HEDG

HEALTH, ECONOMETRICS AND DATA GROUP

WP 17/20

Long-term care reform and the labor supply of informal caregivers – evidence from a quasi-experiment

Johannes Geyer and Thorben Korfhage

August 2017

http://www.york.ac.uk/economics/postgrad/herc/hedg/wps/

Long-term care reform and the labor supply of informal caregivers – evidence from a quasi-experiment

Johannes Geyer*, Thorben Korfhage†

2017

Germany introduced a new insurance for long-term care in 1995 as part of its social security system. Benefits from the long-term care insurance are not means tested and only depend on the required level of care. The new scheme improved the situation for households wanting to organize informal care at home and it changed the incentives for potential informal caregivers. We exploit this reform as a quasi-experiment and examine its effect on the labor supply of caregivers who live in the same household as the care recipient. We find strong negative labor market effects for men but not for women.

Keywords: labor supply, long-term care, long-term care insurance, quasi-experimental approach
JEL Classification: J22, H31, I13

*DIW-Berlin, jgeyer@diw.de
†RWI, thorben.korfhage@rwi-essen.de

1. Introduction

By 2050, the share of the population aged 80 and over is expected to reach 10% in OECD countries – and reaching 15% in Germany and Japan (Colombo et al., 2011). Growing social demand for long-term care (LTC) is expected to pose a real challenge to the LTC systems of ageing societies. Today, average public and private spending on LTC already accounts for as much as 1.5% of GDP across OECD countries. This percentage is projected to double or even triple by 2050 (Colombo et al., 2011).

Individuals in need of care generally prefer to receive support from close family members at home. Across OECD countries, about 10% of the population aged 50 or older report providing informal care on a regular basis (Colombo et al., 2011). Correspondingly, existing LTC systems rely to a large extent on (unpaid) family care and on informal networks. However, it can be challenging for caring relatives to reconcile market work and care obligations. Intensive long-term care is especially associated with a substantial reduction of labor supply. Informal caregivers face additional costs from wage penalties, higher poverty risks and detrimental health effects (Lilly et al., 2007; Colombo et al., 2011; Bauer and Sousa-Poza, 2015). From a budgetary perspective, negative labor supply effects can severely reduce the comparative cost advantage of informal over formal care arrangements.

To avoid this effect, various countries have implemented policies supporting family care arrangements. Our study contributes to the existing literature by providing causal evidence on the impact of these family care policies on caregivers' labor market behavior. Disentangling the ways care benefits might affect caregivers' trade-off between informal care provision and labor supply is a complex task to perform empirically. First, family care policies usually provide different types of benefits – direct cash benefits, benefits in kind and care leave – that can be combined in various ways. Second, several problems of endogeneity have to be resolved.

We investigate labor supply reactions to the introduction of the German long-term care insurance (LTCI) in 1995. Using quasi-experimental variation created by

this reform, we apply a difference-in-differences (DiD) approach to identify the labor supply effects that avoids many common problems of endogeneity. We compare the labor supply of caregivers (treatment group) before and after the LTCI reform with labor supply of similar individuals without care obligations. In the period under study, no other major reform that could yield different macro trends for treatment and control groups, was implemented in Germany. The majority of caregivers we analyze are partners or grown-up children who provide care to their spouse or parent. We conduct various robustness tests to test the validity of our results. The analysis is based on data from the German Socioeconomic Panel Study (SOEP) from 1991-2011.

We contribute to the rich literature on labor supply and caregiving by explicitly analyzing the labor supply effects of care policies. Moreover, the German institutional setting offers a number of interesting features that are important for other countries as well. German LTCI provides benefits solely based on recipients' needs of care. Recipients can choose between cash benefits, benefits in kind and a combination of the two. Benefits are designed to complement – not replace – informal care so that those eligible for care benefits rely on complementary private care. As a consequence, the introduction of LTCI had a nontrivial impact on the incentives to trade off time spent on informal care and working. We think that understanding the labor supply effects of the LTCI in Germany is of general interest. The German LTC system shares many features of systems in other countries. Most countries rely on a similar mix of informal and formal care and many countries provide cash support to family carers. For caregivers it remains a challenge to combine care and work, therefore it is crucial to study the labor market effects of these policies.

The paper is structured as follows: The next section provides a short literature review, Section 3 introduces the relevant institutional setting of Germany, Section 4 describes our data, and Section 5 presents our empirical approach. We discuss the DiD-assumptions in Section 6, where we also provide graphical representations and

descriptive statistics. Our key regression results are presented in Section 7. Section 8 then provides a set of robustness checks. Section 9 discusses our empirical findings, and Section 10 draws some conclusions for future research and policy strategies.

2. Related literature

Studies on the relation of caregiving and labor supply generally face a fundamental identification problem. On the one hand, caregiving reduces available time resources and is therefore difficult to combine with paid work; thus negatively affecting labor supply. On the other hand, individuals who do not work, as well as part-time employees, have fewer time restrictions and are thus more likely to take up caring responsibilities (Bauer and Sousa-Poza, 2015; Heitmueller, 2007). Not all studies follow a causal identification strategy. However, most studies assume the caregiving decision to be endogenous with respect to labor supply. A key concern of research is to find appropriate instruments or research designs to adequately measure this causal link. One strategy is to use panel data in order to control for unobserved characteristics. Another strategy is to use instruments that are assumed to be correlated with the amount of informal care, but have no direct impact on labor market status. Typical instruments include, for example, the health status of potential care receivers, age of the parents, number of siblings, as well as the geographical distance between children and parents.[1]

In addition to the general methodological challenge, we can distinguish studies on the general impact of caregiving on labor supply (for an overview, see Bauer and Sousa-Poza (2015) and Lilly et al. (2007)) from studies focusing on the relationship between institutions and caregiver's labor supply. The following literature review follows this distinction. The majority of studies identify a negative effect of care activities on labor participation (e.g. Bolin et al., 2008; Berecki-Gisolf et al., 2008; Carmichael and Charles, 2003; Carmichael et al., 2008; King and Pickard, 2013; Kot-

[1]For a critical discussion, see, e.g., Heitmueller (2007) and Hassink and Van den Berg (2011).

sadam, 2011; Leigh, 2010; Lilly et al., 2010; Moscarola, 2010; Nguyen and Connelly, 2014; Schneider et al., 2001; Van Houtven et al., 2013; Viitanen, 2010). However effects are usually small or modest. For example, Bolin et al. (2008) estimate for Europe that a 10% increase in time spent on caregiving can be associated with a 3.7 percentage points decrease in caregivers' employment probability. Effects are stronger when focusing on intensive care (Bauer and Sousa-Poza, 2015; Lilly et al., 2007) and on co-residential caregivers (Heitmueller, 2007). Other studies, however, cannot find any significant negative effect (e.g. Lilly et al., 2007, 2010; Meng, 2013; Stern, 1995; Wolf and Soldo, 1994). Results on the intensive margin are more consistent throughout the literature (Bauer and Sousa-Poza, 2015; Lilly et al., 2007). Caregivers are more likely to work fewer hours than non-caregivers (e.g. Berecki-Gisolf et al., 2008; Bolin et al., 2008; Carmichael et al., 2008; Ettner, 1995, 1996; Colombo et al., 2011; Johnson and Lo Sasso, 2006; Kotsadam, 2011; Lilly et al., 2007, 2010; Spiess and Schneider, 2003; Van Houtven et al., 2013). Analyzing several European countries, Bolin et al. (2008) find an average elasticity of -0.26 for working hours with respect to hours spent on caring activities.

Compared to the literature on the general relationship between caregiving and work, few studies explicitly analyze the influence of LTCI or other relevant institutions, such as e.g. early retirement schemes, on caregivers' labor supply so far. For Germany, Meng (2012) estimates the effect of caregiving on retirement decisions and finds respondents' propensity to retire to increase significantly if they are engaged in LTC. Van Houtven et al. (2013) and Jacobs et al. (2014) find similar results for the US and Canada. For elder caregivers who have difficulties reconciling work and care, early retirement programs might offer an attractive option to resolve the problematic situation permanently. Country comparisons also show that differences in institutional settings and culture do matter in this context. Controlling for country effects, Bolin et al. (2008) and Kotsadam (2011), e.g., find negative labor supply effects to be more pronounced in southern Europe than in the north. Trying

to explain these differences, Heger (2014) compares caregivers' labor supply across different institutional settings using data from the Survey for Health, Aging and Retirement in Europe (SHARE). While the effect of caregiving on labor supply is negative in countries with less developed LTC systems, the effect is insignificant in countries with a more generous LTC system.

A small number of studies refers explicitly to policy reforms or uses structural models to identify behavioral effects. Løken et al. (2014) find for Norway that daughters specifically reacted to an increase in formal care supply with reduced absence from work. For Japan, Sugawara and Nakamura (2014) show that the negative relationship between caregiving and female labor supply diminished following the introduction of the LTCI system in 2000. In contrast to Germany, the Japanese LTCI provides only formal services and no cash allowances. Thus, incentives are clearly positive for caregivers to increase labor supply.[2] For Germany, Geyer and Korfhage (2015) also found a positive labor supply effect of benefits in kind. Using data from the German SOEP for the period 2001–2010, the authors analyze the already existing LTCI using a structural labor supply model. Their extended choice model includes recipients' decision on the type of LTCI benefits they use. The authors find that benefits in kind have a small positive labor supply effect, whereas cash benefits have a non-trivial negative labor supply effect. Skira (2015) estimates a dynamic structural model for the US and, in addition to a negative labor supply effect of care allowances, finds that care leave can strengthen caregivers' labor market attachment.

[2] Campbell et al. (2010) compare the German and Japanese LTCI in more detail and discuss possible conclusions for the US.

3. Institutional setting

Germany introduced its mandatory universal-coverage social insurance program for LTC in 1995, which considerably changed the situation of people in need of care.[3] In particular, the new insurance system provides benefits exclusively on the basis of recipients' degree of care dependency, regardless of age or financial status. Prior to the reform, only very few elements of the social system directly supported people in need of LTC. In particular, LTC benefits were provided by means-tested social assistance (*Hilfe zur Pflege*), requiring recipients first to exhaust all private assets and income resources. Moreover, close family members are generally obliged to provide financial support to each other first before social assistance chips in.[4]

People with permanent (at least six months) impairments in at least two activities of daily living (ADL) and one instrumental activity of daily living (IADL) are generally eligible for benefits from LTCI (for more details, see Table 1). Depending on the degree of impairments, three care levels are distinguished (see Schulz, 2010, for more details).[5] Care recipients can generally choose between cash benefits, benefits in kind, or a mix of both types of benefits.[6]

It is important to note that coverage of LTCI comprises only part of the care

[3]The long-term care expenses are financed by income-related contributions split equally between employees and employers. Employers received a once-only compensation from the government by stopping the observance of Penance Day from a national holiday to a regular working day; except in the Federal State of Saxony. In 1995, the contribution rate was 1%, which was then raised to 1.7% in 1996 and to 1.95% in 1998.

[4]Private care insurance had been available since the mid-1980s. However, it only played a minor role and failed to reach the majority of the population (Götting et al., 1994). Cuellar and Wiener (2000) discuss lessons learned from the introduction of LTCI in Germany.

[5]Care-levels are assessed by the Medical Service of the Health Funds (MDK) or by other independent evaluators.

[6]Since July 1996, LTCI also provides benefits for nursing home care. For two reasons, however, this extension is of minor importance to our analysis. First, home care is generally preferred over nursing home care, the latter being chosen only when it has become inevitable. At the end of 1996, 75% of all recipients of LTCI benefits received care at home. Second, we only analyze multi-person households. According to Klein (1998), single person households are significantly more likely to enter nursing home care. An additional household member reduces the probability of entering a nursing home by 38%. If the person in need of LTC is married to this second household member, the effect even amounts to about 80%. Hence, we assume the introduction of nursing home care support to have only negligible effects on our estimation sample.

risk, i.e. care recipients always need a certain amount of additional help that is usually provided informally. In addition to the two benefit schemes, LTCI offers a set of other benefits to employed caregivers and to intensive caregivers. Employed caregivers can take an unpaid leave of up to six months and emergency leave for medical reasons up to ten days per year. In addition, caregivers have a right to take a leave from caregiving (respite care) of up to four weeks per year. Furthermore, there is the possibility to apply for short-term nursing home care of up to four weeks per year. All these benefits help caregivers to deal with their care obligations and should be positively related to the employment probability.[7]

Table 1: Benefits from the LTCI by care level in 1995 (monthly amounts)

	Care level		
	I	II	III
Cash benefits	205	410	665
Benefits in kind	384	921	1432
No. of beneficiaries	532,000	490,000	143,000
in %	45.6	42.1	12.3
benefits/earnings	10.2%	20.5%	33.2%
Necessary care:	Limitations in at least two ADL (personal hygiene, feeding, mobility; so called "basic care" (*Grundpflege*) and limitations in at least one IADL. Average care needed per day of at least 90 minutes. More than 45 minutes have to be necessary for basic care.	Average care needed per day of at least 180 minutes. More than 120 minutes have to be necessary for basic care.	Average care needed per day of at least 300 minutes. More than 240 minutes have to be necessary for basic care.

Note: The person in need of care can choose between the two types of benefits or a combination of both. Cash benefits are directly paid to the individual while benefits in kind are used to reimburse formal care services. Relative benefits are displayed in relation to the average gross monthly earnings (national accounts). The number of beneficiaries refers to the number of individuals in ambulant care. The total amounts of benefits remained stable until 2008 and have been raised several times since then. For example, in 2015 cash benefits in care level I amount to 244 euro per month.

During the first year, more than 1,000,000 people received benefits from the new

[7]Moreover, intensive carers receive a small amount of additional pension entitlements. Since 2008, workers in firms with more than 15 employees can request a reduction in working hours (unpaid) for a period of up to six months (renewable once); however, this is outside of our observation period.

insurance scheme (Table 1). In 1995, monthly cash benefits ranged from 205 euro (care level I) to 665 euro (care level III) and could be used to pay family caregivers. Cash benefits amounted to 10% of average gross earnings in care level I up to 33% in care level III. Cash benefits are not earmarked and spending is not monitored. When recipients choose benefits in kind, they receive formal care and the nursing service is reimbursed directly by the LTCI. In 1997, about 77% of benefit recipients received cash benefits, 7% relied on benefits in kind, and 10% combined both types of benefits (BT-Drucksache 13/9528, 1997).

4. Data, estimation sample and variables

We use data from the SOEP, a representative longitudinal study of households and individuals collected annually since 1984. The waves between 1993 and 1997 contained annual data on about 13,000 adult individuals and their children living in more than 6,500 households.[8] The SOEP data set allows for the identification of individuals in need of LTC prior to 1995, the year when LTCI was introduced. A large set of socio-economic variables can be used to control for changing group compositions (treatment and control group) over time.

Sample

The LTCI was adopted in two steps in 1995. Cash benefits for home care were introduced in January, benefits in kind six months later in July. To clearly identify decisions made before and after the reform had been fully implemented, we exclude all observations gathered in 1995 from our analysis. For the same reason, we also exclude a subsample of SOEP (i.e., Sample D, a special migration sample) surveyed for the first time in the 1995 wave. During the 1990s, the East German labor

[8] In 2013, SOEP contained data on about 24,000 individuals and their children living in more than 14,000 households. For detailed information about SOEP, see Wagner et al. (2007) and http://diw.de/en/diw_02.c.221180.en/research_data_center_soep.html.

market went through the transition process from state-planned socialism to a market economy. Hence, we have to restrict the analysis to West Germany.

To define our sample of co-residential caregivers, we limit the sample to respondents aged between 45 and 65, who have a paid job, who are currently available for work or who are retired.[9] In Germany, retirement decisions are permanent and working after retirement is very uncommon. Therefore, it could be more convenient for elder caregivers, who would prefer to reduce labor supply, to turn to early retirement and exit the labor marked permanently instead of relying on unemployment insurance (Meng, 2012; Van Houtven et al., 2013; Løken et al., 2014).[10]

All respondents in the sample live in multi-person households. Those respondents who report having a household member in need of care are assigned to the treatment group. Note that we do not observe if and to what extent they actually provide LTC. Respondents with no household member in need of any care are assigned to the control group. SOEP contains information about the receipt of LTCI benefits only since 2001. To identify the treatment group, we have to apply a more general LTC indicator to identify household members in need of LTC. We rely on a question posed to the head of the household and asking if any household member permanently receives care for reasons of old age or health impairment. We assume all households falling in this category to be affected by the LTCI reform. Note that we are unable to observe if people in the treatment group are actually receiving benefits from the insurance. The measured effect of the DiD estimator can therefore be interpreted as an intention to treat effect (ITT) instead of the average treatment effect on the treated (ATT). As the question is unrelated to the benefit receipt from the LTCI, there is no incentive for survey participants to self-select themselves into a certain group or change answering behavior after treatment.

[9] We choose this age range for two reasons. First, the majority of working caregivers is 45 years and older. In 1992, about 80% of informal caregivers were older than 45 (Schneekloth and Wahl, 2005). Second, retirement age was 65 in Germany in 1995. Since 2012, it has been lifted stepwise for new pensioners and will reach 67 in 2029.

[10] At the time the LTCI was introduced, early retirement was possible at the age of 60.

For the baseline specification we pool all observations falling in a period of two years prior to and two years after the treatment. Overall, we observe 2,523 males (2,383 females) before 1995 and 2,360 (2,312) after. Thereby, 83 males (82 females) report having a household member in need of care before treatment and are part of the pre-treatment treatment group and 67 (98) belong to the post-treatment treatment group (see Tables 2 and 3 below). As a test of robustness we extend the sample and increase the number of treated individuals significantly (see Figures 3 and 5 and Tables 6 and 9). If we include all years between 1991 and 2007, we observe 829 treated females and 830 treated males.

Variables

We measure respondents' labor market status using a binary employment indicator.[11] To avoid biases from changing group composition, we control for several covariates that might affect respondents' labor supply decision. We include age,[12] migration background, working experience, education, non-labor income, and self-reported health status of the potential caregiver.[13]

We also include a number of covariates that control for several household characteristics. To begin with, we control for household size by using a dummy variable for households with more than two people. Additional household members might influence labor supply in different ways. On the one hand, the presence of children in the household might affect how much time respondents can spend additionally for informal care. On the other hand, additional adult household members are likely to provide additional financial resources or time resources for informal care. Finally,

[11] Unfortunately, the sample is not large enough to analyse working hours of caregivers using the same identification strategy.

[12] Age enters the estimation model as a polynomial. However, changing portions of individuals who have access to early retirement programs might significantly influence our estimation results. Therefore, we add a dummy variable indicating whether a person is aged 60 or older.

[13] We use the following question to measure respondents' health status: *"How would you describe your health at present? Very good, good, satisfactory, poor, very poor."* Unfortunately, the question was not included in the 1993 wave. To generate the variable for the 1993 wave, we use respondents' answers from the previous wave or, if not available, from the following wave.

we also include community size as a control variable. The reason is that households living in small communities might have poorer access to formal care infrastructure than households residing in larger communities.

A third block of covariates controls for the amount of support actually needed by the household member who is in need of care. Differences in the amount of care regularly required are likely to influence respondents' labor supply. We include dummy variables to capture the type of support regularly needed by the household member depending on care. To build the dummy variables, we rely on a question covering four categories of impairments in activities of daily living. The four answer categories are arranged hierarchically in ascending order, which means that a person in need of major care is assumed to need the other three types of minor care as well. As a consequence, each household can only be assigned to one of the four categories (or none), and shares always add up to 100%.[14]

5. Estimation Strategy

We treat the introduction of the LTCI as a quasi-experiment and exploit the exogenous variation induced by this reform using a DiD estimation. In order to construct a counterfactual, the sample is divided into the following groups: in a first step, a control group that is not influenced by the reform is split into observations made before and after treatment (pre- and post-treatment control-group). In a second step, a treatment group that is affected by the reform is split into observations made before and after treatment as well (pre- and post-treatment treatment-group). Among the four groups, only the post-treatment treatment group is actually affected by the exogenous policy change.

In order to estimate causal effects, it is crucial for our data to meet a number

[14]The response categories are the following: *"Needs assistance with – 1. errands outside of the house; 2. running the household, preparing meals and drinks; 3. minor care, such as help with dressing himself, washing up, combing hair, shaving; 4. major care, such as getting in and out of bed."*.

of identifying assumptions, in particular the *stable unit treatment value assumption* and the *common trend assumption*.[15]

According to the *stable unit treatment value assumption*, treatment must only affect the post-treatment treatment group. Neither should the treatment group be affected by the policy reform prior to its implementation nor should the control group be affected thereby at any time by means of interactions between the members of the population. We innocuously assume the need of care to be exogenous. Accordingly, we presuppose co-residential caregivers (our treatment group) not to select themselves into the treatment group by changing households because of the reform. As a robustness check, we estimate our model only for households whose composition remained unchanged over the whole observation period. For the same reason, a small part of the control group might be affected by the reform. Since we cannot identify extra-residential carers, we have some in the control group. For two reasons we think that this is not a severe problem for our identification strategy. First, the group is small: only about 3% of all working age individuals provided extra-residential care on a regular basis (data from 2001, see Geyer and Schulz (2014)). Second, on average they provide less intensive care than co-residential carers and previous studies showed that their labor market behavior is not affected by the provision of LTC (Heitmueller, 2007). The control group is also affected by the compulsory contribution to finance the insurance. As the contribution rate of 1% at the beginning is rather low, we assume that its effect on labor supply is negligible. It is also unlikely that labor market decisions made by the treatment group are large enough to affect equilibrium wages and, therefore, the labor supply of individuals in the control group. The number of people treated seems too small to affect the entire labor market.

The *common trend* assumption implies that the potential non-treatment outcomes follow the same trend independently of group membership. That is, both the

[15]For more background information on the identifying assumptions of the DiD approach, see Lechner (2011) and Blundell and Dias (2009).

treatment- and control-groups ought to be influenced by the same macro-trends. This assumption can be relaxed if the group compositions of treatment and control group differ and if covariates that capture all variables that would otherwise lead to different time trends can be found. Then, the common-trend assumption must hold conditional on the covariates. Even though the *common trend* assumption cannot be tested, in Section 6 we use a graphical analysis to support this assumption.

If the identifying assumptions hold, the treatment effect can be estimated in a regression framework. Thereby, biases resulting from permanent differences between the treatment and control groups, as well as biases resulting from macro trends that are unrelated to the change in policy regulation, are removed (Imbens and Wooldridge, 2009). To do so, dummy variables are constructed to indicate group membership. $Tr_i \in \{0, 1\}$ indicates whether individual i belongs to the treatment group ($Tr = 1$) and $Post_t \in \{0, 1\}$ indicates whether observations are made after treatment has occurred ($Post_t = 1$).

Our estimation strategy can be summarized in the following equation:

$$y_{it} = \alpha + \beta(Tr_i \times Post_t) + \lambda Tr_i + \delta Post_t + \mathbf{X_{it}'}\gamma + e_{it}, \tag{1}$$

where y_{it} measures the employment status of individual i at time t, α is a constant, $\mathbf{X_{it}}$ is a vector of covariates, and e_{it} represents the error-term. The coefficients β, λ, δ, and γ are to be estimated. The coefficient of interest, β, captures the causal labor supply effect of the LTCI.

When estimating binary dependent variables, generally, non-linear models such as *probit* or *tobit* come to mind. However, in a DiD framework these models cannot be applied without further assumptions. As Lechner (2011) shows, the common-trend assumption only holds in a non-linear specification, if there is no group specific difference in the dependent variable. This means that the treatment and control groups ought to start at the same average initial levels of labor market participation before treatment. As this assumption is not valid in our case, we continue with a

linear specification. We analyze the labor supply behavior separately for men and women.

6. Descriptives and discussion of identifying assumptions

In this section, we provide descriptive information on all groups we compare in our econometric model. Moreover, we show how LTCI and caregivers' labor supply decisions might interact. Finally, we discuss whether all necessary identifying assumptions are met.

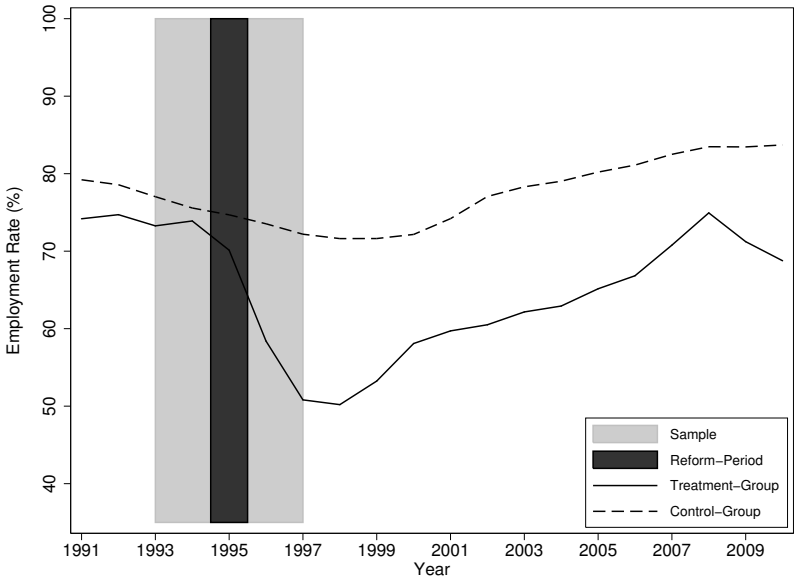
Graphical illustration

To begin with, Fig. 1 (men) and Fig. 2 (women) report employment rates over time for each group, covering the time period from 1991–2010. Trends are plotted separately for men and women. The dashed line represents the control group; the treatment group is indicated by the continuous line. The gray background indicates the time period covered by the main econometric specification that we present in Section 7.[16]

Fig. 1 shows labor market participation rates of male respondents. In the control group, participation rates range between about 75% and 85% over the whole observation period. Due to the smaller sample size, volatility is higher in the treatment group. Within this group, employment rates are constantly lower as compared to

[16] Note that the survey questions used to collect information on the type of care household members regularly need were changed after 1990, then using a broader definition of care. As a consequence, comparisons of pre-1991 findings with later results can be misleading, and our graphs only display developments from 1991–2010. While the general question remained unchanged (*"Is there anyone in your household who is receiving care because of old age or health reasons?"*), response categories were changed in the 1991 wave. Prior to 1991, LTC was classified into *(1) "bed-ridden"* and *(2) "not bed-ridden, but in need of help with daily domestic tasks in the household"*. Since 1991, a broader definition of care was applied. For the 1991 and following waves, we can distinguish help with *(1) "running errands outside the house, (2) running the household, meals and drinks, (3) simple tasks, e.g. help with dressing, washing, etc., and (4) complex tasks, e.g. moving from the bed, bowel movements, etc."*.

Fig. 1: Employment rates of treatment and control group between 1991 and 2010 (men)



Note: The dashed line and the solid line represent the control and the treatment group, respectively. The dark-gray background indicates the time period that is used to estimate the DiD models in Section 7. To smooth out fluctuations due to the small sample we used a moving average: $y = 0.5(x_t + 0.5x_{t-1} + 0.5x_{t+1})$.
Source: SOEPv30, own calculations.

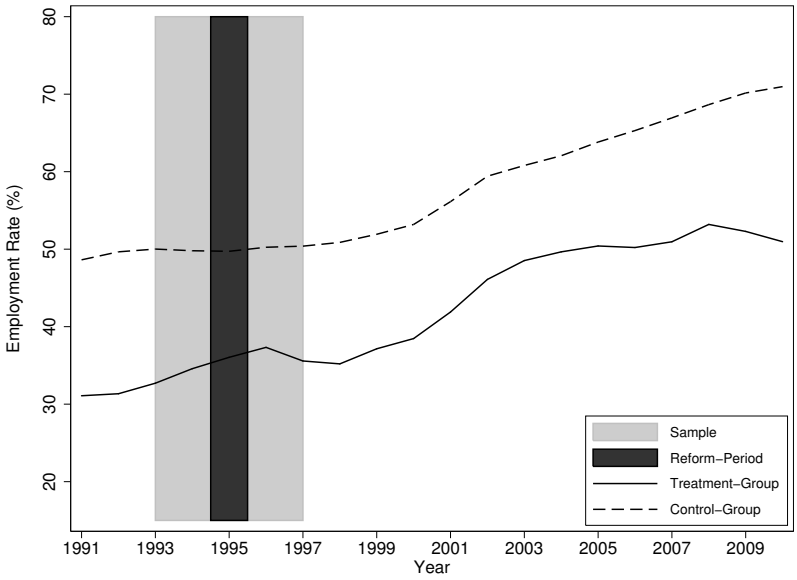
the control group, although the magnitude of the gap varies over time. It was smaller prior to the LTCI reform than after its implementation. The gap has remained relatively constant since 1997. Results are in line with our key assumption that the introduction of LTCI would affected labor supply. Fig. 1 suggests an immediate effect that lasted for several years after the LTCI reform had been implemented in 1995. Moreover, trends were almost parallel in the treatment group and control group from 1991-1995, which supports the *common trend assumption*. The slight drop of the employment rate within the treatment group at the end of the observation period might indicate the impact of the first *extensive* reform of the LTCI since 1995. The reform, which not only increased financial support but also provided more generous professional assistance to informal caregivers, came into force in July 2008.[17]

Fig. 2 displays female employment rates for the time period from 1991–2010. Within the control group, employment rates steadily increased from about 50% to 70% between 1991 and 2010. Even though the employment rate was 20 percentage points lower in the treatment group in 1991, it then constantly developed parallel to the control group's. In contrast to our male subsample, we do not find any visual indication that the introduction of LTCI in 1995 affected women's labor supply.

In sum, our graphical analysis suggests male and female caregivers to respond differently to the LTC reform. While we can observe a decrease in labor supply for men, no visible effect can be found for women. The difference might be explained by the generally low employment rate of female caregivers. When the LTCI was introduced their employment rate was only about 30%, which limits further labor supply reductions. However, reported averages are unconditional on covariates and not weighted. Hence, also the observed drop in male labor supply might vanish in the econometric model that includes the covariates introduced in Section 4.

[17] The so-called *Pflegeweiterentwicklungsgesetz*. For details see e.g. Kostorz et al. (2010).

Fig. 2: Employment rates of treatment and control group between 1991 and 2010 (women)



Note: The dashed line and the solid line represent the control and the treatment group, respectively. The dark-gray background indicates the time period that is used to estimate the DiD models in Section 7. To smooth out fluctuations due to the small sample we used a moving average: $y = 0.5(x_t + 0.5x_{t-1} + 0.5x_{t+1})$.
Source: SOEPv30, own calculations.

Summary statistics

To get a more detailed view of group compositions and to discuss the model assumptions, in Tables 2 and 3, we present the main summary statistics of treatment and control groups before and after treatment for men and women.

Table 2 provides summary statistics for men. The category *before treatment* refers to the years 1993 and 1994, and the category *after treatment* comprises the years 1996 and 1997. While remaining almost constant in the control group, employment clearly dropped from 72% to 49% in the treatment group. As previously discussed, a part of that reduction seems to be driven by individuals who choose to leave the labor market, turning to early retirement. The fraction of new retirees increases from 16% to 30% in the treatment group after the reform, while it stays almost constant at 17% in the control group. When we try to estimate causal effects of the reform, it is important to avoid biases through changing group composition. Thus, in regard to pensioners, we check for changing age structure in the sample. In Table 2, it can be seen that the share of men aged 60 or older increases from 25% in the *pre-treatment treatment group* to 34% after treatment. However this change does not prove to be statistically significant at the 5% level.

The availability of formal care services most likely depends on the size of the community where a household lives. As Table 2 documents, the fraction of households living in large communities (>100,000 inhabitants) remains constant at around 25% in the treatment group before and after treatment. However, a shift can be observed in small (<20,000 inhabitants) and medium sized communities (20,000–100,000 inhabitants) when comparing the treatment group before and after treatment. Before treatment, 30% of all households in the treatment group live in small communities compared to 15% after treatment. Thus, to avoid a downward bias of the estimated reform effect, community size has to be controlled for.

Further changes in the group composition of the treatment group can be observed with regard to the amount of care needed. Before treatment, 35% of all household

Table 2: Descriptive statistics for treatment and control group two years before and after treatment (males)

Variables	Treatment group		Control group	
	before 1995	after 1995	before 1995	after 1995
Employed	0.72	0.49*	0.75	0.71*
Working hours	45.65	43.80	42.91	42.93
Retired	0.16	0.30*	0.16	0.17
Age	55.47	56.43	53.93	54.09
Age≥60	0.25	0.34	0.19	0.20
Migration background	0.20	0.25	0.35	0.32*
Working experience in years	32.43	33.24	31.71	31.65
Years of education	11.17	10.84	10.99	11.21*
Health status:				
good – very good	0.35	0.34	0.44	0.40*
satisfying	0.39	0.39	0.34	0.36
poor – very poor	0.27	0.27	0.22	0.24
Married	0.95	0.93	0.89	0.86*
Other household income /1000	15.07	16.81	23.74	22.01*
Household size	3.55	3.57	2.91	2.90
Community size:				
<20,000	0.30	0.15*	0.12	0.13
20,000–100,000	0.45	0.57	0.53	0.56
>100,000	0.25	0.28	0.34	0.32*
Care spell in years	3.69	4.49		
HH-member needs help with:				
no help or not known	0.01	0.03		
getting around outside the house	0.11	0.09		
household chores, preparing meals	0.16	0.12		
washing, dressing, etc.	0.37	0.30		
getting into and out-of-bed, etc.	0.35	0.46		
Observations	83	67	2,523	2,360

Note: Reported means are not weighted. Working hours are only reported for working individuals. The health status is self-reported. Other household income is the sum of gross annual income other than the persons own labor income, its own retirement benefits and the household's benefits from the LTCI. It is reported in 1000 euro and is inflation adjusted (base year = 2006). * indicates if differences in means before and after treatment are statistically significant at the 5% level.

Source: SOEPv30, own calculation

members in the treatment group rely on major help (with the highest category being "getting into and out of bed, bowel movement etc."). After treatment, their share amounts to 46%. Obviously, this change in group composition is likely to drive the demand for formal and informal care respectively. Although t-tests for the differences turn out not to be significant at the 5%-level, we add control variables to capture this change. Further control variables are migration background, whose share increased from 20% in the pre-treatment treatment period to 35% in the post-treatment treatment period, and other household income that increases from about 15,000 euro to nearly 17,000 euro when comparing the pre-treatment and post-treatment treatment group.

Table 3 provides summary statistics for women and reveals a different picture. Employment rates remain relatively constant in both groups over time. At the same time, however, the employment rate is considerably higher in the control group (49%) than in the treatment group (around 35%). Interestingly, the fraction of retirees increases in the treatment group after the reform. However, different from the male subsample, this change occurs even though the share of women aged 60 years and above decreases from 35% to 30% over the reform period. In general, descriptive statistics reveal less variation between groups in the female subsample as compared to the male subsample. Yet, women in the treatment group are older on average than women in the control group – in particular the fraction of individuals above 60 years being larger. Also, working experience is about two years longer in the *post-treatment treatment group* than it is before treatment and the portion of household members needing major care (getting into and out-of-bed, etc.) increases from 27% to 39%.

In summary, we find employment rates of treatment and control group to be relatively parallel before treatment. Male employment rates drop significantly after 1995. This drop might be related to changing group composition. However, most differences between treatment and control group turn out to be insignificant.

Table 3: Descriptive statistics for treatment and control group two years before and after the reform (females)

Variables	Treatment group		Control group	
	before 1995	after 1995	before 1995	after 1995
Employed	0.35	0.36	0.49	0.49
Working hours	32.07	30.05	31.27	30.78
Retired	0.12	0.18	0.14	0.15
Age	56.45	55.33	53.92	53.83
Age≥60	0.35	0.30	0.20	0.20
Migration background	0.24	0.28	0.31	0.30
Working experience in years	16.65	19.07	19.54	20.39*
Years of education	9.87	9.84	10.17	10.36*
Health status:				
good – very good	0.26	0.30	0.36	0.33*
satisfying	0.46	0.39	0.39	0.42
poor – very poor	0.28	0.32	0.25	0.25
Married	0.87	0.85	0.81	0.79
Other household income /1000	33.95	29.93	37.70	35.69
Household size	3.23	3.28	2.66	2.64
Community size:				
<20,000	0.23	0.24	0.12	0.12
20,000–100,000	0.50	0.48	0.53	0.55
>100,000	0.27	0.28	0.36	0.32
Care spell in years	3.67	3.52		
HH-member needs help with:				
no help or not known	0.02	0.01		
getting around outside the house	0.11	0.08		
household chores, preparing meals	0.20	0.13		
washing, dressing, etc.	0.40	0.40		
getting into and out-of-bed, etc.	0.27	0.39		
Observations	82	88	2,383	2,312

Note: Reported means are not weighted. Working hours are only reported for working individuals. The health status is self-reported. Other household income is the sum of gross annual income other than the persons own labor income, its own retirement benefits and the household's benefits from the LTCI. It is reported in 1000 euro and is inflation adjusted (base year = 2006). * indicates if differences in means before and after treatment are statistically significant at the 5% level.

Source: SOEPv30, own calculation

Furthermore, it cannot be related to other policy reforms, since in the years of our analysis no major additional reform was implemented that could yield different macro trends for treatment and control group. Therefore in the following section we test the treatment effect in a regression framework that takes group characteristics into account.

7. Results

Regression results are presented in Tables 4, 5. All models are estimated using ordinary least squares (OLS). Standard errors are clustered on household level. We report five models for each group and dependent variable using different sets of control variables.

Table 4 presents the results of the regression DiD estimates of the long-term care reform effect on male employment status. Of particular interest is the interaction term $Post95 \times Tr$ measuring the treatment effect. For all models, its coefficient is negative and statistically significant at 5% or at 10%. Model 1 suggests a reduction of 19.3 percentage points in male labor supply by the introduction of LTCI. The regression framework shows that the large estimate – in relative terms it means a reduction in employment by about 30% – is very imprecise with a standard error of 0.082. Nonetheless, the point estimate remains remarkably stable when introducing more controls. It is slightly reduced to 14.8 percentage points but still significant when we introduce a dummy for all carers above the age of 59 (Model 2). The estimated coefficient does not change considerably when introducing further controls. In the full specification (Model 5), which also controls for community size and level of impairments, the point estimate of the interaction is at 14.6 percentage points with a standard error of 6.8 percentage points.

The coefficient of Tr indicates that initial differences between treatment and control group are not significant in any of the models. No matter what, $Post95$ is always

Table 4: Regression DiD estimates of long-term care reform effects on male employment

	(1)	(2)	(3)	(4)	(5)
Post95	-0.035**	−0.026*	-0.027**	−0.025*	-0.027**
	(0.012)	(0.010)	(0.010)	(0.010)	(0.010)
Tr	-0.026	0.022	0.030	0.029	-0.072
	(0.063)	(0.059)	(0.056)	(0.055)	(0.105)
Post95× Tr	-0.196*	−0.153*	-0.145†	−0.151*	-0.162*
	(0.086)	(0.077)	(0.075)	(0.074)	(0.072)
Age[(a)]		✓	✓	✓	✓
Individual characteristics[(b)]			✓	✓	✓
Household characteristics[(c)]				✓	✓
Care variables[(d)]					✓
Obs.	5033	5033	5033	5033	5033
Obs. in treatment group	150	150	150	150	150
R^2	0.01	0.26	0.36	0.36	0.37

Note: Reported values denote estimated coefficients. Standard errors are clustered on household level and reported in parentheses. Significance levels: † p <0.10, ∗ p <0.05, ∗∗ p <0.01 , ∗ ∗ ∗ p <0.001.

Column (1) presents the results from a regression DiD estimation without further covariates. Columns (2) to (5) include further control variables: (a) Caregiver's age: age, age squared and a dummy variable indicating whether a person is aged 60 or older. (b) Individual characteristics: migration background, working experience, education, and self-reported health status. (c) Houeshold characteristics: household size, community size, and other household income. (d) Care variables: dummies for type of care needed by the care recipient.

Source: SOEPv30, own calculations.

negative and significant at 1% or 5%, respectively.[18]

While we find a robust and large negative labor supply effect for men, we do not find any significant effect for women (Table 8). The coefficient of $Post95 \times Tr$ is never statistically significant. Model (1) again resembles the graphical analysis, which showed no indication of an effect of the LTC reform. The descriptives showed that women in the treatment group work less than women in the control group throughout the observation period. The coefficient of Tr is significantly negative only in model (1). The significant effect disappears as soon as Age is included into the model. This indicates that the unconditional differences are mainly driven by different age structures.[19]

In sum, even when controlling for various covariates, the negative effect on male employment remains significant. Women seem not to react to the reform. In order to test whether this effect is robust we conducted a series of tests of our model, which are documented in the next section.

[18] Most of the common covariates to explain labor supply have the expected signs and most are significant. The full estimation results are documented in the Appendix in Table 7.

[19] The full estimation results are reported in Table 8 in the Appendix.

Table 5: Regression DiD estimates of long-term care reform effects on female employment

	(1)	(2)	(3)	(4)	(5)
Post95	0.001	0.001	-0.017	−0.017	-0.016
	(0.013)	(0.011)	(0.011)	(0.011)	(0.011)
Tr	-0.136*	−0.051	0.025	0.030	0.050
	(0.063)	(0.058)	(0.054)	(0.054)	(0.099)
Post95× Tr	0.009	−0.026	-0.054	−0.055	-0.066
	(0.073)	(0.064)	(0.063)	(0.063)	(0.064)
Age[(a)]		✓	✓	✓	✓
Indivividual charcteristics[(b)]			✓	✓	✓
Household characteristics[(c)]				✓	✓
Care variables[(d)]					✓
Obs.	4865	4865	4865	4865	4865
Obs. in treatment group	170	170	170	170	170
R^2	0.00	0.16	0.36	0.36	0.36

Note: Reported values denote estimated coefficients. Standard errors are clustered on household level and reported in parentheses. Significance levels: † p <0.10, ∗ p <0.05, ∗∗ p <0.01, ∗ ∗ ∗ p <0.001.

Column (1) presents the results from a regression DiD estimation without further covariates. Columns (2) to (5) include further control variables: (a) Caregiver's age: age, age squared and a dummy variable indicating whether a person is aged 60 or older. (b) Individual characteristics: migration background, working experience, education, and self-reported health status. (c) Houeshold characteristics: household size, community size, and other household income. (d) Care variables: dummies for type of care needed by the care recipient.

Source: SOEPv30, own calculations.

8. Sensitivity tests

In this section, we test whether regression results are robust to variations in the estimation sample and to different model specifications. Moreover, we conduct a series of placebo regressions. We only discuss results for the male subsample. Results for the female subsample are documented in the Appendix A.1.

One concern is that the introduction of the LTCI could have influenced the decision to move between households (or nursing homes and households) in order to provide family care. If this choice was related to the labor supply decision, we would yield biased estimates. Accordingly, we re-estimate the model for a reduced sample omitting all households whose household composition changed during the observation period (i.e., between 1993 and 1997). Results are documented in column (1) of Table 6. We find a slightly larger negative point estimate than in the baseline model of Table 4 (column (5)). Although the sample gets very small in this specification the effect remains significant at the 10% level.

The common trend assumption is generally not testable but becomes more plausible if treatment and control group have similar characteristics. We use propensity score matching in order to improve the balance between treatment and control group (column (2) of Table 6). We perform five-to-one nearest neighbor matching on the probability of belonging to the treatment group *before* the LTCI was introduced in 1995. Matching covariates include the majority of control variables used in the baseline estimation and are listed in the table notes. The outcome variable that measures employment status is also included. This procedure guarantees that the comparison group is very similar in all observed characteristics before treatment – including the labor supply decision. We balance the panel using only individuals in the post-treatment period who also belong to the matched pre-treatment sample. Note that the matching procedure reduces the number of observations considerably. The point estimate is significant at the 10% level and its magnitude (-16.5 percentage points) is almost identical to the baseline specification.

Table 6: Robustness tests for the treatment effect on male employment

	(1) Constant HH	(2) DiD Matching	(3) Linear Time-Trend
Post95	-0.028*	0.007	
	(0.013)	(0.043)	
Tr	-0.095	0.141	-19.339
	(0.135)	(0.116)	(14.278)
Post95× Tr	-0.181†	−0.165†	-0.195*
	(0.098)	(0.098)	(0.080)
Tr× Year			0.010
			(0.007)
Year			0.002*
			(0.001)
Age variables	✓	✓	✓
Indivividual characteristics	✓	✓	✓
Household characteristics	✓	✓	✓
Care variables	✓	✓	✓
Obs.	3,030	322	28,737
Obs. in Tr	73	67	830
Years in Sample	1993–1997	1994–1996	1991–2007

Note: Reported values denote estimated coefficients. Standard errors are clustered on household level and reported in parentheses. Significance levels: † p <0.10, ∗ p <0.05, ∗∗ p <0.01, ∗ ∗ ∗ p <0.001.

(1) This estimation includes only households whose composition does not change during the observation period. (2) This model is based on a propensity score matching in order to improve the balance between treatment and control group. In the pre-treatment year 1994 we estimate a probit model on the treatment dummy. Thereby, the explanatory variables include the employment decision, retirement status, age, migration background, working experience, years of education, health status, marital status, household size, and community size. We use five-to-one nearest neighbor matching with replacement. In order to impose common support we drop treatment observations whose propensity score is higher than the maximum or less than the minimum propensity score of the controls. If individuals of the treatment group drop out after treatment, their matching partners are dropped as well. (3) This model uses data from the maximum time window of 16 years and includes an interaction of a linear time trend with the treatment dummy (Tr × Year).

Source: SOEPv30, own calculation.

In the third model (column (3)) we use all waves from 1991 through 2007 and interact a linear time trend with the treatment dummy. This robustness check allows treatment and control group to follow a different time trend (Angrist and Pischke, 2008). It is reassuring if the treatment effect is unchanged by the inclusion of this trend. Our estimated treatment effect remains almost stable; it is only slightly larger than in the original model and significant at the 5% level.

In summary, the three models address potential concerns regarding our identifying assumptions and support our main results as they prove to be stable in all specifications.

Time span

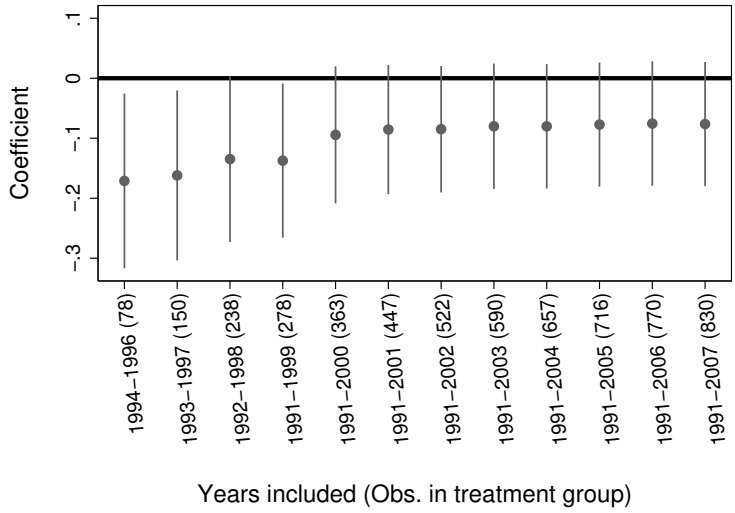
In a next step we extend the sample by including more time spans before and after treatment. We can include waves from 1991 to 2007. We start with the smallest sample and include only waves 1994 and 1996. Then we proceed and increase the time window of observations.

Fig. 3 reports results for the male sample. Dots represent point estimates controlling for the full vector of covariates. Vertical lines represent the 95% confidence interval. The treatment effect on employment is stable with time-spans of one to up to four years before and after treatment. The more we extend the post treatment period, the smaller does the estimated effect get. When we extend the sample to include more than 5 period after treatment we yield marginally insignificant point estimates in large samples (when we control for the full set of covariates).[20]

Labor supply effects of the LTCI introduction are for several reasons likely to change over time. Firstly, the new LTCI lead to an increase of ambulatory home care services in the years after 1995. Secondly, the level of LTC benefits remained unchanged until 2008. Consequently, the real value of the benefits continuously decreased over the years. For instance, monthly benefits in cash of 205 euro in care

[20] All estimates are statistically significant without covariates; results are available upon request from the authors.

Fig. 3: Estimated treatment effect using different time spans (men)



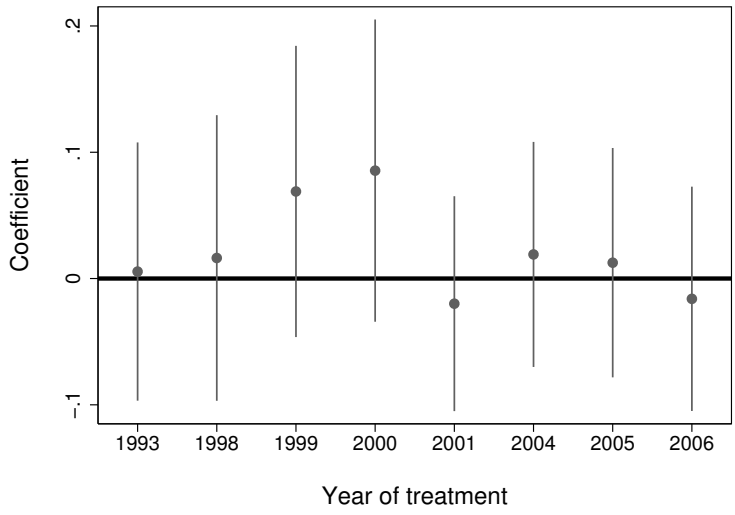
Note: Dots represent point estimates controlling for the full vector of controls. Vertical lines represent the 95% confidence interval. The year 1995 is omitted in all models. Standard errors are clustered on household level.
Source: SOEPv30, own calculations.

level I decreased by approximately 35 euro in purchasing power from 1995 through 2007. Hence, the income effect of the LTCI decreased as well. Thirdly, Germany introduced deductions for early retirement starting with cohort 1937, which made it less attractive to leave the labor market.

Placebo regressions

The available data allow only one placebo regression in the pre-treatment period. All other placebo regressions are on post-treatment samples. All regressions are performed with assumed pre- and post-treatment periods covering a time-span of two years. We stop the placebo regressions after 2006, because the first fundamental reform of the LTCI (*Pflegeweiterentwicklungsgesetz*) was introduced in 2008 (Kostorz et al., 2010) and, hence, estimates could pick up this reform. Furthermore, we do not report regressions that include the period of the LTCI reform in 1995. Consequently, the years 1994–1997 are not reported.

Fig. 4: Placebo Regressions (male)



Note: Dots represent point estimates controlling for the full vector of controls. Vertical lines represent the 95% confidence interval. Each model uses observations of two years before and after the assumed year of treatment. The year reported on the x-axis is always the first year of treatment. Standard errors are clustered on household level.
Source: SOEPv30, own calculations.

Fig. 4 summarizes placebo regressions on male employment status. We do not find any significant effects in any of the placebo specifications for men.

Care provision / care source

The labor supply results we find are remarkably stable even though the sample of treated individuals is relatively small. It is neither driven by sample characteristics nor by changing group composition after 1995. However, one important aspect is not directly observable in our data set: the care provision. That is, we have to assume that the co-residential household members are caregivers. Unfortunately there is no other data source on informal care provision from that time span. We cannot test this assumption for the time of the reform. SOEP added questions on the sources of care from 1997. If we use the same sample definition as in our empirical analysis, about 97% of all persons in need of care report to receive care from within the household.

9. Discussion of results

The aim of our empirical analysis is to use the introduction of the LTCI in Germany to study the effect of long-term care policies on the labor supply of informal caregivers. LTCI generally provides various measures to support caring relatives. In Germany, recipients can choose both benefits in kind and cash benefits as well as a combination of the two, leaving the choice to the household. While cash benefits increase household non working income, benefits in kind provide a substitute for informal care. Accordingly, the newly introduced LTCI provides incentives for both the reduction as well as the extension of labor supply. However, surveys also show that many family carers would be likely to provide care even in the absence of cash benefits. For this group the LTCI increases household income but might not change behavior (a point, also raised by Campbell et al., 2010).

The insurance does not cover all care needs; a certain amount of additional informal care and/or co-payment is always needed. We cannot observe the actual choice of benefits by the household but we focus only on multi-person households for which it is reasonable to assume that most of them choose benefits in cash – in particular after the introduction of the LTCI. Surveys show that people prefer care by family members such as their spouses or children over formal care services (Schupp and Künemund, 2004). Studies also show that marital status is a strong predictor of having a family carer. Himes et al. (2001, 2000), e.g., find that being married is associated with less reliance on formal services and a greater likelihood of receiving care from family members.

Our results suggest that male and female co-residential caregivers reacted differently to the introduction of the LTCI. The estimates show a negative effect on labor supply for male but not for female caregivers. Female caregivers already had a very low employment rate before the LTCI came into effect, and average female employment rates were also low. When the LTCI was introduced there was little scope for them to adjust labor supply. It is plausible to assume that this group would have provided care even in the absence of the LTCI. The situation was different for men: the employment rate of male caregivers was about as high as the average employment rate before the reform. The cash benefit – at the time of the introduction of the LTCI – was relatively high (see Table 1) and replaced up to 33% of average gross earnings. Thus men had strong additional incentives to reduce labor supply. It is plausible that one important channel – at least for men near the age of 60 and older – were early retirement programs. In principle people could retire even before 60 because the unemployment insurance provided benefits for up to 32 months for elder unemployed. And it was possible to retire without or with very low actuarial deductions starting at the age of 60.

10. Conclusion

LTC systems rely to a large extent on informal care by family members and friends. This is often viewed as a cost-saving alternative to formal care arrangements. However, a considerable share of informal caregivers is of working age and has to reconcile care obligations and market work. The empirical literature suggests that, at a minimum intensive long-term care is associated with a substantial reduction of labor supply, wage penalties, higher poverty risks and detrimental health effects. As the demand for LTC is growing various countries introduced policies to support family care arrangements.

In this paper we provide causal evidence of the labor supply reaction of family caregivers to the introduction of the LTCI in Germany. This large reform generated quasi-experimental variation that we exploit in a DiD-framework. We assign working age individuals from households with a person in need of care to the treatment group. Respondents with no household member in need of any care are assigned to the control group. We compare labor supply before and after treatment had occurred. We find that while the insurance did not have a significant effect on the labor supply of women, a negative effect can be found on the labor supply of men. The effect turns out to be robust in various specifications. The point estimate seems to be rather large but very imprecisely estimated due to the small sample size and our estimation approach.

One goal of the LTCI was to improve the availability of home based care and allow family members to care for their relatives (BMG, 2007). An immediate consequence was an increase in resources to organize care at home. Given data limitations, we can only provide indirect evidence for the effect of the LTCI on the provision of informal care. At least for men, our results suggest that this goal seems to be met, as the insurance has large effect on their labor supply. We do not however have data on care provision so it remains an assumption that men increase time devoted to care when they reduce labor supply.

The results reveal a trade off for policy makers that is important for future reforms – in particular for countries that base their LTC system mainly on informal care. An increasing number of male caregivers in the future raises the question if they are able to combine care obligations and market work. The same will be relevant for female cohorts as they continue to increase labor market participation. This is particularly true if the decision to provide care is linked to the retirement decision. A possibility to support working family carers are policies that help to combine caregiving and market work, e.g. care leave and respite care. In Germany, these policies are already in place but take-up rates are very low. A possible reason is that theses measures are – financially – less attractive than comparable policies for young families with children.

References

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Bauer, J. M. and Sousa-Poza, A. (2015). Impacts of informal caregiving on caregiver employment, health, and family. *Journal of Population Ageing*, 8(3):113–145.

Berecki-Gisolf, J., Lucke, J., Hockey, R., and Dobson, A. (2008). Transitions into informal caregiving and out of paid employment of women in their 50s. *Social Science & Medicine*, 67(1):122 – 127.

Blundell, R. and Dias, M. C. (2009). Alternative approaches to evaluation in empirical microeconomics. *Journal of Human Resources*, 44(3):565–640.

BMG, Bundesministerium für Gesundheit (German Federal Ministry of Health) (2007). Erster Bericht über die Entwicklung der Pflegeversicherung seit ihrer Einführung am 01. Januar 1995.

Bolin, K., Lindgren, B., and Lundborg, P. (2008). Your next of kin or your own career?: Caring and working among the 50+ of europe. *Journal of Health Economics*, 27(3):718–738.

BT-Drucksache 13/9528 (1997). Erster Bericht über die Pflegeversicherung seit ihrer Einführung am 01.Januar 1995.

Campbell, J. C., Ikegami, N., and Gibson, M. J. (2010). Lessons from public long-term care insurance in Germany and Japan. *Health Affairs*, 29(1):87–95.

Carmichael, F. and Charles, S. (2003). Benefit payments, informal care and female labour supply. *Applied Economics Letters*, 10(7):411–415.

Carmichael, F., Hulme, C., Sheppard, S., and Connell, G. (2008). Work – life imbalance: Informal care and paid employment in the UK. *Feminist Economics*, 14(2):3–35.

Colombo, F., Llena-Nozal, A., Mercier, J., and Tjadens, F. (2011). *Help Wanted?: Providing and Paying for Long-Term Care*. OECD health policy studies. OECD Publishing, Paris.

Cuellar, A. E. and Wiener, J. M. (2000). Can social insurance for long-term care work? The experience of Germany. *Health Affairs*, 19(3):8–25.

Ettner, S. L. (1995). The impact of "Parent care" on female labor supply decisions. *Demography*, 32(1):63–80.

Ettner, S. L. (1996). The opportunity costs of elder care. *The Journal of Human Resources*, 31(1):189–205.

Geyer, J. and Korfhage, T. (2015). Long-term care insurance and carers' labor supply - a structural model. *Health Econ.*, 24(9):1178–1191.

Geyer, J. and Schulz, E. (2014). Who cares? Die Bedeutung der informellen Pflege durch Erwerbstätige in Deutschland. *DIW Wochenbericht*, 81(14):294–301.

Götting, U., Haug, K., Hinrichs, K., Götting, U., Haug, K., and Hinrichs, K. (1994). The long road to long-term care insurance in germany. *Journal of Public Policy*, 14(03):285.

Hassink, W. H. and Van den Berg, B. (2011). Time-bound opportunity costs of informal care: Consequences for access to professional care, caregiver support, and labour supply estimates. *Social Science & Medicine*, 73(10):1508–1516.

Heger, D. (2014). Work and well-being of informal caregivers in Europe. *Ruhr Economic Papers*, 512.

Heitmueller, A. (2007). The chicken or the egg?: Endogeneity in labour market participation of informal carers in England. *Journal of Health Economics*, 26(3):536–559.

Himes, C. L., Schneider, U., and Wolf, D. A. (2001). The dynamics of long-term care service use in Germany. *Vierteljahrshefte zur Wirtschaftsforschung*, 70(1):153–158.

Himes, C. L., Wagner, G. G., Wolf, D. A., Aykan, H., and Dougherty, D. D. (2000). Nursing home entry in Germany and the United States. *Journal of cross-cultural gerontology*, 15(2):99–118.

Imbens, G. W. and Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1):5–86.

Jacobs, J. C., Laporte, A., Houtven, C. H. V., and Coyte, P. C. (2014). Caregiving intensity and retirement status in canada. *Social Science & Medicine*, 102:74–82.

Johnson, R. W. and Lo Sasso, A. T. (2006). The impact of elder care on women's labor supply. *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, 43(3):195–210.

King, D. and Pickard, L. (2013). When is a carer's employment at risk? longitudinal analysis of unpaid care and employment in midlife in england. *Health & Social Care in the Community*, 21(3):303–314.

Klein, T. (1998). Der Heimeintritt alter Menschen und Chancen seiner Vermeidung: Ergebnisse einer Repräsentativerhebung in den Einrichtungen der stationären Altenhilfe. *Zeitschrift für Gerontologie und Geriatrie*, 31(6):407–416.

Kostorz, P., Aufderhorst, U., Fecke, M., Hesa, M., Mört, D., Wübbeler, M., and Ziegler, S. (2010). Das Pflegeweiterentwicklungsgesetz und seine praktische Umsetzung. *Wege zur Sozialversicherung – Zeitschrift für die Sozialversicherungs-Praxis*, 65(6):179–188.

Kotsadam, A. (2011). Does informal eldercare impede women's employment? the case of european welfare states. *Feminist Economics*, 17(2):121–144.

Lechner, M. (2011). The estimation of causal effects by difference-in-difference methods. *University of St. Gallen, Discussion Paper*, 2010-28.

Leigh, A. (2010). Informal care and labor market participation. *Labour Economics*, 17(1):140–149.

Lilly, M. B., Laporte, A., and Coyte, P. C. (2007). Labor market work and home care's unpaid caregivers: A systematic review of labor force participation rates, predictors of labor market withdrawal, and hours of work. *Milbank Quarterly*, 85(4):641–690.

Lilly, M. B., Laporte, A., and Coyte, P. C. (2010). Do they care too much to work? the influence of caregiving intensity on the labour force participation of unpaid caregivers in canada. *Journal of Health Economics*, 29(6):895–903.

Løken, K. V., Lundberg, S. J., and Riise, J. (2014). Lifting the burden: State care of the elderly and labor supply of adult children. Technical report.

Meng, A. (2012). Informal caregiving and the retirement decision. *German Economic Review*, 13(3):307–330.

Meng, A. (2013). Informal home care and labor-force participation of household members. *Empirical Economics*, 44(2):959–979.

Moscarola, F. C. (2010). Informal caregiving and women's work choices: Lessons from the netherlands. *LABOUR*, 24(1):93–105.

Nguyen, H. T. and Connelly, L. B. (2014). The effect of unpaid caregiving intensity on labour force participation: Results from a multinomial endogenous treatment model. *Social Science & Medicine*, 100:115–122.

Schneekloth, U. and Wahl, H. W. (2005). Möglichkeiten und Grenzen selbständiger Lebensführung in privaten Haushalten (MuG III). Integrierter Abschlussbericht im Auftrag des Bundesministeriums für Familie, Senioren, Frauen und Jugend.

Schneider, T., Drobnic, S., and Blossfeld, H.-P. (2001). Pflegebedürftige Personen im Haushalt und das Erwerbsverhalten verheirateter Frauen. *Zeitschrift für Soziologie*, 30(5):362–383.

Schulz, E. (2010). The long-term care system in Germany. Discussion Papers of DIW Berlin 1039.

Schupp, J. and Künemund, H. (2004). Private Versorgung und Betreuung von Pflegebedürftigen in Deutschland. *DIW Wochenbericht*, 71(20):289–294.

Skira, M. M. (2015). Dynamic wage and employment effects of elder parent care: Dynamic effects of caregiving. *International Economic Review*, 56(1):63–93.

Spiess, C. K. and Schneider, A. U. (2003). Interactions between care-giving and paid work hours among european midlife women, 1994 to 1996. *Ageing and Society*, 23(1):41–68.

Stern, S. (1995). Estimating family long-term care decisions in the presence of endogenous child characteristics. *Journal of Human Resources*, 30(3):551–580.

Sugawara, S. and Nakamura, J. (2014). Can formal elderly care stimulate female labor supply? The Japanese experience. *Journal of the Japanese and International Economies*, 34(C):98–115.

Van Houtven, C. H., Coe, N. B., and Skira, M. M. (2013). The effect of informal care on work and wages. *Journal of Health Economics*, 32(1):240–252.

Viitanen, T. K. (2010). Informal eldercare across europe: Estimates from the european community household panel. *Economic Analysis and Policy*, 40(2):149–178.

Wagner, G. G., Frick, J. R., and Schupp, J. (2007). The German Socio-Economic Panel study (SOEP) – evolution, scope and enhancements. *Schmollers Jahrbuch*, 127(1):139–170.

Wolf, D. A. and Soldo, B. J. (1994). Married women's allocation of time to employment and care of elderly parents. *Journal of Human Resources*, 29(4):1259–1276.

A. Tables

Table 7: Regression DiD estimates of long-term care reform effects on male employment

	(1)	(2)	(3)	(4)	(5)
Post95	−0.035**	−0.026*	−0.027**	−0.025*	−0.027**
	(0.012)	(0.010)	(0.010)	(0.010)	(0.010)
Tr	−0.026	0.022	0.030	0.029	−0.072
	(0.063)	(0.059)	(0.056)	(0.055)	(0.105)
Post95× Tr	−0.196*	−0.153*	−0.145†	−0.151*	−0.162*
	(0.086)	(0.077)	(0.075)	(0.074)	(0.072)
Age≥60		−0.083*	−0.079*	−0.078*	−0.076*
		(0.035)	(0.033)	(0.033)	(0.033)
Age		0.245***	0.206***	0.199***	0.200***
		(0.029)	(0.029)	(0.029)	(0.029)
Age^2/100		−0.257***	−0.231***	−0.224***	−0.224***
		(0.028)	(0.027)	(0.027)	(0.027)
Migration background			−0.004	−0.009	−0.005
			(0.018)	(0.018)	(0.018)
Working experience in years			0.017***	0.017***	0.017***
			(0.002)	(0.002)	(0.002)
Years of education			0.033***	0.032***	0.033***
			(0.003)	(0.003)	(0.003)
Health status:					
good - very good (base)					

Table 7: Continued

satisfying	−0.041**	−0.040**	−0.039**
	(0.013)	(0.013)	(0.013)
poor – very poor	−0.174***	−0.169***	−0.167***
	(0.020)	(0.020)	(0.020)
Married	0.010	0.005	−0.003
	(0.024)	(0.024)	(0.024)
Other household income /1000		0.001**	0.001**
		(0.000)	(0.000)
Household size≥3		−0.041*	−0.036*
		(0.016)	(0.016)
Community size:			
<20,000 (base)			
20,000–100,000			0.033
			(0.024)
>100,000			−0.027
			(0.026)
HH-member needs help with:			
getting around outside the house (base)			
household chores, preparing meals			−0.049
			(0.166)
washing, dressing, etc.			0.132
			(0.121)
getting into and out-of-bed, etc.			0.125

Table 7: Continued

					(0.113)
Care spell in years					0.005
					(0.013)
R^2	0.01	0.26	0.36	0.36	0.37
Obs. in treatment group	150	150	150	150	150
Obs.	5033	5033	5033	5033	5033

Note: Reported values denote estimated coefficients. Standard errors are clustered on household level and reported in parentheses. Significance levels: † p <0.10, ∗ p <0.05, ∗∗ p <0.01, ∗ ∗ ∗ p <0.001.

Source: SOEPv30, own calculations.

Table 8: Regression DiD estimates of long-term care reform effects on female employment

	(1)	(2)	(3)	(4)	(5)
Post95	0.001	0.001	−0.017	−0.017	−0.016
	(0.013)	(0.011)	(0.011)	(0.011)	(0.011)
Tr	−0.136*	−0.051	0.025	0.030	0.050
	(0.063)	(0.058)	(0.054)	(0.054)	(0.099)
Post95× Tr	0.009	−0.026	−0.054	−0.055	−0.066
	(0.073)	(0.064)	(0.063)	(0.063)	(0.064)
Age≥60		−0.081*	−0.091**	−0.090**	−0.088**
		(0.036)	(0.033)	(0.033)	(0.033)
Age		0.115**	0.084*	0.084*	0.085*

Table 8: Continued

	(0.039)	(0.034)	(0.034)	(0.034)
Age^2/100	−0.133***	−0.107***	−0.108***	−0.109***
	(0.036)	(0.032)	(0.032)	(0.032)
Migration background		−0.019	−0.018	−0.023
		(0.021)	(0.021)	(0.021)
Working experience in years		0.017***	0.017***	0.017***
		(0.001)	(0.001)	(0.001)
Years of education		0.015***	0.015***	0.015***
		(0.004)	(0.004)	(0.004)
Health status:				
good - very good (base)				
satisfying		−0.064***	−0.064***	−0.063***
		(0.016)	(0.016)	(0.016)
poor – very poor		−0.129***	−0.129***	−0.129***
		(0.020)	(0.020)	(0.020)
Married		−0.073***	−0.077**	−0.074**
		(0.021)	(0.024)	(0.024)
Other household income /1000			0.000	0.000
			(0.000)	(0.000)
Household size≥3			0.020	0.019
			(0.018)	(0.019)
Community size:				
<20,000 (base)				

Table 8: Continued

20,000–100,000					0.024
					(0.029)
>100,000					0.040
					(0.030)
HH-member needs help with:					
getting around outside the house (base)					
household chores, preparing meals					−0.130
					(0.133)
washing, dressing, etc.					−0.064
					(0.098)
getting into and out-of-bed, etc.					0.018
					(0.122)
Care spell in years					0.008
					(0.011)
R^2	0.00	0.16	0.36	0.36	0.36
Obs. in treatment group	170	170	170	170	170
Obs.	4865	4865	4865	4865	4865

Note: Reported values denote estimated coefficients. Standard errors are clustered on household level and reported in parentheses. Significance levels: † p <0.10, ∗ p <0.05, ∗∗ p <0.01, ∗ ∗ ∗ p <0.001.

Source: SOEPv30, own calculations.

A.1. Robustness tests for the female sample

Table 9: Robustness tests for the treatment effect on female employment

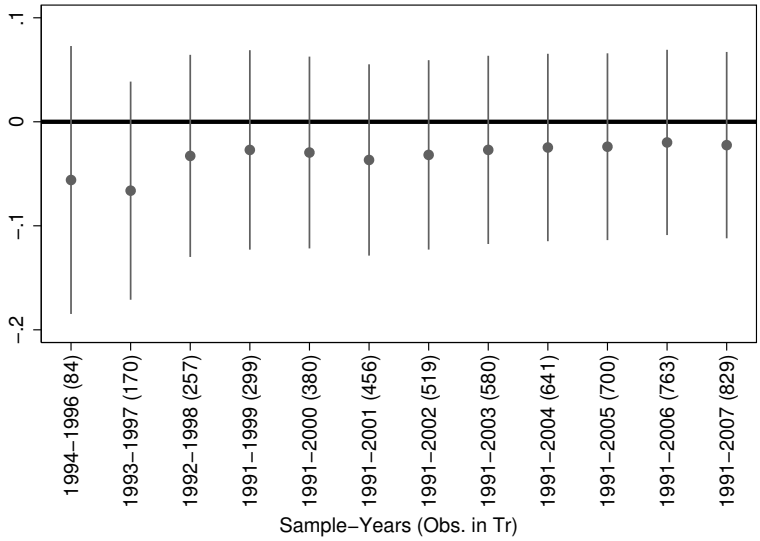
	(1) Constant HH	(2) DiD Matching	(3) Linear Time-Trend
Post95	-0.019	−0.049	
	(0.014)	(0.044)	
Tr	0.047	0.194	-5.062
	(0.147)	(0.174)	(12.528)
Post95× Tr	-0.015	0.012	-0.061
	(0.075)	(0.099)	(0.069)
Tr× Year			0.003
			(0.006)
Year			0.003**
			(0.001)
Age Variables	✓	✓	✓
Indivividual Characteristics	✓	✓	✓
Household Characteristics	✓	✓	✓
Care Variables	✓	✓	✓
Obs.	3114	320	29401
Obs. in Tr	93	60	829
Years in Sample	1993–1997	1994–1996	1991–2007

Note: Reported values denote estimated coefficients. Standard errors are clustered on household level and reported in parentheses. Significance levels: † p <0.10, ∗ p <0.05, ∗∗ p <0.01, ∗ ∗ ∗ p <0.001.

(1) This estimation includes only households whose composition does not change during the observation period. (2) This model is based on a propensity score matching in order to improve the balance between treatment and control group. In the pre-treatment year 1994 we estimate a probit model on the treatment dummy. Thereby, the explanatory variables include the employment decision, retirement status, age, migration background, working experience, years of education, health status, marital status, household size, and community size. We use five-to-one nearest neighbor matching with replacement. In order to impose common support we drop treatment observations whose propensity score is higher than the maximum or less than the minimum propensity score of the controls. If individuals of the treatment group drop out after treatment, their matching partners are dropped as well. (3) This model uses data from the maximum time window of 16 years and includes an interaction of a linear time trend with the treatment dummy (Tr × Year).

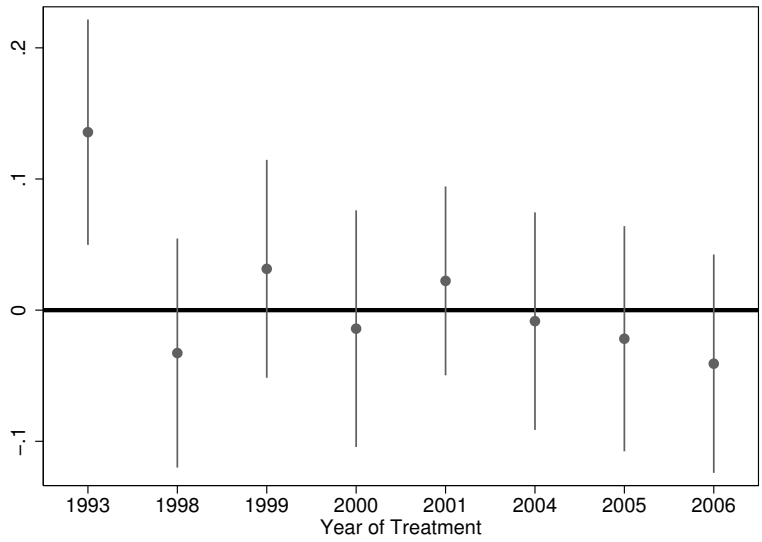
Source: SOEPv30, own calculation.

Fig. 5: Estimated treatment effect using different time spans (women)



Note: Dots represent point estimates controlling for the full vector of controls. Vertical lines represent the 95% confidence interval. The year 1995 is omitted in all models. Standard errors are clustered on household level.
Source: SOEPv30, own calculations.

Fig. 6: Placebo r egressions (women)



Note: Dots represent point estimates controlling for the full set of controls. Vertical lines represent the 95% confidence interval. Each model uses observations of two years before and after the assumed year of treatment. The year reported on the x-axis is always the first year of treatment. Standard errors are clustered on household level.
Source: SOEPv30, own calculations.