other  
23 domestic tasks because of the increased care burden. Finding (5) adds to the literature by revealing that  
24 physical deterioration caused by increased care load is greater and more lasting than mental  
25 deterioration. Although the psychological benefits of providing care may reduce (and in some cases  
26 outweigh) the secondary stressors, the physical load of caregiving is not alleviated. This suggests that

1 the consequences of the primary stressors caused by increasing care load may be more pronounced,  
2 necessitating a comprehensive solution to relieve them.

3 Findings (6) and (7) add to the literature by demonstrating a completely gender-distinct pattern in  
4 changes in well-being as care burden increases. Female caregivers' physical and mental health are both  
5 vulnerable to reductions in formal care benefits, whereas male caregivers suffer little physically and  
6 perform better mentally. Based on a closer examination of the gender-specific reform impact across  
7 caregivers' characteristics, we conclude that the gender disparity is likely driven by two factors that the  
8 conceptual frameworks highlight: opportunity costs and psychological rewards of providing care.

9 For example, the notable decline in well-being of working female caregivers can be attributed to a  
10 larger opportunity cost of caregiving (Carmichael and Charles, 2003; Schmitz and Westphal, 2015). The  
11 cost might be incurred directly by wage loss or indirectly by everyday life compromises. Working  
12 female caregivers were much younger and more likely to be married than their counterparts (Appendix  
13 Table A5). It is easy to see how hard it would be for them to balance their career, social, and married  
14 lives while also providing more intensive care, and it is no surprise that they suffer more following the  
15 reform in musculoskeletal systems, financial stresses, and marriage life (Fu et al., 2017; Schmitz and  
16 Stroka, 2013; Van Houtven et al., 2013). Meanwhile, the notable well-being decline of female caregivers  
17 caring for parents-in-law is likely because they receive fewer psychological benefits due to the less, if  
18 not the least, intimate caregiver-recipient relationship (Heger, 2017; Wakabayashi and Kureishi, 2018).  
19 Although they may not have to strike a balance between career and caregiving as those who work, the  
20 small psychological rewards may make them less motivated to provide care and more vulnerable to  
21 reductions in formal care benefits. The improvement in mental health among nonworking male  
22 caregivers lends credence to our hypothesis. Given that they do not need to work, their opportunity cost  
23 of caregiving could be quite modest. Furthermore, as male caregivers primarily look after their wives  
24 or parents (Appendix Table A5), they may feel better rewarded for engaging in more intimate  
25 relationships while providing care. Because of the modest opportunity cost and strong psychological

1 rewards, nonworking male caregivers benefit from informal care provision (Collins and Feeney, 2000;  
2 Penning and Wu, 2016).

3 In addition to the two factors, we hypothesize that discrepancies in the type of care provided by  
4 men and women may also explain the gender disparity in well-being changes. According to the literature,  
5 women are more likely than men to carry out demanding care tasks, which may be especially true given  
6 Japan's predominant gender role (Sharma et al., 2016). As a result, a similar increase in care hours may  
7 have a more detrimental impact on female caregivers. While suggestive, this hypothesis is supported  
8 by the gender disparity among caregivers with high incomes or caring for spouses. Despite having a  
9 comparable income level (i.e., opportunity costs) or caregiver-recipient relationship (i.e., psychological  
10 rewards), female caregivers are nevertheless more vulnerable to reductions in formal care benefits than  
11 male caregivers. This disparity is most likely the result of women providing more demanding care than  
12 men.

13 This study has two limitations owing to the lack of data. First, we could not assess the care burden  
14 using a broader range of measurements. Care intensity may reflect a significant portion of the care  
15 burden, but not all of it. This is especially true for female caregivers with high income or caring for  
16 spouses. Their great decline in well-being is likely because of care burdens that were not quantified by  
17 care intensity. Second, since we could only observe caregivers of non-institutionalized recipients,  
18 the results may suffer from selection bias. Further studies using more comprehensive datasets and  
19 longer data periods are necessary to clarify our results.

20 Our findings highlight that, while the advantage of providing formal care to recipients with  
21 moderate care needs may appear to be smaller than the cost, it does help the caregivers of those  
22 recipients. Therefore, well-being loss from caregivers shall be taken in account while evaluating the  
23 reform of contracting formal care availability based on care needs. The findings further underline that  
24 the magnitudes of the well-being loss observed in our study may represent a lower bound of the potential  
25 impacts of a reform that targets not only recipients with moderate care needs, but also those in more  
26 severe conditions. Contracting formal care benefits may thus have more negative consequences if future



1 reforms are aimed at a wider range of recipients. We propose alternatives to the need-based method,  
2 such as affordability-based methods, to contain public spending on LTC. In fact, the government  
3 launched an affordability-based reform in 2018, raising the co-insurance cost for affluent recipients.  
4 Future research is needed to assess the impact of this reform. Last but not least, the findings suggest  
5 policymakers to take the heterogeneous impact of providing informal care into consideration. Female  
6 caregivers, particularly those working, are susceptible to an increased care burden, requiring a  
7 comprehensive measure to protect them against financial issues and health issues. Concurrently,  
8 because the benefit of informal care on well-being is pronounced for nonworking men or with high  
9 income, policies that steer them to a fruitful life, including altruistically looking after their family, may  
10 be beneficial to both individual and social well-being.

11

## 12 **References**

- 13 Agree, E.M., Glaser, K., 2009. Demography of informal caregiving, in: *International Handbook of*  
14 *Population Aging*. Springer Netherlands, pp. 647–668. [https://doi.org/10.1007/978-1-4020-8356-](https://doi.org/10.1007/978-1-4020-8356-3_29)  
15 [3\\_29](https://doi.org/10.1007/978-1-4020-8356-3_29).
- 16 Bauer, J.M., Sousa-Poza, A., 2015. Impacts of informal caregiving on caregiver employment, health,  
17 and family. *J. Popul. Ageing*. 8, 113–145. <https://doi.org/10.1007/s12062-015-9116-0>.
- 18 Bonsang, E., 2009. Does informal care from children to their elderly parents substitute for formal care  
19 in Europe? *J. Health Econ*. 28, 143–154. <https://doi.org/10.1016/J.JHEALECO.2008.09.002>.
- 20 Brandt, M., Haberkern, K., Szydlik, M., 2009. Intergenerational help and care in Europe. *Eur. Sociol.*  
21 *Rev.* 25, 585–601. <https://doi.org/10.1093/esr/jcn076>.
- 22 Bremer, P., Challis, D., Hallberg, I.R., Leino-Kilpi, H., Saks, K., Vellas, B., Zwakhalen, S.M.G.,  
23 Sauerland, D., RightTimePlaceCare Consortium, 2017. Informal and formal care: Substitutes or  
24 complements in care for people with dementia? Empirical evidence for 8 European countries.  
25 *Health Policy*. 121, 613–622. <https://doi.org/10.1016/J.HEALTHPOL.2017.03.013>.
- 26 Brenna, E., Di Novi, C., 2016. Is caring for older parents detrimental to women’s mental health? The  
27 role of the European North–South gradient. *Rev. Econ. Household*. 14, 745–778.  
28 <https://doi.org/10.1007/s11150-015-9296-7>.
- 29 Brown, R.M., Brown, S.L., 2014. Informal caregiving: A reappraisal of effects on caregivers. *Soc.*  
30 *Issues Policy Rev.* 8, 74–102. <https://doi.org/10.1111/sipr.12002>.

- 1 Campbell, J.C., Ikegami, N., Gibson, M.J., 2010. Lessons from public long-term care insurance in  
2 Germany and Japan. [https://Doi.Org/10.1377/Hlthaff.2009.0548](https://doi.org/10.1377/Hlthaff.2009.0548). 29, 87–95.  
3 <https://doi.org/10.1377/HLTHAFF.2009.0548>.
- 4 Carmichael, F., Charles, S., 2003. The opportunity costs of informal care: Does gender matter? J.  
5 Health Econ. 22, 781–803. [https://doi.org/10.1016/S0167-6296\(03\)00044-4](https://doi.org/10.1016/S0167-6296(03)00044-4).
- 6 Collins, N.L., Feeney, B.C., 2000. A safe haven: An attachment theory perspective on support seeking  
7 and caregiving in intimate relationships. J. Pers. Soc. Psychol. 78, 1053–1073.  
8 <https://doi.org/10.1037//0022-3514.78.6.1053>.
- 9 Colombo, F., Llena-Nozal, A., Mercier, J., Tjadens, F., 2011. Help Wanted? Providing and Paying for  
10 Long-Term Care. OECD Publishing. <https://doi.org/10.1787/9789264097759-en>.
- 11 De La Maisonnette, C., Martins, J.O., 2013. Public spending on health and long-term care: A new set  
12 of projections. <https://www.oecd-ilibrary.org/content/paper/5k44t7jwwr9x-en?crawler=true>.
- 13 Di Novi, C., Jacobs, R., Migheli, M., 2015. The quality of life of female informal caregivers: From  
14 Scandinavia to the Mediterranean Sea. Eur. J. Population 31, 309–333.  
15 <https://doi.org/10.1007/s10680-014-9336-7>.
- 16 Dong, J., Pollack, H., Konetzka, R.T., 2019. Effects of long-term care setting on spousal health  
17 outcomes. Health Serv. Res. 54, 158–166. <https://doi.org/10.1111/1475-6773.13053>.
- 18 Fu, R., Noguchi, H., 2019. Moral hazard under zero price policy: Evidence from Japanese long-term  
19 care claims data. Eur. J. Health Econ. 20, 785–799. <https://doi.org/10.1007/s10198-019-01041-6>.
- 20 Fu, R., Noguchi, H., Kawamura, A., Takahashi, H., Tamiya, N., 2017. Spillover effect of Japanese  
21 long-term care insurance as an employment promotion policy for family caregivers. J. Health  
22 Econ. 56, 103–112. <https://doi.org/10.1016/j.jhealeco.2017.09.011>.
- 23 Gonçalves, J., von Hafe, F., Filipe, L., 2021. Formal home care use and spousal health outcomes. Soc.  
24 Sci. Med. 287, 114373. <https://doi.org/10.1016/J.SOCSCIMED.2021.114373>.
- 25 Heger, D., 2017. The mental health of children providing care to their elderly parent. Health Econ. 26,  
26 1617–1629. <https://doi.org/10.1002/HEC.3457>.
- 27 Heitmueller, A., Inglis, K., 2007. The earnings of informal carers: Wage differentials and opportunity  
28 costs. J. Health Econ. 26, 821–841. <https://doi.org/10.1016/J.JHEALECO.2006.12.009>.
- 29 Lazarus, R.S., Folkman, S., 1984. Stress, Appraisal, and Coping. Springer, New York.
- 30 Marino, V.R., Haley, W.E., Roth, D.L., 2017. Beyond hedonia: A theoretical reframing of caregiver  
31 well-being. Transl. Issues Psychol. Sci. 3, 400–409. <https://doi.org/10.1037/TPS0000134>.
- 32 McEwen, B.S., 1998. Protective and damaging effects of stress mediators. N. Engl. J. Med. 338, 171–  
33 179. <https://doi.org/10.1056/NEJM199801153380307>.
- 34 Ministry of Health, Labour and Welfare (MHLW), 2014. Results of the Comprehensive survey of  
35 living Conditions, in: *Ministry of Health, Labour and Welfare*, p. FY2013.
- 36 MHLW, 2019. Report on long-term care insurance. <https://www.mhlw.go.jp/toukei/list/84-1.html>.

- 1 Miyawaki, A., Kobayashi, Y., Noguchi, H., Watanabe, T., Takahashi, H., Tamiya, N., 2020. Effect of  
2 reduced formal care availability on formal/informal care patterns and caregiver health: A quasi-  
3 experimental study using the Japanese long-term care insurance reform. *BMC Geriatr.* 20, 207.
- 4 OECD, 2017. Long-term care expenditure. *Health at a Glance*. 2017: OECD. Indicators.  
5 [https://doi.org/https://doi.org/10.1787/health\\_glance-2017-81-en](https://doi.org/https://doi.org/10.1787/health_glance-2017-81-en).
- 6 Pearlman, L.I., Mullan, J.T., Semple, S.J., Skaff, M.M., 1990. Caregiving and the stress process: An  
7 overview of concepts and their measures. *Gerontologist.* 30, 583–594.  
8 <https://doi.org/10.1093/geront/30.5.583>.
- 9 Penning, M. J., & Wu, Z. (2016). Caregiver stress and mental health: Impact of caregiving  
10 relationship and gender. *The gerontologist*, 56(6), 1102-1113.  
11 <https://doi.org/10.1093/geront/gnv038>
- 12 Pinquart, M., Sörensen, S., 2006. Gender differences in caregiver stressors, social resources, and  
13 health: An updated meta-analysis. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 61, P33–P45.
- 14 Puhani, P.A., 2012. The treatment effect, the cross difference, and the interaction term in nonlinear  
15 “difference-in-differences” models. *Econ. Lett.* 115, 85–87.  
16 <https://doi.org/10.1016/J.ECONLET.2011.11.025>.
- 17 Ranci, C., Pavolini, E., 2013 Ranci, C., Pavolini, E. (Eds.). *Reforms in Long-Term Care Policies in*  
18 *Europe*. Springer, New York. <https://doi.org/10.1007/978-1-4614-4502-9>.
- 19 Roth, D.L., Fredman, L., Haley, W.E., 2015. Informal caregiving and its impact on health: A  
20 reappraisal from population-based studies. *Gerontologist.* 55, 309–319.  
21 <https://doi.org/10.1093/geront/gnu177>.
- 22 Rotter, J.B., 1966. Generalized expectancies for internal versus external control of reinforcement.  
23 *Psychol. Monogr.* 80, 1–28. <http://www.ncbi.nlm.nih.gov/pubmed/5340840>.
- 24 Ruppner, L., Bostean, G., 2014. Who cares? Caregiver well-being in Europe. *Eur. Sociol. Rev.* 30,  
25 655–669. <https://doi.org/10.1093/esr/jcu065>.
- 26 Schmitz, H., Stroka, M.A., 2013. Health and the double burden of full-time work and informal care  
27 provision — Evidence from administrative data. *Lab. Econ.* 24, 305–322.  
28 <https://doi.org/10.1016/J.LABECO.2013.09.006>.
- 29 Schmitz, H., Westphal, M., 2015. Short- and medium-term effects of informal care provision on  
30 female caregivers’ health. *J. Health Econ.* 42, 174–185.  
31 <https://doi.org/10.1016/J.JHEALECO.2015.03.002>.
- 32 Sharma, N., Chakrabarti, S., & Grover, S. (2016). Gender differences in caregiving among family-  
33 caregivers of people with mental illnesses. *World journal of psychiatry*, 6(1), 7.  
34 <https://doi.org/10.5498/wjp.v6.i1.7>
- 35 Shigeoka, H., 2014. The effect of patient cost sharing on utilization, health, and risk protection. *Am.*  
36 *Econ. Rev.* 104, 2152–2184. <https://doi.org/10.1257/aer.104.7.2152>.

- 1 Suanet, B., Van Groenou, M.B., Van Tilburg, T., 2012. Informal and formal home-care use among  
2 older adults in Europe: Can cross-national differences be explained by societal context and  
3 composition? *Ageing Soc.* 32, 491–515. <https://doi.org/10.1017/S0144686X11000390>.
- 4 Simes, R. J. (1986). An improved Bonferroni procedure for multiple tests of  
5 significance. *Biometrika*, 73(3), 751-754. <https://doi.org/10.1093/biomet/73.3.751>.
- 6 Tamiya, N., Noguchi, H., Nishi, A., Reich, M.R., Ikegami, N., Hashimoto, H., Shibuya, K., Kawachi,  
7 I., Campbell, J.C., 2011. Population ageing and wellbeing: Lessons from Japan’s long-term care  
8 insurance policy. *Lancet*. 378, 1183–1192. [https://doi.org/10.1016/S0140-6736\(11\)61176-8](https://doi.org/10.1016/S0140-6736(11)61176-8).
- 9 Ucheddu, D., Gauthier, A.H., Steverink, N., Emery, T., 2019. The pains and reliefs of the transitions  
10 into and out of spousal caregiving. A cross-national comparison of the health consequences of  
11 caregiving by gender. *Soc. Sci. Med.* 240, 112517.  
12 <https://doi.org/10.1016/j.socscimed.2019.112517>.
- 13 Van Houtven, C.H., Coe, N.B., Skira, M.M., 2013. The effect of informal care on work and wages. *J.*  
14 *Health Econ.* 32, 240–252. <https://doi.org/10.1016/J.JHEALECO.2012.10.006>.
- 15 Verbakel, E., 2014. Informal caregiving and well-being in Europe: What can ease the negative  
16 consequences for caregivers? *J. Eur. Soc. Policy.* 24, 424–441.  
17 <https://doi.org/10.1177/0958928714543902>.
- 18 Verbakel, E., Metzelthin, S.F., Kempen, G.I.J.M., 2018. Caregiving to older adults: Determinants of  
19 informal caregivers’ subjective well-being and formal and informal support as alleviating  
20 conditions. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 73, 1099–1111.  
21 <https://doi.org/10.1093/geronb/gbw047>.
- 22 Wagner, M., Brandt, M., 2015. Loneliness among informal caregivers aged 50, in: Börsch-Supan, A.,  
23 Kneip, T., Litwin, H., Myck, M., Weber, G. (Eds.), *Ageing in Europe—Supporting Policies for*  
24 *an Inclusive Society*. De Gruyter, pp. 179–188.
- 25 Wakabayashi, M., Kureishi, W., 2018. Differences in the effects of informal family caregiving on  
26 health and life satisfaction between wives and husbands as caregivers. *Rev. Dev. Econ.* 22,  
27 1063–1080. <https://doi.org/10.1111/RODE.12390>.
- 28

1 **Table 1. Model specifications per  $\Lambda_p$**

	<b>Model 1</b>	<b>Model 2</b>	<b>Model 3</b>	<b>Model 4</b>
Prefecture-treatment trend $c\lambda_p$	No	Yes	No	Yes
Prefecture-year trend $t\lambda_p$	No	No	Yes	Yes

2

3

1 **Table 2. Basic statistics**

	Before		After		T-test	
	Control N=2,325 (1)	Treatment N=838 (2)	Control N=4,008 (3)	Treatment N=1,060 (4)	(2)- (1) (5)	(4)- (2) (6)
<b>Outcomes</b>						
Care intensity (%)	52.30	11.81	53.77	13.87	-21.87	1.32
Any symptom (%)	50.67	44.31	49.65	49.29	-3.25	2.20
Outpatient visits (%)	56.54	48.75	59.75	60.13	-3.96	5.04
ADL limitations (%)	22.61	17.98	23.43	21.69	-2.77	1.96
Self-rated poor health (%)	28.18	23.32	31.97	30.32	-2.82	3.50
Being stressed (%)	75.12	65.63	73.95	64.48	-5.27	-0.52
Specific symptoms (%)						
Musculoskeletal system	36.80	29.02	36.03	35.63	-4.18	3.11
Systemic symptoms	25.99	20.31	23.70	24.41	-3.37	2.16
Chest	9.71	5.25	7.14	7.63	-4.10	2.13
Limb	18.54	14.62	18.56	18.57	-2.64	2.33
Respiratory	10.18	6.70	9.71	11.88	-3.07	3.91
Eye and ear	22.50	17.30	20.51	19.79	-3.26	1.41
Digestive system	16.19	12.83	15.47	15.74	-2.38	1.82
Gum	11.27	7.37	9.93	9.05	-3.29	1.35
Skin	9.50	8.59	7.56	8.01	-0.80	-0.47
Urinary tract	7.78	7.14	8.71	9.99	-0.61	2.23
Injury	2.94	2.79	2.10	2.26	-0.23	-0.74
Reasons of stress (%)						
Financial strain	11.81	10.50	13.74	15.52	-1.02	3.17
Childcare and domestic work	6.62	4.06	5.74	6.07	-2.68	1.95
Social networks	5.73	5.01	5.24	8.16	-0.77	2.69
Having no free time	14.32	7.28	14.13	8.26	-5.29	0.78
Marriage, love, and sexual life	1.79	1.67	1.19	1.39	-0.22	-0.48
Reason for living	5.41	4.77	6.09	6.47	-0.71	1.56
Own health and LTC issues	35.12	31.03	19.48	19.20	-2.14	-5.92
Family relationships	13.78	16.59	13.74	15.42	1.97	-0.68
Family health and LTC issues	59.02	41.53	57.99	34.43	-8.77	-3.14
<b>Caregiver's characteristics</b>						
Female (%)	77.70	71.41	72.82	67.05	-3.84	-2.13
Recipient-caregiver relationship (%)						
Spouse	39.90	31.69	40.12	37.06	-4.42	2.55
Parent	23.91	27.41	30.73	33.84	2.11	3.14
Parent-in-law	32.17	37.69	25.22	25.51	3.05	-5.98
Working (%)	38.41	52.04	40.26	44.65	7.23	-3.33
Low HH-Exp (%)	48.37	47.57	48.02	47.03	-0.39	-0.23
Age (years)	62.93 (29.24)	60.11 (12.75)	64.43 (18.93)	63.41 (13.37)	-2.85	5.68
Marital Status (%)						
Married	85.10	82.76	81.43	79.05	-1.68	-2.12
Single	7.17	9.21	10.15	11.19	1.99	1.47
Widowed	4.67	3.85	4.37	3.94	-1.03	0.10
Divorced	3.06	4.18	3.91	5.73	1.61	1.61
Health Insurance (%)						
National health insurance	62.80	58.56	48.19	43.51	-2.27	-6.84
Employee health insurance	35.66	40.37	32.53	35.68	2.53	-2.18
Advance elderly health insurance	0.00	0.00	17.71	19.55	n.a.	15.02
Others	1.53	1.08	1.57	1.26	-1.01	0.39
Pension (%)						
Basic pension	17.83	13.49	22.30	19.44	-3.04	3.61

Basic pension and employee pension	10.43	8.35	17.88	14.07	-1.81	4.06
Basic pension and mutual aid pension	1.81	1.18	2.67	2.33	-1.30	1.95
National pension	9.30	6.85	3.77	3.76	-2.27	-3.15
Employee pension	11.03	9.74	11.36	10.57	-1.09	0.62
Mutual aid pension	2.82	3.21	2.43	3.67	0.61	0.57
Governmental pension	0.68	0.75	0.24	0.36	0.20	-1.21
Household size (person)	3.74	3.77	3.52	3.33	0.59	-6.36
	(1.69)	(1.68)	(1.57)	(1.50)		

1 *Notes:* Standard deviations are in parentheses for continuous variables. The t-statistic for testing the treatment-control  
2 difference in the proportion of advance elderly health insurance prior to the reform is not available because of the zero  
3 proportions.

4

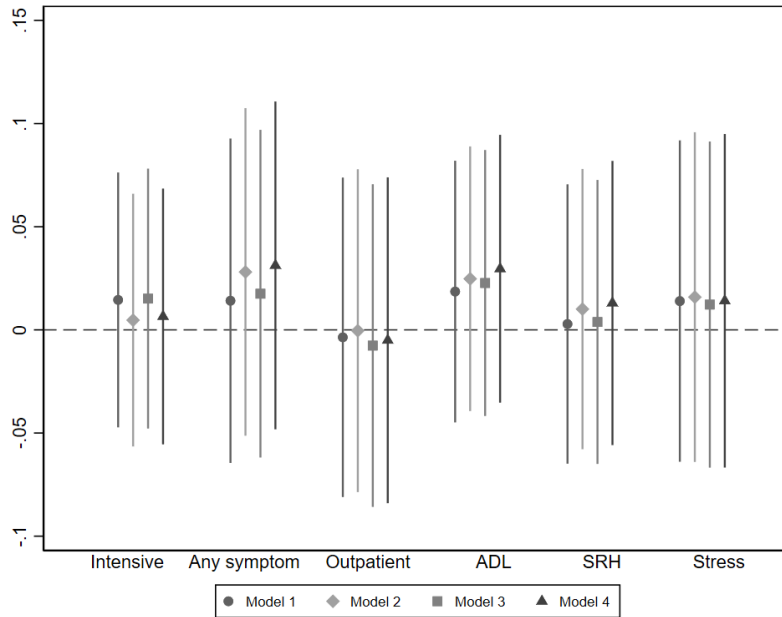
1 **Table 3. Gender disparities in the impact of 2006 reform**

	Female (N=6,339)			Male (N=3,362)		
	Coef.	S.E.	p-value	Coef.	S.E.	p-value
<b>Panel A: Main outcomes</b>						
Care intensity	0.183	***	(0.031)	0.000	0.143	** (0.055) 0.009
Any symptom	0.061	*	(0.034)	0.072	0.050	(0.062) 0.425
Outpatient visits	0.079	**	(0.036)	0.031	-0.016	(0.061) 0.790
ADL limitations	0.062	*	(0.032)	0.052	-0.058	(0.053) 0.278
Self-rated health	0.036		(0.033)	0.287	-0.009	(0.055) 0.877
Being stressed	0.033		(0.036)	0.359	-0.128	** (0.064) 0.045
<b>Panel B: Specific symptoms</b>						
Musculoskeletal system	0.100	**	(0.037)	0.008	0.081	(0.057) 0.159
Systemic symptoms	0.069	**	(0.032)	0.034	0.054	(0.050) 0.275
Chest	0.065	***	(0.019)	0.001	0.003	(0.034) 0.920
Limb	0.043		(0.029)	0.143	0.031	(0.046) 0.507
Respiratory	0.074	**	(0.023)	0.001	-0.034	(0.037) 0.361
Eye and ear	0.045		(0.030)	0.141	-0.006	(0.047) 0.900
Digestive system	0.041		(0.028)	0.136	-0.012	(0.043) 0.787
Gum	0.034		(0.022)	0.116	-0.009	(0.038) 0.808
Skin	0.044	**	(0.021)	0.034	-0.085	** (0.036) 0.018
Urinary tract	0.031		(0.020)	0.117	-0.056	(0.039) 0.148
Injury	-0.009		(0.012)	0.463	0.002	(0.018) 0.919
<b>Panel C: Reasons for being stressed</b>						
Financial strain	0.059	**	(0.026)	0.022	-0.017	(0.041) 0.676
Domestic work	0.040	**	(0.019)	0.033	-0.006	(0.026) 0.805
Social networks	0.029		(0.019)	0.123	0.019	(0.027) 0.486
Having no free time	0.016		(0.025)	0.529	0.013	(0.033) 0.697
Marriage, love and sexual life	0.014		(0.009)	0.107	-0.022	(0.022) 0.320
Reason for living	0.007		(0.018)	0.692	-0.020	(0.027) 0.456
Own health and LTC issues	0.022		(0.033)	0.514	-0.097	* (0.055) 0.080
Family relations	-0.021		(0.029)	0.476	-0.030	(0.040) 0.441
Family health and LTC issues	-0.017		(0.040)	0.667	-0.106	* (0.064) 0.096

2 *Notes:* All estimates were derived based on subsample estimation using Model 3 of Equation (2), which controlled for the full  
3 set of caregiver characteristics and recipients' gender and age, as well as year FE, care level FE, prefecture FE, and prefecture-  
4 year trend. N stands for number of observations, S.E. for standard errors that were clustered at the municipal level. \*Inference:  
5 \*\*\* p < 0.01; \*\* p < 0.05; \* p < 0.1.

6

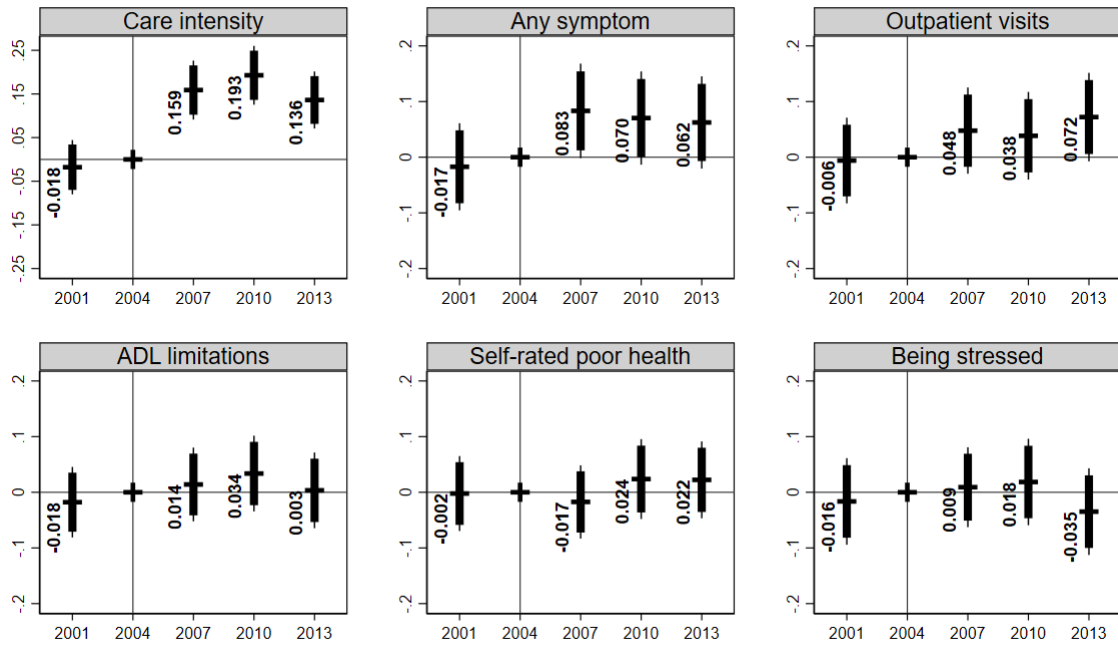




1  
2  
3  
4  
5  
6  
7  
8  
9

**Figure 1. Placebo impacts on informal care provision and caregiver well-being across models**

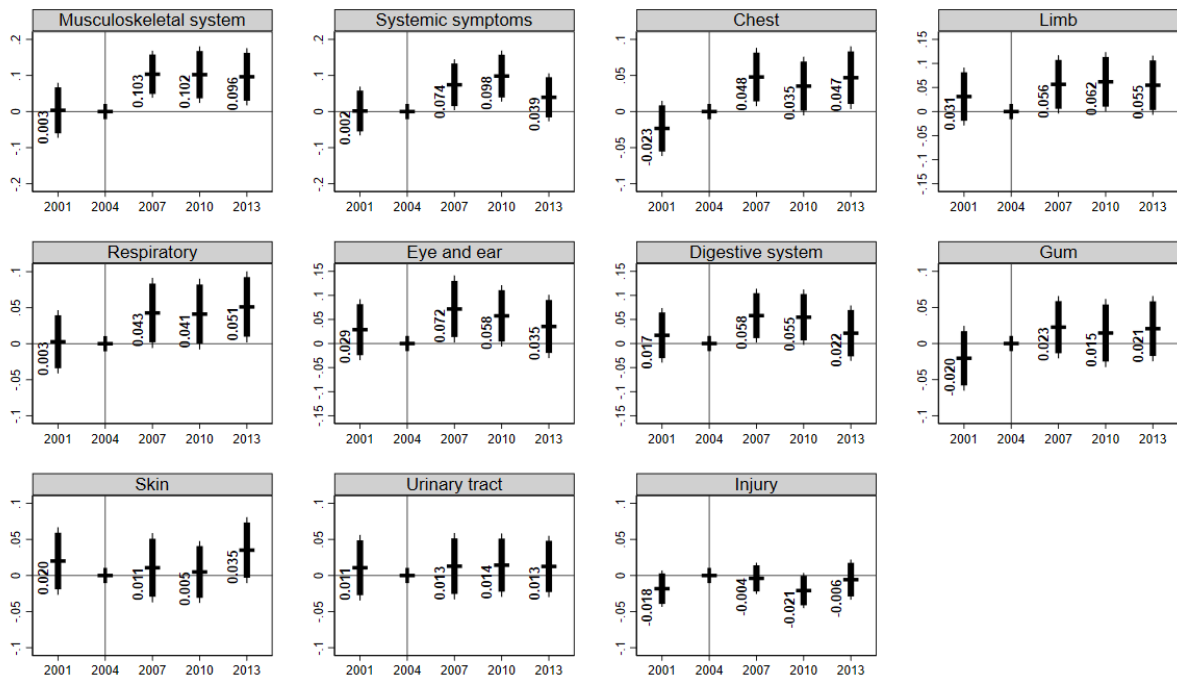
Notes: The markers with 95th confidence intervals represent the estimates of  $\theta$  from Equation (4). Model 1 controlled for the full set of caregiver characteristics and recipients' gender and age, as well as year FE, care level FE, and prefecture FE. Based on Model 1, Model 2 further controlled for prefecture-treatment trend, Model 3 further controlled for prefecture-year trend, and Model 4 further controlled for both trends. Standard errors are clustered at the municipal level. Standard errors are clustered at the municipal level. "intensive" stands for care intensity, "outpatient" for outpatient visits, "ADL" for limitations in ADLs, "SRH" for self-report poor health. Standard errors were clustered at the municipality level.



1  
2  
3  
4  
5  
6  
7  
8  
9

**Figure 2. Dynamic impacts of the 2006 reform on informal care provision and well-being of caregivers**

Notes: All estimates were derived based on Model 3 of Equation (1), which controlled for the full set of caregiver characteristics and recipients' gender and age, as well as year FE, care level FE, prefecture FE, and prefecture-year trend. The y-axes share a uniform scale for visual comparability, except for "care intensity." The markers stand for the estimates of  $\alpha_t$  in Equation (1). The thicker and thinner bars represent confidence intervals at the 90% and 95% levels, respectively. Standard errors were clustered at the municipality level.

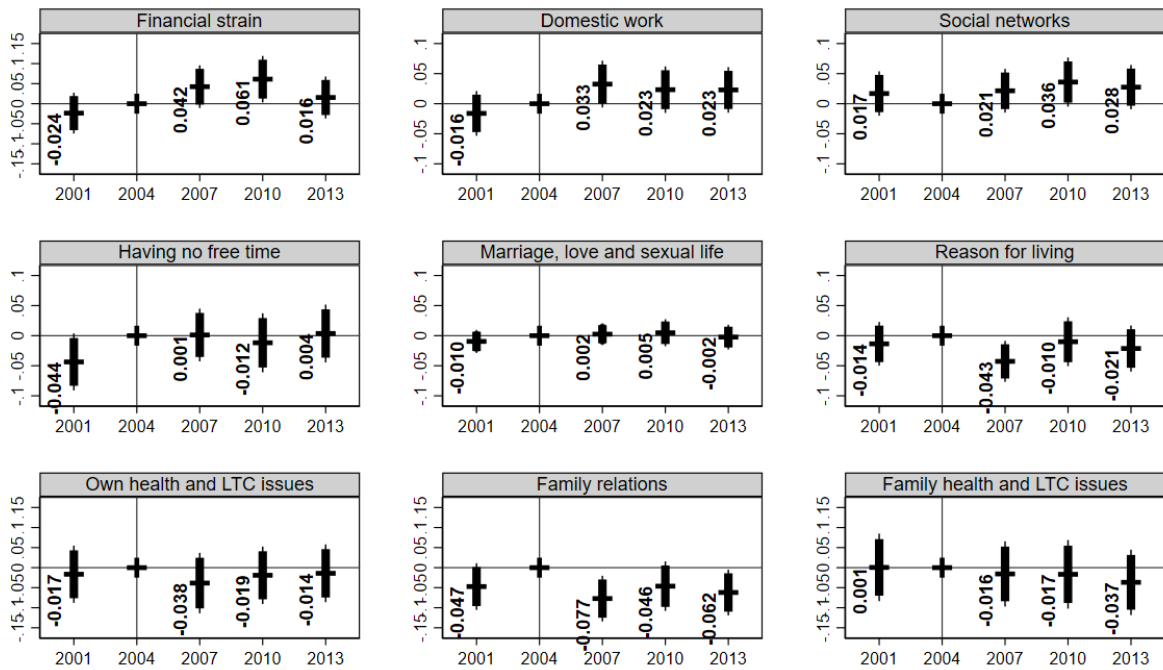


1

2 **Figure 3. Dynamic impacts of the 2006 reform on specific symptoms**

3 *Notes:* All estimates were derived based on Model 3 of Equation (1), which controlled for the full set of caregiver  
 4 characteristics and recipients' gender and age, as well as year FE, care level FE, prefecture FE, and prefecture-year trend. The  
 5 y-axes share a uniform scale for visual comparability, except for "care intensity." The markers stand for the estimates of  $\alpha_t$  in  
 6 Equation (1). The thicker and thinner bars represent confidence intervals at the 90% and 95% levels, respectively. Standard  
 7 errors were clustered at the municipality level.

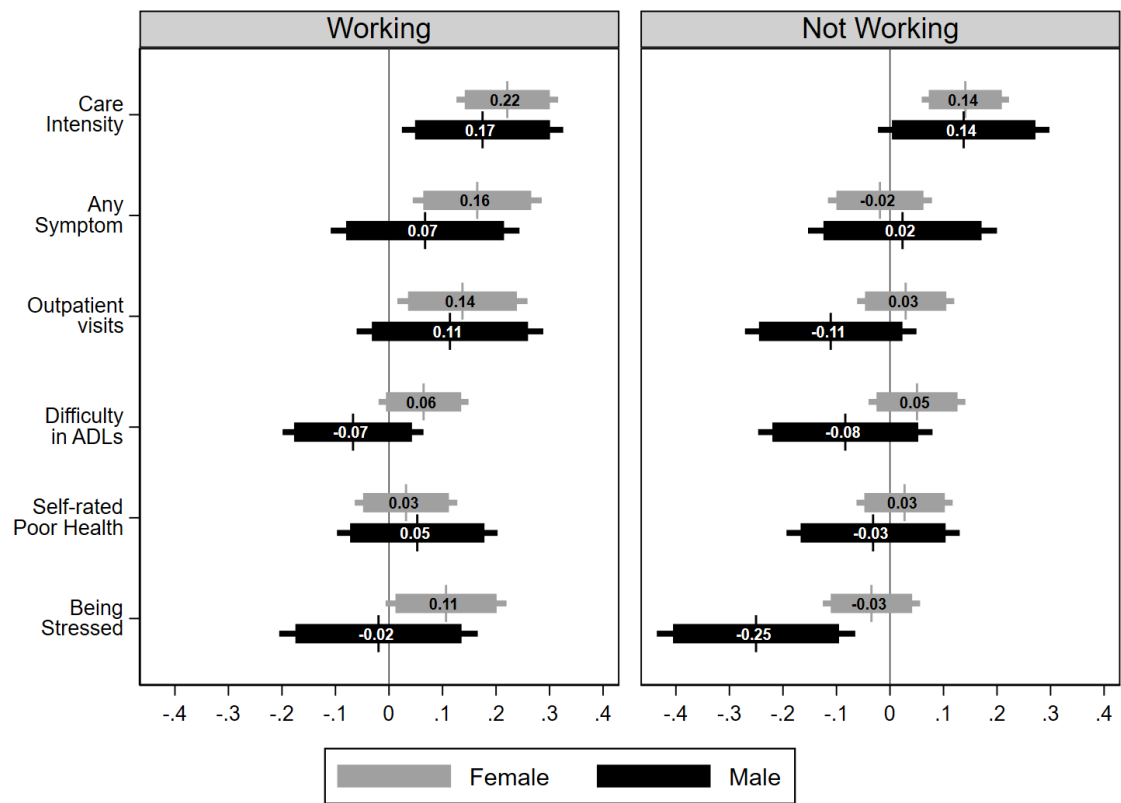
8



1  
2  
3  
4  
5  
6  
7  
8  
9

**Figure 4. Dynamic impacts of the 2006 reform on specific reasons of stress**

Notes: All estimates were derived based on Model 3 of Equation (1), which controlled for the full set of caregiver characteristics and recipients' gender and age, as well as year FE, care level FE, prefecture FE, and prefecture-year trend. The y-axes share a uniform scale for visual comparability, except for "care intensity." The markers stand for the estimates of  $\alpha_t$  in Equation (1). The thicker and thinner bars represent confidence intervals at the 90% and 95% levels, respectively. Standard errors were clustered at the municipality level.

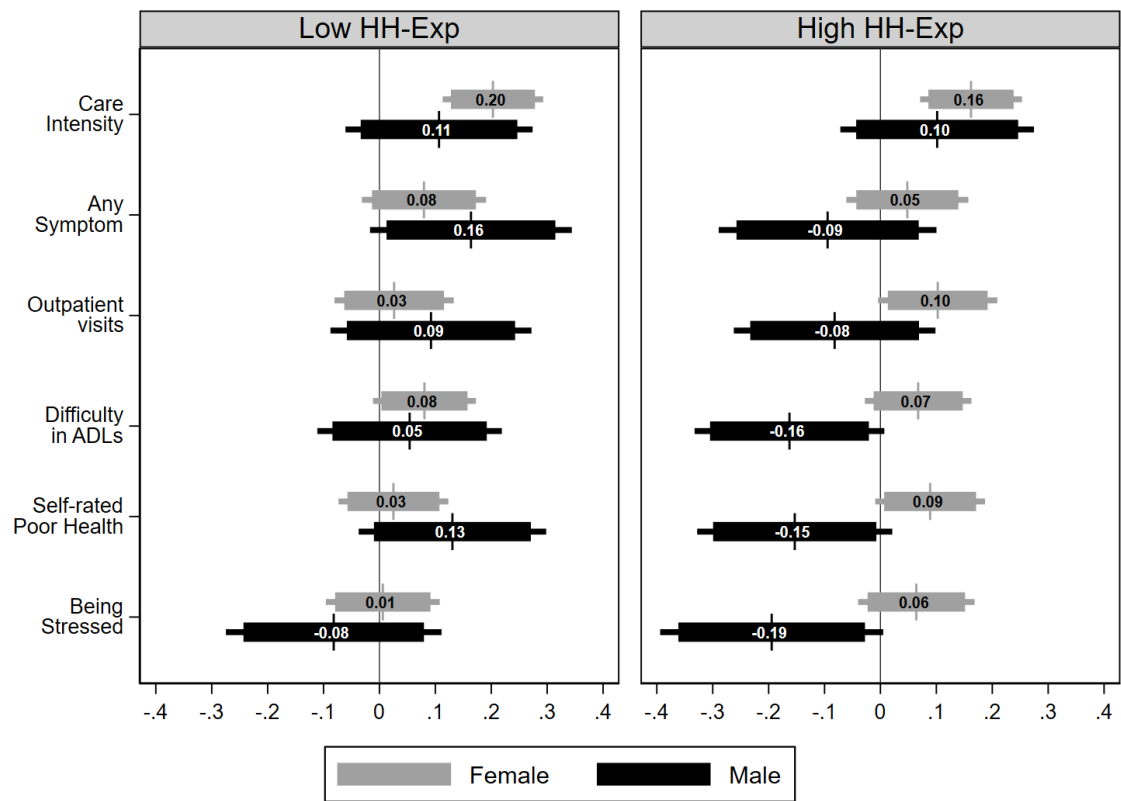


1

2 **Figure 5. Gender disparities in impact of the 2006 reform: caregivers' working status**

3 *Notes:* All estimates were derived based on Model 3 of Equation (2), which controlled for the full set of caregiver  
 4 characteristics and recipients' gender and age, as well as year FE, care level FE, prefecture FE, and prefecture-year trend. The  
 5 vertical bars stand for the estimates of  $\gamma$  in Equation (2). The thicker and thinner horizontal bars represent confidence intervals  
 6 at the 90% and 95% levels, respectively. Standard errors were clustered at the municipality level.

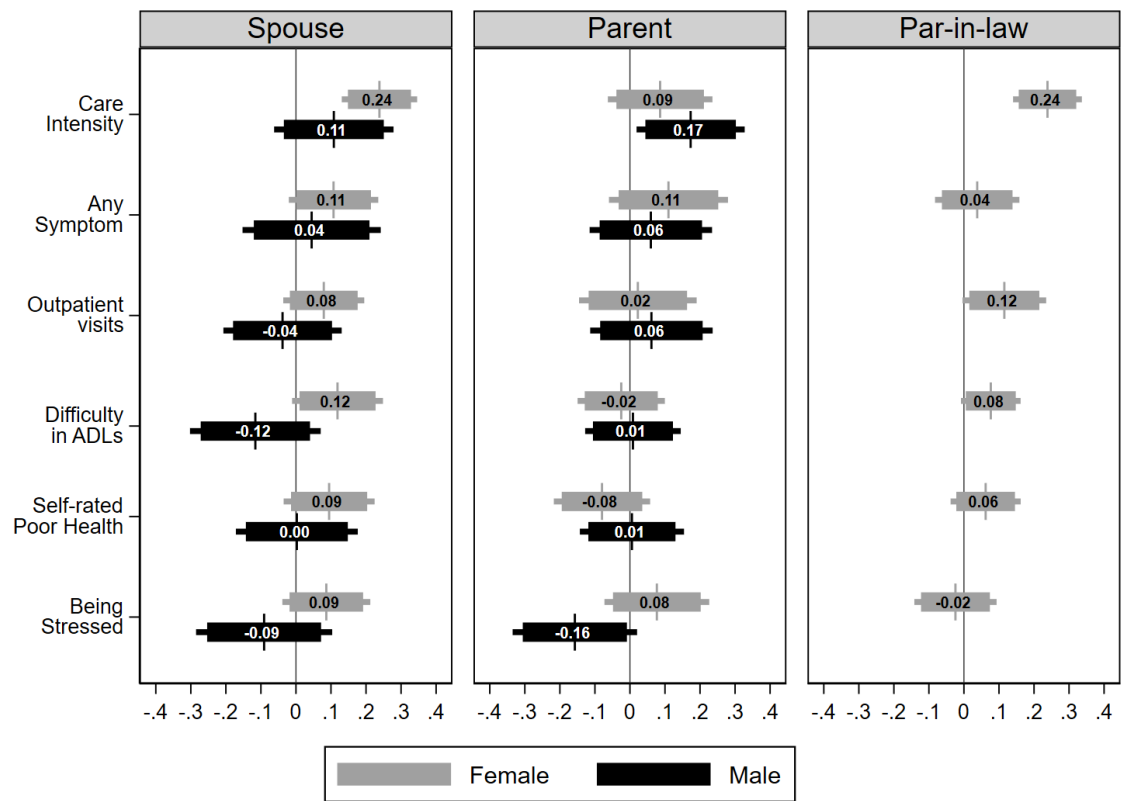
7



1

2 **Figure 6. Heterogeneous impact of the 2006 reform by gender: caregivers' income level**

3 *Notes:* All estimates were derived based on Model 3 of Equation (2), which controlled for the full set of caregiver  
 4 characteristics and recipients' gender and age, as well as year FE, care level FE, prefecture FE, and prefecture-year trend. The  
 5 vertical bars stand for the estimates of  $\gamma$  in Equation (2). The thicker and thinner horizontal bars represent confidence intervals  
 6 at the 90% and 95% levels, respectively. Standard errors were clustered at the municipality level.



1

2 **Figure 7. Heterogeneous impact of the 2006 reform by gender: caregiver-recipient relationships**

3 *Notes:* All estimates were derived based on Model 3 of Equation (2), which controlled for the full set of caregiver  
 4 characteristics and recipients' gender and age, as well as year FE, care level FE, prefecture FE, and prefecture-year trend. The  
 5 vertical bars stand for the estimates of  $\gamma$  in Equation (2). The thicker and thinner horizontal bars represent confidence intervals  
 6 at the 90% and 95% levels, respectively. Standard errors were clustered at the municipality level.

7