
# The Effects of an Increase in the Retirement Age on Health - Evidence from Administrative Data

**Mara Barschkett** (DIW Berlin and FU Berlin)
**Johannes Geyer** (DIW Berlin and Netspar)
**Peter Haan** (DIW Berlin, FU Berlin and Netspar)
**Anna Hammerschmid** (DIW Berlin)

Discussion Paper No. 302

November 29, 2021

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Georg Weizsäcker, Ph.D., Humboldt University Berlin, 10117 Berlin, Germany
info@rationality-and-competition.de

# The Effects of an Increase in the Retirement Age on Health – Evidence from Administrative Data*

Mara Barschkett[1,2], Johannes Geyer[1,3], Peter Haan[1,2,3], and Anna Hammerschmid[1]

[1]DIW Berlin

[2]FU Berlin

[3]Netspar

November 26, 2021

**Abstract**

This study analyzes the causal effect of an increase in the retirement age on health. We exploit a sizable cohort-specific pension reform for women using two complementary empirical approaches – a Regression Discontinuity Design and a Difference-in-Differences approach. The analysis is based on official records covering all individuals insured by the public health system in Germany and including all certified diagnoses by practitioners. This enables us to gain a detailed understanding of the multi-dimensionality in these health effects. The empirical findings reflect the multi-dimensionality but allow for deriving two broader conclusions. We provide evidence that the increase in the retirement age negatively affects health outcomes as the prevalence of several diagnoses, e.g., mental health, musculoskeletal diseases, and obesity, increases. In contrast, we do not find support for an improvement in health related to a prolonged working life since there is no significant evidence for a reduction in the prevalence of any health outcome we consider. These findings hold for both identification strategies, are robust to sensitivity checks, and do not change when correcting for multiple hypothesis testing.

**Keywords:** Germany, Retirement, Pension reform, Health, ICD-10, Regression Discontinuity Design, Difference-in-Differences

**JEL classification:** I10, I12, I18, J14, J18, J26.

*We are grateful to the National Association of Statutory Health Insurance Physicians (Kassenärztliche Bundesvereinigung, KBV) for data access and for their excellent support. We further thank Lena Janys, Adam Lederer, and Marius Opstrup Morthorst, as well as the participants at the Essen Health conference 2020, The Econometric Society/Bocconi University World Congress 2020, and at internal seminars at DIW Berlin. Moreover, we gratefully acknowledge funding from the German Science Foundation through the CRC/TRR190 (Project number 280092119) and Project HA5526/4-2. Mara Barschkett acknowledges funding from the Research Network of the German Pension Insurance Fund through a doctoral scholarship (Forschungsnetzwerk Alterssicherung, FNA). The authors declare that there is no conflict of interest. Corresponding author: Mara Barschkett, mbarschkett@diw.de, DIW Berlin, Mohrenstrasse 58, 10117 Berlin

# 1 Introduction

Aging populations present immense challenges for public pension systems due to growing numbers of beneficiaries and declining numbers of contributors. To sustain the systems' financial stability, policy makers across the OECD have introduced pension reforms which raised retirement ages. While postponing retirement has the potential to increase pension contributions and to reduce the share of pension benefit recipients, a prolonged working life might also have consequences for the health of individuals. Thus, to understand and to assess the overall impact of changes to the pension system, it is crucial to quantify and fully understand the health implications of pension reforms.

In this paper, we study the effects of an increase in the retirement age using official data on certified diagnoses by practitioners based on the International Classification of Diseases (ICD-10) for the period from 2009 to 2018. Thus, we assess effects on specific diagnoses and groups of diseases, thereby accounting for the impact of an increase in the retirement age on health in a multi-dimensional way. This detailed analysis is important since broader health measures might disguise potentially negative or positive implications for different health dimensions.

To identify the causal effect of an increase in the retirement age on diagnoses, we exploit a sizable and cohort-specific pension reform which was implemented in 1999. The reform abolished an early retirement program for women born after 1951[1] and thereby effectively increased the early retirement age (ERA) for women from age 60 to at least 63. It provides a clean quasi-experimental setting as it induces a substantial discontinuity in retirement ages for two adjacent cohorts (women born in 1951 versus women born in 1952). Using the same variation, Geyer and Welteke (2021) and Geyer et al. (2020), analyzing the employment effects as well as distributional consequences of the pension reform, show that the reform led to substantial individual labor market responses, including increased employment between age 60 and 62. Geyer and Welteke (2021) also show that labor market behavior before age 60 does not differ between cohorts. That is, affected women did not adjust participation rates or working hours before age 60 in anticipation of a higher retirement age. However,

[1]The majority of previous studies on the link between health and retirement use age discontinuities in the retirement age to instrument the individual's retirement status (see van Ours and Picchio (2020) for an overview of methodologies of previous studies). Only a few studies exploit direct variation from pension reforms (e.g., Bloemen et al., 2017; Charles, 2004; Etgeton and Hammerschmid, 2019; Grip et al., 2011; Kuhn et al., 2019).

Etgeton et al. (2021) show that the reform had negative effects on private savings.[2] Using data covering 2009 through 2018, we can consistently analyze the health effects for women aged 59, i.e. before the reform had a direct effect on employment (age-59-effects), for women aged 60–62 (main effects) and for women aged 63–65, which we define as post employment period.

We use two identification strategies, namely a regression discontinuity (RDD) and a Difference-in-Differences (DiD) design. The medical and demographic literature documents that health outcomes are correlated with month of birth as well as with cohort effects (e.g., Boland et al., 2015; Doblhammer and Vaupel, 2001). Therefore, it is crucial to account for cohort and seasonality (month of birth) effects to isolate the causal effect of the pension reform on health. First, we follow e.g. Geyer and Welteke (2021) and use an RDD. The pension reform leads to an arbitrary and distinct cutoff for women born before and after December 31, 1951, which determines the assignment into treatment and control groups. Second, we use an additional identification strategy and account for potential month of birth effects using cohorts not affected by the reform in a control group. Specifically, similar to Schönberg and Ludsteck (2014), we define a treatment group (women born between October 1951 and March 1952) and a control group (women born between October 1950 and March 1951) and use the pension reform in a DiD framework.

In the analysis, we focus on three dimensions of health: mental health, physical health, and health care consumption. Within these dimensions, we concentrate on groups of diseases that are most likely affected by lifestyle choices and that have been used in existing studies on the link between health and retirement. Within these groups, we select the diagnoses most frequently causing rehabilitation treatments prescribed by the pension insurance in the application process of invalidity benefits ("Erwerbsminderungsrente"). More precisely, we analyze the impact of the increase in the retirement age on mood (affective) disorders and on neurotic, stress-related, and somatoform disorders (hereafter: stress-related diseases) to assess the effects on mental health. For the physical health dimension, we consider the group of metabolic and cardiovascular diseases (diabetes mellitus, obesity, hypertensive diseases, ischaemic heart diseases, and cerebrovascular diseases (strokes)) as well as the group of musculoskeletal diseases (arthrosis and other dorsopathies). In addition, we study

[2]To date, few other studies exploit variation from the 1999 pension reform: e.g., Gohl et al. (2020) use the reform to test the human capital theory and Fischer and Müller (2020) analyzes its impact on informal care provision. Moreover, Etgeton and Hammerschmid (2019) study the effects of retirement on self-reported health, in particular across educational groups, using a two-sample-2SLS approach.

hypertension since this is the most common physical disease within our sample, but is not captured using the rehabilitation criterion. To estimate the impact on health care consumption, we examine the annual number of treatment cases.

Our empirical findings provide evidence that the increase in the retirement age has a negative effect on health outcomes as the prevalence of several diagnoses, e.g. mental health, musculoskeletal diseases, and obesity, increases. In contrast, we do not find support for an improvement in health related to a prolonged working life since there is no significant evidence for a reduction in the prevalence of any health outcome we consider. These findings hold for both identification strategies, are robust to sensitivity checks, and do not change when correcting for multiple hypothesis testing. In particular, we find that the pension reform increased the prevalence of both mental diseases in 60–62 year old women. The effect sizes range from 3.6 to 4.8 percent for stress-related diseases and from 4.8 to 8.3 percent for mood disorders relative to the respective pre-treatment means. The effects for 59 year old women are of similar magnitude and significance. Within the physical health dimension, our results suggest that raising the retirement age increases the prevalence of arthrosis and obesity at ages 60–62 years as well as 59 years. For other physical health outcomes, our results are less clear but, as mentioned above, we do not find significant evidence of an improvement in physical health in response to the reform. Furthermore, we find significant effects of the reform on healthcare consumption for 59 year olds. Overall, our findings reflect the multi-dimensionality of health but allow us to conclude that the reform had negative and significant effects on some health outcomes and did not have positive and significant effects on any of the considered health outcomes. Additional analyses on post-employment effects suggest that the majority of the effects persist into retirement (at age 63–65), but effect sizes are smaller compared to the direct effects on 60–62 year old women.

The existing literature on the health effects of retirement and pension reforms can be divided into four strands: Studies using survey data and exploring effects of retirement on i) mental health or ii) physical or general health, and studies using administrative data considering iii) mortality or iv) health care usage or diagnoses as outcome variables. We discuss the relation of our paper to these four strands in the following:[3]

**Survey data: Mental health**

A number of studies find positive effects of retirement on mental health (e.g., Atalay and

[3]For a more detailed overview of the literature please refer to e.g. Garrouste and Perdrix (2021)

Barrett, 2014; Atalay et al., 2019; Belloni et al., 2016; Charles, 2004; Eibich, 2015; Gorry et al., 2018; Grip et al., 2011; Leimer, 2017). Atalay and Barrett (2014), for example, exploit variation of a pension reform for women in Australia and find positive effects of retirement on mental health. They emphasize that the effects can mostly be attributed to a reduction in mood disorders. Eibich (2015) uses data from the German Socio-economic Panel (SOEP) to estimate an RDD exploiting age thresholds in the German pension system. He also finds positive effects of retirement on mental health and explains this by a reduction in work-related stress and more frequent exercise (cf. Celidoni and Rebba (2017)). Applying a similar methodology van Ours and Picchio (2020) find heterogeneous effects for the Netherlands. They find positive effects of retirement on the mental health of men and their partners but no effects for women or singles.

In contrast, there are also studies showing no, if not negative, effects of retirement on mental health. For example, Heller-Sahlgren (2017) conducts a cross-country analysis using the Survey of Health, Ageing and Retirement in Europe (SHARE) and employs an RDD approach. He finds no effects on mental health in the short-run but a large and negative long-run impact. Similarly, Rohwedder and Willis (2010) find negative effects on cognitive abilities in a cross-national study in the US and Europe. These results are also supported by Mazzonna and Peracchi (2017), who find a decline in cognitive abilities following retirement for most workers using SHARE data. Atalay et al. (2019) find a negative but modest effect on cognition, the effect is larger for men than for women.

**Survey data: Physical and general health**

The relationship between physical or general health and retirement is also ambiguous in the literature. Coe and Zamarro (2011) and Gorry et al. (2018) find positive effects of retirement on self-reported health status in Europe using SHARE data. Shai (2018) reports similar findings for Israel. Leimer (2017) uses SHARE data and reports a reduction in mobility limitations and the number of limitations in activities of daily living along with an increase in maximum grip strength following retirement. Close to our study, in particular in terms of the same reform being used for identification, is Etgeton and Hammerschmid (2019). They focus on the effects of retirement on broad, self-reported health, in particular across educational groups, based on SOEP and SHARE data. Using a two-sample 2SLS approach, they identify the impact of retirement on health using the 1999 pension reform in Germany. Their findings point toward non-detrimental general health effects of retirement, with less

educated women benefiting more than the average. In addition to positive effects on mental health, Atalay and Barrett (2014) also find positive effects on physical health, namely on hypertension, migraine, back pain, and disc disorders for women in Australia.

Negative effects of retirement on physical health are found, for example, by Godard (2016) (increase in BMI with SHARE data) and Behncke (2012). Specifically, they discover an increase in risk of being diagnosed with a chronic condition and an increase in risk of developing a cardiovascular disease in the UK following retirement.

Two examples of studies assessing the effect of retirement on health care consumption are Zhang et al. (2018) for China and Eibich (2015) for Germany. While Zhang et al. report increased health care utilization following retirement, Eibich finds a decrease in both hospitalization and number of doctor visits.

Compared to our paper, the first two literature strands discussed here use survey data and mostly self-assessed health measures. Thus, studying administrative diagnose data, our paper is complementary to this part of the existing literature.

The reasons for the discrepancies in the literature are not comprehensively and systematically studied yet, but contributing factors seem to be, for instance, differences in empirical methods, data sources, pension systems, health care systems, effect heterogeneity in subpopulations, and differing outcome variables (Nishimura et al., 2018; Pilipiec et al., 2020).[4] Furthermore, heterogeneity in the effects of retirement on different health dimensions could potentially also contribute to explaining the contradictory results. There is ex ante no reason to believe that the effects of retirement (reforms) on different health dimensions are indeed homogeneous and go into the same direction. Some aspects of mental or physical health may be positively affected whereas others may be negatively affected.

**Administrative data: Mortality**

Analyses using detailed administrative data including objective health measures have the potential to explore this issue. So far, only a small number of studies use this kind of data. Three examples of studies looking at the effect of retirement on mortality are Kuhn et al. (2019), who find negative effects for Austrian men, Fitzpatrick and Moore (2018) for the US, and Brockmann et al. (2009) in the German context. Brockmann et al. (2009) use German health insurance data from one specific health insurance fund and find heterogeneous effects

[4]Nishimura et al. (2018) show that the choice of empirical method plays a key role in explaining why estimated results differ across studies.

across individuals with good and poor health. Healthy people benefit from retirement while individuals with poor health tend to have decreased life expectancy following early retirement. In contrast, Hallberg et al. (2015) use a pension reform for military officers that decreased the retirement age from 60 to 55 in Sweden. They find support that early retirement leads to a reduction in mortality. Hernaes et al. (2013) find no effect of a series of retirement reforms that reduced the retirement age on mortality in Norway.

It is important to note that death is a specific and extreme outcome. Mortality rates are rather low around retirement age. Potential effects on mortality might only establish later in the long run. Thus, it is difficult to estimate mortality effects of recent pension reforms, such as the 1999 reform studied in this paper.

**Administrative data: Health care consumption and diagnoses**

Studies using administrative data and considering health outcomes other than mortality are rare; these mostly find positive effects of retirement on health. The following studies are closely related to our study:

Kuusi et al. (2020) use Finish registry data (a random sample covering 11% of the population) and an IV approach to assess the effect of retirement on mental health and physical health. They measure mental health with antidepressant purchases and physical health by hospital visits associated with cardiovascular or musculoskeletal diseases. They find substantial positive effects on mental health and small effects on physical health. Similarly, Nielsen (2019) uses Danish full population data to assess the effect of retirement on general practitioner (GP) visits, hospitalization, comorbidities, and mortality using IV and RDD approaches. He finds a reduction in GP visits and hospitalization following the reform, but no effect on comorbidities and mortality. Hagen (2018) conducts a similar study in Sweden but does not find an impact of retirement on health. He uses Swedish data for women in the public sector to estimate the effect of a pension reform on drug prescriptions, hospitalizations, mortality, and cause-specific health indices in a DiD framework. There are only a few studies outside the Nordic countries relying on administrative data (e.g. Frimmel and Pruckner, 2020; Horner and Cullen, 2016; Rose, 2020). Horner and Cullen (2016) use administrative data from the US on a specific group, manufacturing workers in an aluminum production company, to evaluate the impact of retirement on hypertension, diabetes, asthma, arthritis, and major depression. They find a reduction in asthma following retirement but no effects on the other outcome variables. Frimmel and Pruckner (2020)

study the effect of two Austrian pension reforms on individual inpatient and outpatient healthcare utilization in Austria and find that retirement decreases service utilization and healthcare expenditure. Rose (2020) uses a combination of administrative and survey data from the UK to study a variety of outcomes: She generally finds a positive association between retirement and health, e.g. an increase in self-reported health, a decrease in long-term ailments, lower pulses, more sleep and generally an improvement in healthy behaviors (e.g. reduced smoking and drinking). However, she does not find retirement to impact cognition, mental health, health care utilization and mortality.

Our paper extends the literature in several ways. First, we study a major pension reform that led to a substantial increase of the retirement age of three years. Second, our study is based on unique administrative health records that cover almost the whole German population. Moreover, the data include all recorded diagnoses in outpatient care during the observation period. Thus, in contrast to most of the previous studies we can study the multi-dimensionality of health effects for a very general population. Third, we provide evidence that effects from increasing the retirement age are not bound to the affected age group. Instead, increasing the retirement age implies expectation effects (effects for the age group before reaching the retirement age) and the effects persist into the post-employment period.

The remainder of the paper is structured as follows: Section 2 describes the institutional background in Germany. In Section 3, we give an overview over the data. The empirical strategy is explained in Section 4 and, in Section 5, we present the results and provide several robustness checks. Finally, Section 6 concludes.

## 2 Institutional background - Pension system

To establish the institutional setting of the analysis, we provide an overview on the relevant institutions of the German pension system[5] and discuss the 1999 pension reform, which induced an exogenous increase in the early retirement age for women born after 1951.

The public pension system in Germany covers about 90% of the workforce.[6] Pension benefits account for about two-thirds of gross income of the elderly. It includes old-age pensions,

[5]For a more general description of the German pension system, see the German country profile by the OECD available at `http://oe.cd/pag`.

[6]There are a few exemptions from compulsory insurance: civil servants have a separate tax-financed, non-contributory scheme and most of the self-employed are not compulsory insured.

disability pensions, and survivors' benefits. The system is financed by a pay-as-you-go (PAYG) scheme and has a strong contributory link. The calculation of pension benefits is based on a points system and depends on the entire working history.[7] The statutory pension age (SRA) was 65 for cohorts born before 1947. It is stepwisely raised to age 67 and fully phased in for all cohorts born in 1964 or later. For the 1951 cohort, the SRA was 65 and 5 months, for those born in 1952 it was 65 and 6 months. People qualify for this regular old-age pension after five years of pension contributions.

Retirement before the SRA (with permanent deductions) is possible under certain conditions.[8] There are four alternative pathways to claiming early retirement benefits: the *pension for women*, the disability pension, the pension for the long-term insured, and the pension after unemployment or after partial retirement. There is a fifth option, invalidity benefits ("Erwerbsminderungsrente"), for people with severe health problems who are not able to work more than three hours a day.[9] In general, the calculation of pension benefits does not vary between these alternatives, whereas eligibility criteria differ.[10] The 1999 reform abolished the *pension for women* for cohorts born after 1951. Effectively, the reform raised the ERA for most women from 60 to 63, which implies an extension of the working life of three years. The eligibility criteria of the *pension for women* were: (i) at least 15 years of pension insurance contributions; and (ii) at least 10 years of pension insurance contributions after the age of 40. According to Geyer and Welteke (2021), about 60% of all women born in 1951 were eligible for the old-age pension for women.

Geyer and Welteke (2021) and Geyer et al. (2020) evaluate the labor market effects of the 1999 pension reform. Several findings of these studies are relevant for the subsequent empirical analysis. Most importantly, the increase in the ERA has sizable labor market effects: retirement rates of eligible women aged between 60 and 62 decreased by about 30 percentage points. At the same time, employment rates increased by about 15 percentage points. Inactivity and unemployment increased by about 12 percentage points. Geyer and Welteke (2021) document that the pension reform had no significant effect on labor market activity before the age of 60. Moreover, the employment effect results almost entirely from women staying longer in the respective labor market status; there is no significant evidence

[7] People also acquire pension entitlements during short-term unemployment, for childcare, and for providing elderly care.

[8] There is no change to public health insurance coverage when starting to draw retirement benefits.

[9] People who are able to work more than three hours a day but less than six are eligible for partial invalidity benefits. These benefits are available before the age of 60.

[10] For more details see Geyer et al. (2020).

that the unemployed make more transitions into employment. Finally, the pension reform did not lead to substitution effects into other health-related early retirement pathways (disability pension or invalidity benefits). The prolonged duration in the labor market is the main treatment for the 1952 cohort. Therefore, the reform effect operates through different channels which we cannot differentiate with the data at hand. The majority of treated women stays longer in employment which might affect health. However, the prolonged status in unemployment could as well impact health.

# 3 Data

For the analysis, we use administrative data covering the years 2009-2018, collected by all public health insurance funds in Germany.[11] In the data, physicians record a standardized diagnosis for each claim in order to be reimbursed by the health insurance.

In Germany, health insurance is mandatory and characterized by a public insurance system and a private insurance system. Nearly 90% of the German population is covered by one of the public health insurance funds.[12] Only individuals with earnings exceeding a certain threshold[13] and individuals in specific occupational groups (e.g., civil servants and self-employed) are allowed to opt out of the public system and to sign up with a private insurance company instead.[14]

With the data, in principle we can focus on women born between 1950–1953. However, a major school reform affects many women born after 1952, therefore in the empirical analysis we consider only cohorts 1950–1952.[15] This allows us to construct a placebo group (women born late in 1950 and early in 1951) in addition to the group of women around the cutoff date of the pension reform (women born late in 1951 and early in 1952). We have access

[11] The data are based on the database of claims of all publicly insured individuals in Germany as collected by the Association of Statutory Health Insurance Physicians and then forwarded to the National Association of Statutory Health Insurance Physicians (Kassenärztliche Bundesvereinigung, KBV).

[12] Public health insurance is financed primarily through mandatory contributions from employers and employees, along with tax revenues. Contributions are pooled in the Central Health Fund (Gesundheitsfonds) and reallocated to the sickness funds according to a morbidity-based risk adjustment scheme. There are currently about 109 health insurance funds. For more information about the German health insurance system, see OECD (2019).

[13] The income threshold for 2020 was 62,500 euro ($\approx 74,500$ dollar) per year.

[14] Importantly, similar rules apply for the eligibility of the public health and public pension insurance. Individuals with a private health insurance, e.g., civil servants and the self-employed, have additional private pension plans that were not affected by the pension reform.

[15] Regional schooling reforms in western Germany raised compulsory schooling from 8 to 9 years. Four large federal states changed compulsory schooling within cohort 1953. The reform had positive effects on health outcomes (Kemptner et al., 2011).

to data covering 2009 through 2018, thus we can consistently analyze the health effects for women aged 59, i.e. before the reform had a direct effect on employment (age-59-effects), for women aged 60–62 (main effects) and for women aged 63–65, which we define as post employment period. As mentioned above women born in 1952 or later can enter retirement at age 63.

The data include information about all diagnoses patients received during the observed period. Each diagnosis constitutes a new entry meaning that the number of observations equals the number of diagnoses over the observed time period. Thus, the sample is unbalanced as patients only appear if they received outpatient care including a diagnosis. Based on this information, we construct a balanced sample with yearly information for all publicly insured individuals.[16] The final data set includes about 500,000 women per birth cohort resulting in 1.5 million women overall.[17] The data only includes few demographic characteristics, such as age and region.

Instead of estimating the effect for about 70,000 different diagnoses categorized by the ICD-10 codes, we use clear criteria to select the relevant health outcomes. Specifically, we concentrate on groups of diseases that are most likely affected by lifestyle choices and are used in the existing literature on the link between health and retirement. Within these groups, we select the diagnoses that most frequently caused rehabilitation measures prescribed by the German pension insurance for our age group.[18] In addition, we study hypertension since this is the most common disease within our sample and is not captured using the rehabilitation criterion.

Specifically, we define the following groups:

- Mental and behavioral disorders (ICD-10)

  F30-F39: Mood (affective) disorders

[16] First, we create variables indicating whether an outcome, for example diabetes, was diagnosed or not in a specific period. Secondly, we aggregate the data to a yearly level such that each patient appears only once per year. Finally, we balance the data by imputing information for patients without outpatient care in a specific year. By definition, all outcome variables are zero as the patient did not receive a relevant diagnosis during this year. The definition of our outcome variables is analogous to Van den Berg and Siflinger (2020).

[17] Women who did not receive any outpatient care during the 10 year observation period are not included in our sample. However, RKI (2010) states, that 90% of women receive outpatient care at least once per year. Thus, given that we observe individuals over 10 years, the share of women not receiving any outpatient care should be negligible.

[18] Employees can receive medical rehabilitation benefits if their earning capacity is at considerable risk or already reduced. The goal is that individuals recover such that they can return to the labor market and do not need invalidity benefits. We selected the diseases that were responsible for at least 20% of the prescription cases within a group of diseases. The list is accessible at `https://statistik-rente.de/drv/`.

F40-F48: Neurotic, stress-related and somatoform disorders (stress-related diseases)

- Endocrine, nutritional and metabolic diseases and diseases of the circulatory system (cardiovascular) (ICD-10)

  E10-E14: Diabetes mellitus

  E65-E68: Obesity and other hyperalimentation

  I10-I15: Hypertensive diseases

  I20-I25: Ischaemic heart diseases

  I60-I69: Cerebrovascular diseases (strokes)

- Diseases of the musculoskeletal system and connective tissue (ICD-10)

  M15-M19: Arthrosis

  M50-M54: Other dorsopathies

- Health care consumption

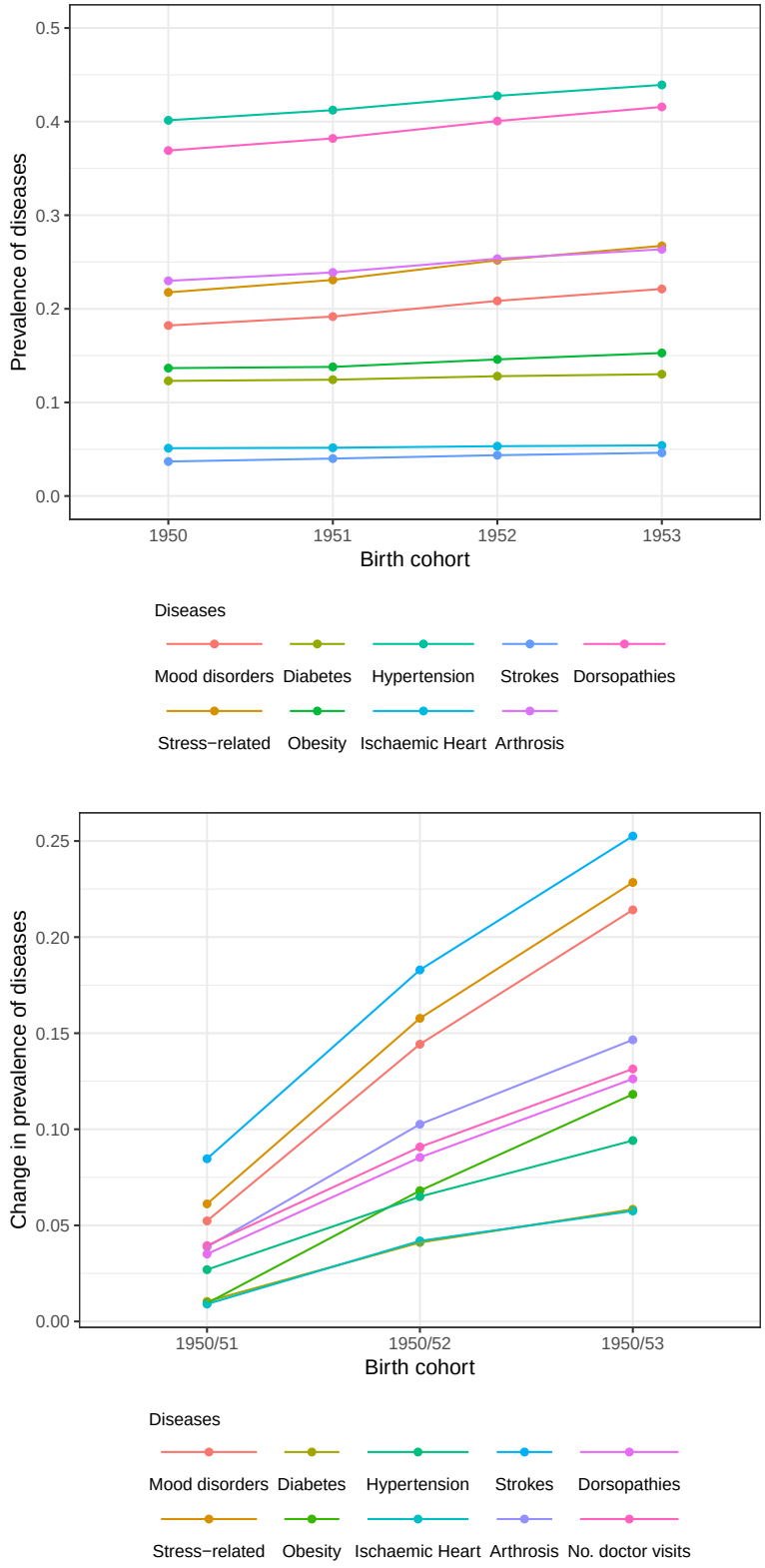
  Doctor visits

Figure 1 shows the prevalence of the selected diagnoses within our sample and how they vary across cohorts. The top panel presents the average share of women suffering from a certain disease by birth cohort. The prevalence of diseases in our sample ranges from about 5% (ischaemic heart diseases and strokes) to more than 40% (hypertension and other dorsopathies). It is also visible that most diseases have a positive and sizable cohort trend, meaning that younger cohorts have a higher likelihood to be diagnosed with one of the diseases. This pattern becomes clearer in the bottom panel of Figure 1, which presents the percentage difference in the prevalence of the diseases compared to cohort 1950. The graphical evidence underlines the importance to control for cohort effects to identify the causal reform effect in the empirical analysis.

# 4 Empirical strategy

We estimate the effect of the 1999 pension reform on health outcomes and use two different identification strategies, an RDD design and a DiD approach. Following Geyer and Welteke

Figure 1: Prevalence of diagnosed diseases and cohort trends




*Notes:* The top figure presents the average prevalence of the different diseases among women between age 60 and 62 by birth cohort. The dots represent means. The bottom figure presents the average change in the prevalence of the diseases and the number of doctor visits compared to birth cohort 1950 for women between age 60 and 62.

*Source: KBV, own calculations*

(2021), we first exploit the reform cutoff in an RDD setting. The pension reform leads to an arbitrary and distinct cutoff for women born on or before December 31, 1951, and those born after; this determines the assignment into treatment and control groups.

As a complementary identification strategy, we use a DiD approach. The medical literature (e.g., Boland et al., 2015; Doblhammer and Vaupel, 2001) documents that the month of birth is correlated with health outcomes. In the RDD, we can only account for seasonality (month of birth effects) by including quarter of birth as a control variable. This, however, requires an observation period of at least 12 months before and after the cutoff, thus exacerbating the challenge to absorb cohort effects, especially if observations are grouped by month of birth. Therefore, we additionally use an alternative DiD strategy that explicitly accounts for potential month of birth effects by differencing them out. Specifically, like Schönberg and Ludsteck (2014), we define a control group (women born between October 1950 and March 1951) and a treatment group (women born between October 1951 and March 1952). Women born between January and March are considered to be born after the cutoff. Thus, the interaction between control group and being born after the cutoff estimates the effect of the pension reform in a DiD setting.

In order to account for correlation between observations of the same individual or individuals born in the same month, we use robust standard errors clustered by month of birth in both specifications. In addition, we perform multiple hypotheses tests to account for the uncertainty related to the relatively large number of outcome variables.

### 4.1 Regression discontinuity approach

More formally, in the RDD, the woman's month of birth is the running variable $M$, which determines treatment $D$ to be one if she was born on or after January 1, 1952 ($c$) and zero otherwise:

$$D_i = \begin{cases} 1, & if\ M_i \geq c \\ 0, & if\ M_i < c \end{cases} \tag{1}$$

For the identification of a causal effect, it is important that no manipulation of the month of birth for women born in 1951 and 1952, i.e. the running variable, and no selection into or out of treatment are possible. As a result, the treatment and control groups should be comparable around the cutoff. Using representative data Geyer and Welteke (2021) and Geyer et al. (2020) show that treatment and control group do not differ with respect to

socio-economic characteristics. Moreover, as discussed, e.g., in Geyer and Welteke (2021), no other relevant policy reform affected women born in 1951 and 1952 differently.[19]

In the main specification, we implement the RDD according to the following equation:

$$y_{it} = \alpha^{RDD} + \beta^{RDD} D_i + \gamma_0^{RDD} f(M_i - c) + \gamma_1^{RDD} D_i f(M_i - c) + X_{it}\delta^{RDD} + \varepsilon_{it}^{RDD} \quad (2)$$

$D_i$ is a dummy specifying treatment that is equal to 1 if a woman is born in January 1952 or later, and 0 otherwise. A woman's month of birth is described by $M_i$ and $c$ is the cutoff date for the increase in early retirement age (ERA, January 1952). The function $f$ represents the trend in the running variable. In our main specification, we include a linear and quadratic cohort trend. This function is interacted with the treatment variable $D_i$ to allow for different slopes before and after the cutoff. In addition, we account for further explanatory variables ($X$), including quarter of birth and age. Controlling for birth quarter allows us to account for month of birth effects in the prevalence of diseases. In the main specification, the outcome variable $y_{it}$ is defined as an indicator variable that is equal to one if the disease of interest was diagnosed at least once during a calendar year.[20]

### 4.2 Difference-in-Differences approach

In the DiD approach, we use the same variation, but we isolate the causal effect of the reform from month of birth and cohort effects by comparing health outcomes of women born in the same calendar months across cohorts affected and not affected by the reform. Importantly, the sample only includes individuals born between October 1951 and March 1952 as well as between October 1950 and March 1951, respectively. Thus, birth months between March and October are not included in the sample. This way, we avoid comparing birth months that are rather far away from the reform cutoff in January.

[19] Geyer and Welteke (2021) discuss threats for identification from other policy reforms which had a differential impact on cohorts 1951 and 1952. The normal retirement age was 65 and 5 months for women born in 1951 and 65 and 6 months for women born in 1952. This difference is only relevant for women who plan not to retire early. Most women who qualified for early retirement used this option. Therefore, there are no visible differences in labor market behavior between cohorts after age 63. The abolishment of the pension after unemployment or after partial retirement for the 1952 cohort had no relevant impact since most eligible women also qualify for the pension for the long-term insured. Finally, the early retirement age for the disability pension increased for the 1952 cohort in monthly steps for those born until June and remained at 60 and 6 months for those born between July and December. This type of pension, however, was only accesible for women with severe disability status and long insurance record. It entailed lower deductions than other early retirement programs and was always financially more attractive than other early retirement options.

[20] As a robustness check, we run regressions including only a linear cohort trend and adding an east/west dummy variable.

Specifically we estimate the following equation:

$$y_{it} = \alpha^{DiD} + \beta_0^{DiD} Winter_i^{5152} + \beta_1^{DiD} JanFebMar_i + \beta_2^{DiD} Winter_i^{5152} \times JanFebMar_i + Z_{it}\delta^{DiD} + \varepsilon_{it}^{DiD} \quad (3)$$

where $Winter_i^{5152}$ indicates whether individual $i$ was born between October 1951 and March 1952. The indicator is zero if individual $i$ was born between October 1950 and March 1951. $JanFebMar_i$ is the reform indicator that is one if individual $i$ was born between January and March and zero otherwise. $Winter_i^{5152} \times JanFebMar_i$ is the interaction between the two indicator variables and turns one for every woman born from January 1952. Thus, the interaction term marks the individuals who are affected by the reform. In addition, we account for age effects captured in $Z_{it}$.[21]

# 5 Empirical results

In the following, we present the estimation results of the RDD and DiD estimation and discuss how an increase in the retirement age affects the health outcomes defined above. We estimate the effects for different age groups. Our main focus is on the group of 60–62 year old women. Effects are most direct for this group because in younger ages women of neither cohort can enter an old age retirement scheme. Women's health might, however, react to the reform already before reaching the age of 60 because they anticipate and expect to retire only three years later than expected. Therefore, we also study effects at age 59. There are two main channels through which the expectation of retiring only at age 63 could affect health at age 59: First, the effect could be caused by the expectation of working three years longer ("real" retirement effect). Second, cohort 1952 could perceive the reform as unfair as their only slightly older peers can retire three years before them (fairness effect). Thus, effects at age 59 are likely a mixture of both a "real" retirement effect and a fairness effect. In Section 5.5 we will turn to women aged 63–65. Women born in 1952 or later can enter retirement at age 63, therefore theses results can be interpreted as post employment effects.

In the data, we neither have information on the working history of women nor on their eligibility for the old-age pension for women. Therefore, we identify an intent-to-treat

[21] As a robustness check, we additionally include an east/west dummy variable.

effect (ITT) of the pension reform. According to Geyer and Welteke (2021), about 60 percent of all women born in 1951 were eligible for the old-age pension for women. In the following, we only present the results of the main specification and show a broad set of robustness checks and placebo tests in the Appendix.

## 5.1 Results – Mental health

We start with the discussion of the effects of the pension reform on two dimensions of mental health: stress-related mental diseases and mood disorders. The first subsection depicts descriptive, graphical evidence. Thereafter, we present the estimated causal effects of the increase in the ERA based on the RDD and the DiD.

### 5.1.1 Graphical analysis – Mental health

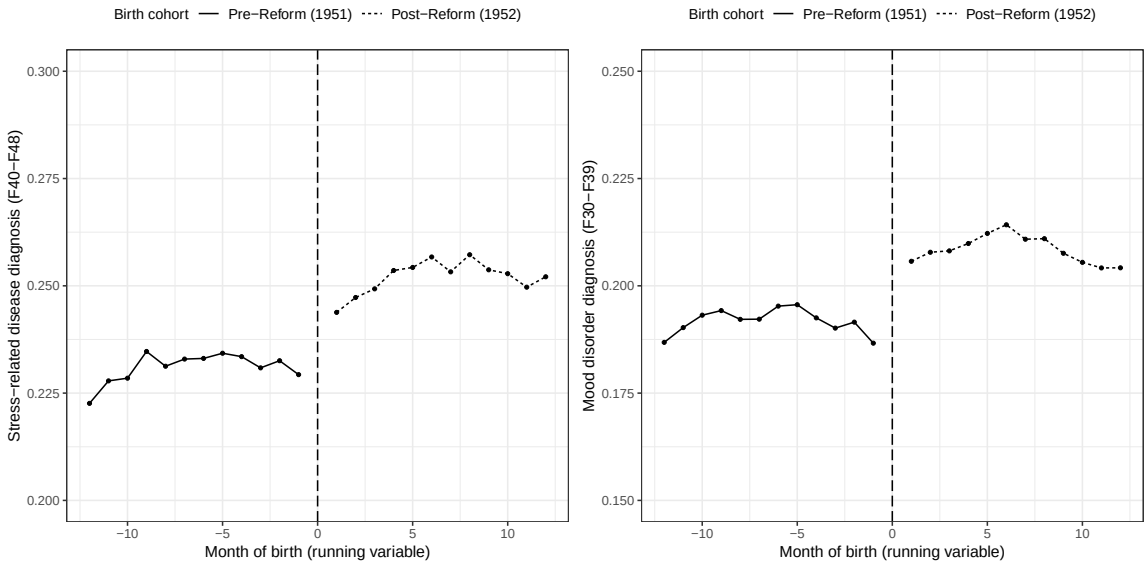
Figure 2 shows the average share of women aged 60 to 62 who are diagnosed with a stress related or a mood disorder diagnosis by month of birth. For both groups of diseases, there is a distinct and clear jump at the reform cutoff that ranges between one and two percentage points. In addition, there is evidence of seasonality in the trend both before and after the cutoff. This underlines the importance of controlling for quarter of birth in addition to potential cohort effects to identify the causal effect of the reform.

### 5.1.2 Regression results – Mental health

The regression results based on the RDD (Table 1) and the DiD (Table 2) confirm the graphical evidence: The first Column of Table 1 shows the estimated causal effects on stress-related mental diagnoses for women aged 60–62 using the RDD estimation approach. The coefficient of interest amounts to 0.011, thus the increase in the ERA increases the probability of a mental diagnosis between ages 60 and 62 by 1.1 percentage points. Relative to the prevalence in the 1951 cohort (pre-treatment mean), the magnitude of this effect amounts to about 4.8 percent. Turning to the estimation results for mood disorders in Column 3 of Table 1, we find that the effects of the reform are slightly larger in this mental health dimension. The effect of the increase in the ERA on being diagnosed with a mood disorder amounts to 1.6 percentage points, which corresponds to 8.3 percent in relation to pre-treatment prevalence.

So far, we focus on the effects of the main group of interest, namely 60–62 year old women.

Figure 2: Diagnoses of mental and behavioral disorders by month of birth




*Notes:* The left figure presents the average share of women between age 60 and 62, who got a F40-F48 diagnosis in a given year, for each birth month. The right figure presents the average share of women between age 60 and 62, who got a F30-F39 diagnosis in a given year, for each birth month. The vertical lines represent the cutoff date (01/1952).

*Source: KBV, own calculations*

As mentioned above, women's health might react to the reform even before reaching the age of 60 because they know that they need to work three years longer. As Columns 2 and 4 in Table 1 show, the point estimates for women aged 59 are comparable to the estimates for women in the main sample. These findings suggest that expectations play a role in the effects of the pension reform on mental health.

Table 1: RDD results: Mental diagnoses

| | *Stress-related* | | *Mood disorder* | |
|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 |
| $D_i$ | 0.011** | 0.013** | 0.016*** | 0.014*** |
| | (0.003) | (0.004) | (0.005) | (0.004) |
| $Birthmonths$ | −0.001* | −0.002** | −0.002*** | −0.004*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| $(Birthmonths)^2$ | −0.0002*** | −0.0002*** | −0.0002*** | −0.0003*** |
| | (0.00004) | (0.0001) | (0.00005) | (0.0001) |
| $D_i \times (Birthmonths)$ | 0.005*** | 0.005*** | 0.005*** | 0.009*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| $D_i \times (Birthmonths)^2$ | −0.00004 | 0.00004 | −0.00002 | −0.00000 |
| | (0.0001) | (0.0001) | (0.0001) | (0.0001) |
| Pre-treatment mean | 0.231 | 0.212 | 0.192 | 0.173 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no |
| Control for birth season | yes | yes | yes | yes |
| Observations | 3,429,155 | 1,235,612 | 3,429,155 | 1,235,612 |

$^{+}$p<0.1;$^{*}$p<0.05; $^{**}$p<0.01; $^{***}$p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) and (3) show the RDD estimates for women aged 60–62 years and include age and birth quarter dummies as control variables. Column (2) and (4) show the RDD estimates for women at age 59 and include birth quarter dummies as control variables. All regressions include linear and quadratic cohort trends in the running variable on both sides of the policy cut-off.
*Source: KBV, own calculations*

As shown in Appendix Table A.2, the magnitude of the estimated effects does not change when a linear cohort trend is used instead of a quadratic trend. Using a linear trend, the estimation is less precise, which makes the age-59-effects insignificant. In a further robustness check, we additionally control for regional effects (East and West Germany) and show that results do not differ (Table A.3).[22] In addition, in the Appendix we present results of the placebo test for which we use the same RDD specification but shift the cutoff date to January 1951. The test shows no significant effect for either health outcome, which suggests that the RDD specification accounts for potential cohort and month of birth effects driving stress-related mental diseases and mood disorders (Table A.1).

The DiD results are very similar to the main RDD results. The effect on stress-related diseases for 60–62 year old women amounts to 0.8 percentage points (3.6 percent relative to the pre-treatment mean) and is only slightly lower than the effect estimated using the RDD approach. For mood disorder diagnoses, the estimated effect in age group 60–62 amounts to 0.9 percentage points, which corresponds to a relative effect of 4.8 percent in relation to the pre-treatment mean. Compared to the RDD effect (1.6 percentage points),

[22] In a further robustness check (not reported) we control for regional differences by Bundeslaender, which accounts for differences in the school system, economic situation, or number of doctors. Results do not differ.

this effect is smaller but still highly statistically significant. The effects for 59 year old women are confirmed in magnitude and significance in this alternative DiD setting. The DiD specification is robust to including a dummy variable indicating West-Germany as an additional control variable (Table A.4).

Table 2: DiD results: Mental diagnoses

| | *Stress-related* | | *Mood disorder* | |
|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 |
| $Winter5152_i \times JanFebMar_i$ | 0.008** | 0.015*** | 0.009*** | 0.014*** |
| | (0.003) | (0.003) | (0.002) | (0.003) |
| $Winter5152_i$ | 0.012*** | 0.007*** | 0.008*** | 0.002 |
| | (0.002) | (0.001) | (0.002) | (0.002) |
| $JanFebMar_i$ | 0.008** | 0.004 | 0.009*** | 0.006* |
| | (0.003) | (0.002) | (0.002) | (0.002) |
| Pre-treatment mean | 0.222 | 0.206 | 0.186 | 0.17 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no |
| Observations | 1,738,083 | 627,391 | 1,738,083 | 627,391 |

$^{+}$p<0.1;*p<0.05; **p<0.01; ***p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) and (3) show the DiD estimates for women aged 60–62 years and include age as control variable. Column (2) and (4) show the DiD estimates for women at age 59.
*Source: KBV, own calculations*

To corroborate our findings, we alter the definition of the outcome variables to test, whether noise of erroneous one-time diagnoses or miss-classifications by the medical personnel drive the results. For this exercise, we follow the so-called M2Q criterion and define a person in a calendar year to suffer from a mental disease only if she was diagnosed with such a condition in two quarters of the calendar year. Compared to the main specification, this alternative definition is more conservative because women who were only diagnosed in one quarter in a specific calendar year are not considered to suffer from the condition in this robustness check. Table A.5 in the Appendix shows the results for this exercise using the RDD specification. For both outcomes, the estimated treatment effects at age 59 and age 60–62 are positive, as in the main specification. Moreover, the effects on mood disorders are significant and confirm the order of magnitude of the main estimation results. For stress-related mental conditions, the main effect on 60–62 year old women is smaller and no longer significant when using the M2Q criterion. The effect on 59 year old women stays similar in size and significant. Using the M2Q criterion in the DiD setting gives similar results for mood disorders compared to the main DiD specification (Table A.6). The main effect on stress-related diseases stays similar as well. The age-59-effect decreases in magnitude but remains positive and statistically significant.

## 5.2 Results – Physical health

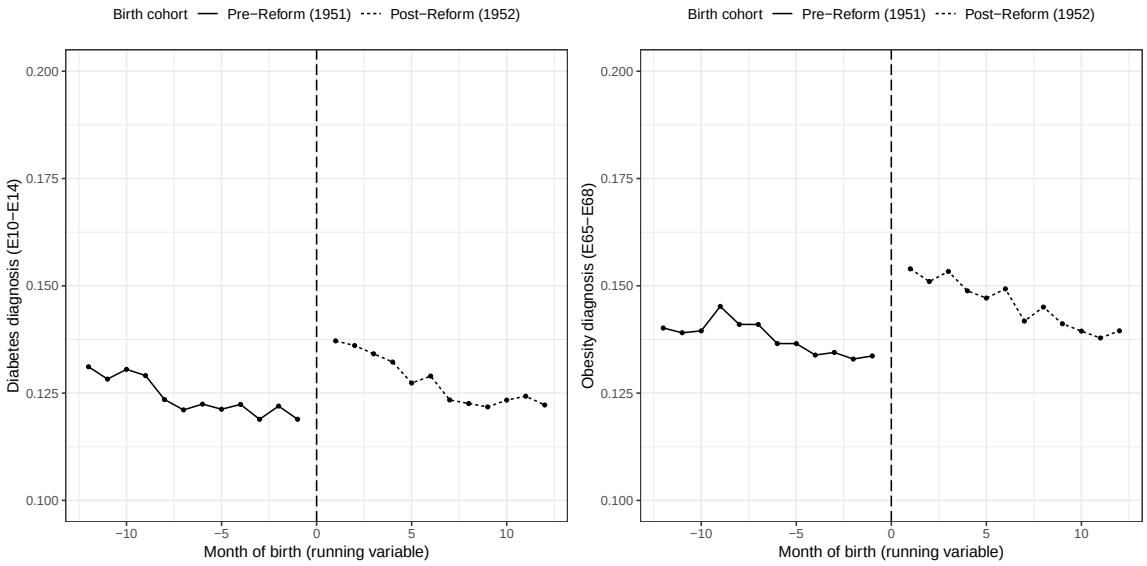
In the next step, we analyze the impact on physical health outcomes. We study three groups of physical health outcomes: Nutritional and metabolic diagnoses (diabetes and obesity), musculoskeletal diagnoses (arthrosis and dorsopathies), as well as circulatory and heart diagnoses (hypertension, ischaemic heart diseases and strokes).

### 5.2.1 Graphical analysis – Physical health

The graphical analysis reveals the importance of seasonality for the different physical health outcomes and provides mixed evidence about the effect of the 1999 pension reform on physical health. Regarding the nutritional and metabolic outcomes, we observe a strong seasonality pattern (Figure 3). Women born early in the year are more likely to be diagnosed with either of the diseases (diabetes and obesity) compared to women born later in the year. This is in line with findings from the medical literature that suggest that environmental reasons, exposure to sunlight, or nutrition are the main drivers for these differences (e.g., Kahn et al., 2009; Phillips and Young, 2000; Vaiserman and Khalangot, 2008; Wattie et al., 2008). Apart from seasonality, there seems to be no clear and strong jump at the reform cutoff.

For circulatory and heart diseases the pattern is similar: The graphical evidence does not indicate sizable reform effects (Figure 4). In line with Boland et al. (2015), we also find a strong seasonality pattern for hypertension whereas the pattern for heart and cerebrovascular diseases is rather stable. Musculoskeletal diagnoses also show quite strong seasonal fluctuations especially for arthrosis (Figure 5). However, there is some evidence of a positive reform effect on both musculoskeletal outcomes under study.

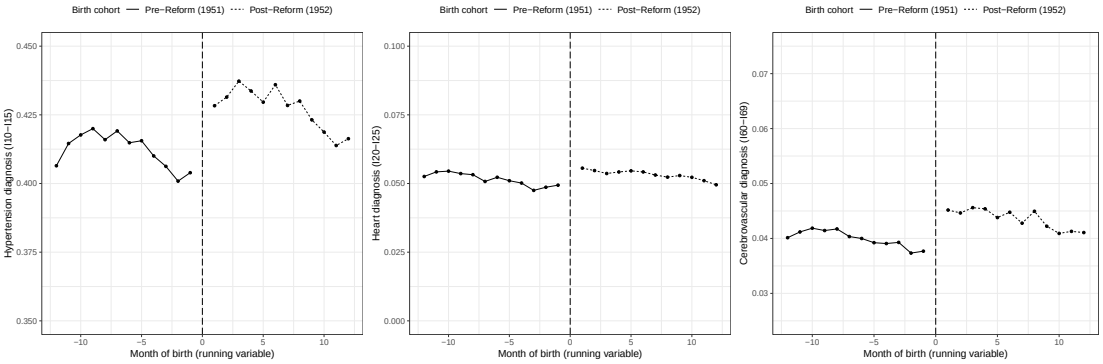
Figure 3: Metabolic/nutritional diagnoses by month of birth




*Notes:* The left figure presents the average share of women between age 60 and 62, who got a E10-E14 diagnosis in a given year, for each birth month. The right figure presents the average share of women between age 60 and 62, who got a E65-E68 diagnosis in a given year, for each birth month. The vertical lines represent the cutoff date (01/1952).

*Source: KBV, own calculations*

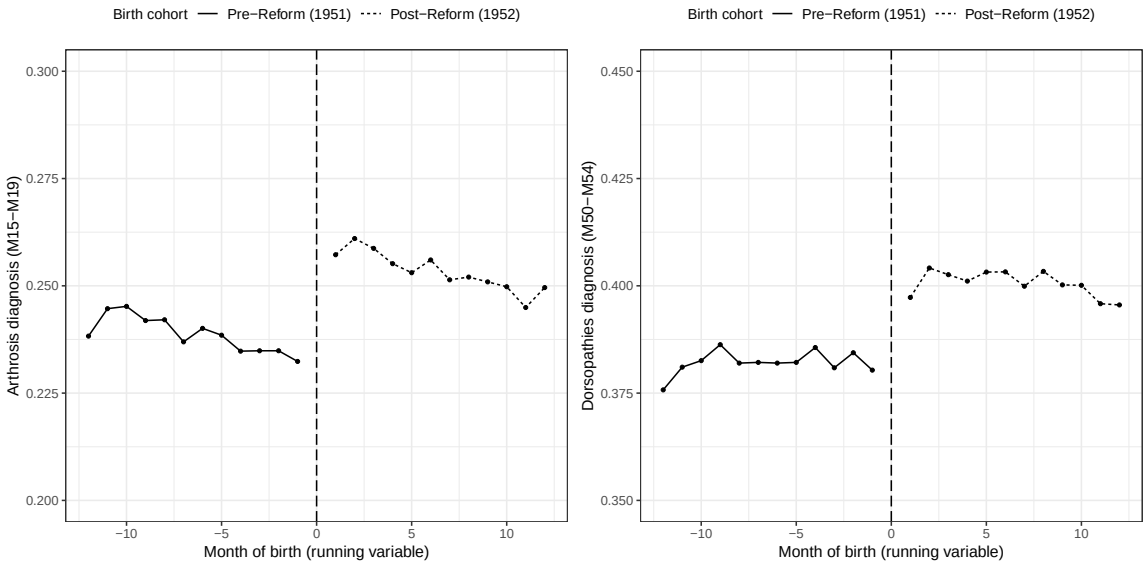
Figure 4: Circulatory/heart diagnoses by month of birth




*Notes:* The left figure presents the average share of women between age 60 and 62, who got a M15-M19 diagnosis in a given year, for each birth month. The figure in the middle the average share of women between age 60 and 62, who got a M50-M54 diagnosis in a given year, for each birth month. The right figure presents the average share of women between age 60 and 62, who got a I60-I69 diagnosis in a given year, for each birth month. The vertical lines represent the cutoff date (01/1952).

*Source: KBV, own calculations*

Figure 5: Musculoskeletal diagnoses by month of birth




*Notes:* The left figure presents the average share of women between age 60 and 62, who got a E10-E14 diagnosis in a given year, for each birth month. The right figure presents the average share of women between age 60 and 62, who got a E65-E68 diagnosis in a given year, for each birth month. The vertical lines represent the cutoff date (01/1952).

*Source: KBV, own calculations*

### 5.2.2 Regression results – Physical health

In the following, we present the causal effects estimated using the RDD and DiD approaches. We first cover metabolic and nutritional diseases, then, in the second subsection, we show the effects on circulatory and heart diseases, and the last subsection presents musculoskeletal diseases. Overall, the regression results largely support the insights from the graphical analysis.

#### Metabolic and nutritional diseases

We find positive effects of the reform on both health outcomes with the RDD (Table 3) and the DiD (Table 4), however the point estimates differ. Specifically for diabetes, the interaction effect in the DiD specification, which captures the effect of the pension reform, is relatively small (0.3 percentage points for the main effect and 0.5 percentage points for the age-59-effect) but still significant. In contrast, the point estimate in the RDD is considerably larger (2 percentage points for the main effect and 1.8 percentage points for the age-59-effect), which might be related to the longer observation period and the potential

influence of cohort effects.[23] This is consistent with the significant and positive effects in the placebo regression (Appendix Table A.7) that are most likely driven by cohort effects. Thus, the results suggest that the pension reform has a significant but small effect on the prevalence of diabetes.

The pattern is similar for obesity. Again, the point estimates of the two specifications differ, however the difference is less pronounced. According to the RDD specification, the pension reform increases the diagnosis of obesity by 1.9 percentage points or 13.7%. The effect measured at age 59 is similar in size (13.6%). The estimates of the DiD are slightly smaller but highly significant and positive.

Overall the results for diabetes and obesity are confirmed when using the more conservative definition of the outcome variable (M2Q-criterion) in the RDD and DiD specifications (Table A.19 and A.20). Furthermore, the results are robust to including only a linear cohort trend in the RDD analysis (Table A.10) and to including regional controls (Table A.13 and A.14).

Table 3: RDD results: Metabolic/nutritional diagnoses

| | *Diabetes* | | *Obesity* | |
|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 |
| $D_i$ | 0.020*** | 0.018*** | 0.019*** | 0.017*** |
| | (0.005) | (0.005) | (0.005) | (0.003) |
| $Birthmonths$ | −0.0004 | −0.0004 | −0.001 | −0.001* |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| $(Birthmonths)^2$ | 0.00004 | 0.00003 | −0.00004 | −0.00004 |
| | (0.00005) | (0.00005) | (0.00004) | (0.00004) |
| $D_i \times (Birthmonths)$ | −0.002$^+$ | −0.001 | −0.001 | −0.001 |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| $D_i \times (Birthmonths)^2$ | 0.0001 | 0.00003 | 0.0001** | 0.0001** |
| | (0.00005) | (0.00005) | (0.00004) | (0.00004) |
| Pre-treatment mean | 0.124 | 0.098 | 0.138 | 0.125 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no |
| Control for birth season | yes | yes | yes | yes |
| Observations | 3,429,155 | 1,235,612 | 3,429,155 | 1,235,612 |

$^+$p<0.1;*p<0.05; **p<0.01; ***p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) and (3) show the RDD estimates for women aged 60–62 years and include age and birth quarter dummies as control variables. Column (2) and (4) show the RDD estimates for women at age 59 and include birth quarter dummies as control variables. All regressions include linear and quadratic cohort trends in the running variable on both sides of the policy cut-off.
*Source: KBV, own calculations*

[23] Note again, given that the birth date is aggregated at a monthly level, it is difficult to identify a more flexible specification of the cohort trend.

Table 4: DiD: Metabolic/nutritional diagnoses

| | *Diabetes* | | *Obesity* | |
|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 |
| $Winter5152_i \times JanFebMar_i$ | 0.003* | 0.005** | 0.010*** | 0.008*** |
| | (0.002) | (0.002) | (0.001) | (0.001) |
| $Winter5152_i$ | 0.003* | 0.002 | 0.003** | $-0.001^{+}$ |
| | (0.001) | (0.002) | (0.001) | (0.001) |
| $JanFebMar_i$ | 0.013*** | 0.011*** | 0.009*** | 0.004*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Pre-treatment mean | 0.124 | 0.098 | 0.135 | 0.125 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no |
| Observations | 1,738,083 | 627,391 | 1,738,083 | 627,391 |

$^{+}$p<0.1;*p<0.05; **p<0.01; ***p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) and (3) show the RDD estimates for women aged 60–62 years and include age and birth quarter dummies as control variables. Column (2) and (4) show the RDD estimates for women at age 59 and include birth quarter dummies as control variables.
*Source: KBV, own calculations*

**Circulatory and heart diseases**

In the RDD and the DiD, we do not find significant effects on hypertension or ischeamic heart diseases for 60–62 year old women (Table 5). In contrast, the RDD specification shows that the probability of being diagnosed with a stroke increases significantly by 0.6 percentage points (15 percent, relative to the low pre-treatment mean of 4%), yet this finding is not confirmed by the DiD, which suggests that the reform effect in the RDD might also capture month of birth and cohort effects.[24] Interestingly, we find small, but significant, age-59-effects for all three diseases: 1.6 percentage points for hypertension, 0.7 percentage points for heart diseases, and 0.5 percentage points for strokes in the RDD analysis. The age-59-effects also persist in the DiD analysis for hypertension (2.4 percentage points) and strokes (0.2 percentage points).

The results for hypertension and heart diseases when using the M2Q-criterion are quite similar to our main specifications. However, for strokes, the estimates turn very small and insignificant in the RDD analysis when applying the M2Q criterion (Table A.21) while the estimates turn significant in the DiD specification (Table A.22). When additionally controlling for region, both specifications tend to produce more significant, but still small, point estimates (Table A.15 and A.16). Including only a linear cohort trend in the RDD, however, reveals similar results to the main specification (Table A.11).

[24] In the placebo test, we find a significant effect of similar magnitude, which again highlights the importance of cohort effects (Table A.8).

Overall, we do not find robust evidence that the increase in the retirement age increases the prevalence of the circulatory and heart diseases under study. Yet, our results show no evidence for a reduction in disease prevalence in response to the reform, since none of the effects is negatively significant.

Table 5: RDD results: Circulatory/heart diagnoses

| | *Hypertension* | | *Heart diagnosis* | | *Stroke* | |
|---|---|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 | Main | Age-59 |
| $D_i$ | 0.012 | 0.016* | 0.006$^{+}$ | 0.007* | 0.006*** | 0.005* |
| | (0.008) | (0.008) | (0.003) | (0.003) | (0.002) | (0.003) |
| $Birthmonths$ | −0.004** | −0.004** | −0.001 | −0.001$^{+}$ | −0.001* | 0.00004 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.0003) | (0.0005) |
| $(Birthmonths)^2$ | −0.0004*** | −0.0004*** | −0.00001 | −0.00005 | −0.0001* | 0.00001 |
| | (0.0001) | (0.0001) | (0.0001) | (0.00004) | (0.00002) | (0.00003) |
| $D_i \times (Birthmonths)$ | 0.009*** | 0.008*** | 0.0004 | 0.001 | 0.001* | −0.0002 |
| | (0.002) | (0.002) | (0.001) | (0.001) | (0.001) | (0.001) |
| $D_i \times (Birthmonths)^2$ | 0.0001 | 0.0001 | −0.00000 | 0.00002 | 0.00002 | −0.00002 |
| | (0.0001) | (0.0001) | (0.00004) | (0.00003) | (0.00002) | (0.00002) |
| Pre-treatment mean | 0.412 | 0.347 | 0.052 | 0.041 | 0.04 | 0.028 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no | yes | no |
| Control for birth season | yes | yes | yes | yes | yes | yes |
| Observations | 3,429,155 | 1,235,612 | 3,429,155 | 1,235,612 | 3,429,155 | 1,235,612 |

$^{+}$p<0.1;*p<0.05; **p<0.01; ***p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1),(3) and (5) show the RDD estimates for women aged 60–62 years and include age and birth quarter dummies as control variables. Column (2), (4) and (6) show the RDD estimates for women at age 59 and include birth quarter dummies as control variables. All regressions include linear and quadratic cohort trends in the running variable on both sides of the policy cut-off.
*Source: KBV, own calculations*

Table 6: DiD: Circulatory/heart diagnoses

| | *Hypertension* | | *Heart diagnosis* | | *Stroke diagnosis* | |
|---|---|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 | Main | Age-59 |
| $Winter5152_i \times JanFebMar_i$ | 0.007$^{+}$ | 0.024*** | 0.0003 | −0.0003 | 0.001 | 0.002*** |
| | (0.004) | (0.004) | (0.001) | (0.001) | (0.001) | (0.001) |
| $Winter5152_i$ | 0.013*** | 0.001 | 0.001 | 0.001 | 0.003*** | 0.001*** |
| | (0.002) | (0.002) | (0.001) | (0.0004) | (0.001) | (0.0004) |
| $JanFebMar_i$ | 0.022*** | 0.009** | 0.006*** | 0.005*** | 0.006*** | 0.004*** |
| | (0.003) | (0.003) | (0.001) | (0.001) | (0.0005) | (0.0004) |
| Pre-treatment mean | 0.402 | 0.343 | 0.051 | 0.04 | 0.038 | 0.027 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no | yes | no |
| Observations | 1,738,083 | 627,391 | 1,738,083 | 627,391 | 1,738,083 | 627,391 |

$^{+}$p<0.1;*p<0.05; **p<0.01; ***p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1), (3) and (5) show the DiD estimates for women aged 60–62 years and include age as control variable. Column (2), (4) and (6) show the DiD estimates for women at age 59.
*Source: KBV, own calculations*

### Musculoskeletal diseases

Results for musculoskeletal diseases (arthrosis and dorsopathies) indicate positive effects of the pension reform. The RDD estimation reveals a significant increase in the risk of arthrosis by 2.1 percentage points, corresponding to a relative effect of 9 percent in relation to the pre-treatment mean (Table 7). The age-59-effect is of similar size and significance.

The effects on arthrosis in the placebo analysis (Table A.9) are smaller and only significant at the ten percent level suggesting that the RDD model for this outcome does not sufficiently capture month of birth effects. In line with this, the DiD results are considerably smaller (0.8 percentage points for the main analysis (3.4%) and 0.7 percentage points for the age-59-effect) and less significant (Table 8).

The effects on dorsopathies are less clear: The RDD analysis shows an increase of dorsopathies of one percentage point, which is only significant at the 10% level and a significant 1.3 percentage point increase for the age-59-effect. In the DiD analysis both effects are highly significant. The main effect is similar in size to the RDD effect (0.8 percentage points; 2.1%), while the age-59-effect is considerably larger (2.1 percentage points; 6%). The placebo analysis reveals insignificant effects for dorsopathies.

Our robustness checks mostly confirm the results of our main analysis for both outcome variables: Including only a linear cohort trend in the RDD analysis leads to similar effects for arthrosis and insignificant effects for dorsopathies compared to the main RDD specification (Table A.12). Controlling for region reveals very similar results as the main analysis (Table A.17 and A.18). Using the M2Q-criterion for the definition of the outcome variables gives similar results for arthrosis, while the estimates for dorsopathies become larger and more significant in the RDD analysis and smaller in the DiD analysis (Table A.23 and A.24). Thus, the effects on arthrosis turn out to be more robust than the results on dorsopathies.

Table 7: RDD results: Musculoskeletal diagnoses

| | *Arthrosis* | | *Dorsopathies* | |
|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 |
| $D_i$ | $0.021^{***}$ | $0.021^{***}$ | $0.010^{+}$ | $0.013^{*}$ |
| | (0.005) | (0.005) | (0.006) | (0.006) |
| $Birthmonths$ | $-0.003^{***}$ | $-0.004^{***}$ | $-0.001^{+}$ | $-0.004^{***}$ |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| $(Birthmonths)^2$ | $-0.0002^{**}$ | $-0.0003^{***}$ | $-0.0002^{**}$ | $-0.0004^{***}$ |
| | (0.0001) | (0.0001) | (0.0001) | (0.0001) |
| $D_i \times (Birthmonths)$ | $0.004^{**}$ | $0.006^{***}$ | $0.005^{***}$ | $0.008^{***}$ |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| $D_i \times (Birthmonths)^2$ | 0.0001 | 0.0001 | −0.00004 | 0.0001 |
| | (0.0001) | (0.0001) | (0.0001) | (0.0001) |
| Pre-treatment mean | 0.239 | 0.203 | 0.382 | 0.354 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no |
| Control for birth season | yes | yes | yes | yes |
| Observations | 3,429,155 | 1,235,612 | 3,429,155 | 1,235,612 |

$^{+}$p<0.1;$^{*}$p<0.05; $^{**}$p<0.01; $^{***}$p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) and (3) show the RDD estimates for women aged 60–62 years and include age and birth quarter dummies as control variables. Column (2) and (4) show the RDD estimates for women at age 59 and include birth quarter dummies as control variables. All regressions include linear and quadratic cohort trends in the running variable on both sides of the policy cut-off.
*Source: KBV, own calculations*

Table 8: DiD: Musculoskeletal diagnoses

| | *Arthrosis* | | *Dorsopathies* | |
|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 |
| $Winter5152_i \times JanFebMar_i$ | $0.008^{**}$ | $0.007^{*}$ | $0.008^{**}$ | $0.021^{***}$ |
| | (0.003) | (0.003) | (0.003) | (0.004) |
| $Winter5152_i$ | $0.008^{***}$ | $0.004^{+}$ | $0.013^{***}$ | 0.002 |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| $JanFebMar_i$ | $0.017^{***}$ | $0.014^{***}$ | $0.011^{***}$ | 0.001 |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Pre-treatment mean | 0.235 | 0.201 | 0.374 | 0.352 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no |
| Observations | 1,738,083 | 627,391 | 1,738,083 | 627,391 |

$^{+}$p<0.1;$^{*}$p<0.05; $^{**}$p<0.01; $^{***}$p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) and (3) show the DiD estimates for women aged 60–62 years and include age as control variable. Column (2) and (4) show the DiD estimates for women at age 59.
*Source: KBV, own calculations*

## 5.3 Results – Multiple hypothesis testing

Given the relatively large number of health outcomes used in the analysis, we perform multiple-hypothesis-tests using a Bonferroni correction adjustments procedure to the single physical and mental health outcomes. We correct for nine hypotheses (number of diagnoses considered).[25] The underlying regression are the RDD and DiD specifications from

[25]We choose the Bonferroni correction as our preferred method since this is the most conservative correction procedure. We implement this by using the R-package *p.adjust*.

Equation 2 including linear and quadratic cohort trends and Equation 3. The multiple hypothesis method confirms our findings of rejecting the null hypothesis for stress-related diseases, mood disorders, obesity, and arthrosis. The results are shown in the Appendix (Table A.25 and A.26).

## 5.4 Results – Health care consumption

In this section, we turn to the effects of the 1999 pension reform on doctor visits. We measure doctor visits as doctor cases, aggregated at the calendar year level (official term: "Artzfälle"). One doctor case is defined as a treatment of an insured person by a doctor in a quarter, billed to one public health insurance fund.[26] Thus, if a person visits two different doctors in a quarter, she has two doctor cases in that specific quarter.[27] We aggregate quarterly cases to the calendar year level, thus counting the number of quarterly doctor cases per year. This means that a patient who visits every quarter only one and the same doctor would have a yearly count of four doctor cases, irrespective of the actual number of visits to this doctor per quarter.[28]

### 5.4.1 Graphical analysis – Health care consumption

Figure 6 shows the average number of doctor visits per year for each birth month around the reform cutoff. There is a jump of almost 0.5 doctor visits at the threshold. However, it is important to take into account that the number of doctor visits also varies by about 0.25 doctor visits over birth months on both sides of the discontinuity. Thus, a formal estimation of the causal effect needs to control for month of birth effects and trends.

### 5.4.2 Regression results – Health care consumption

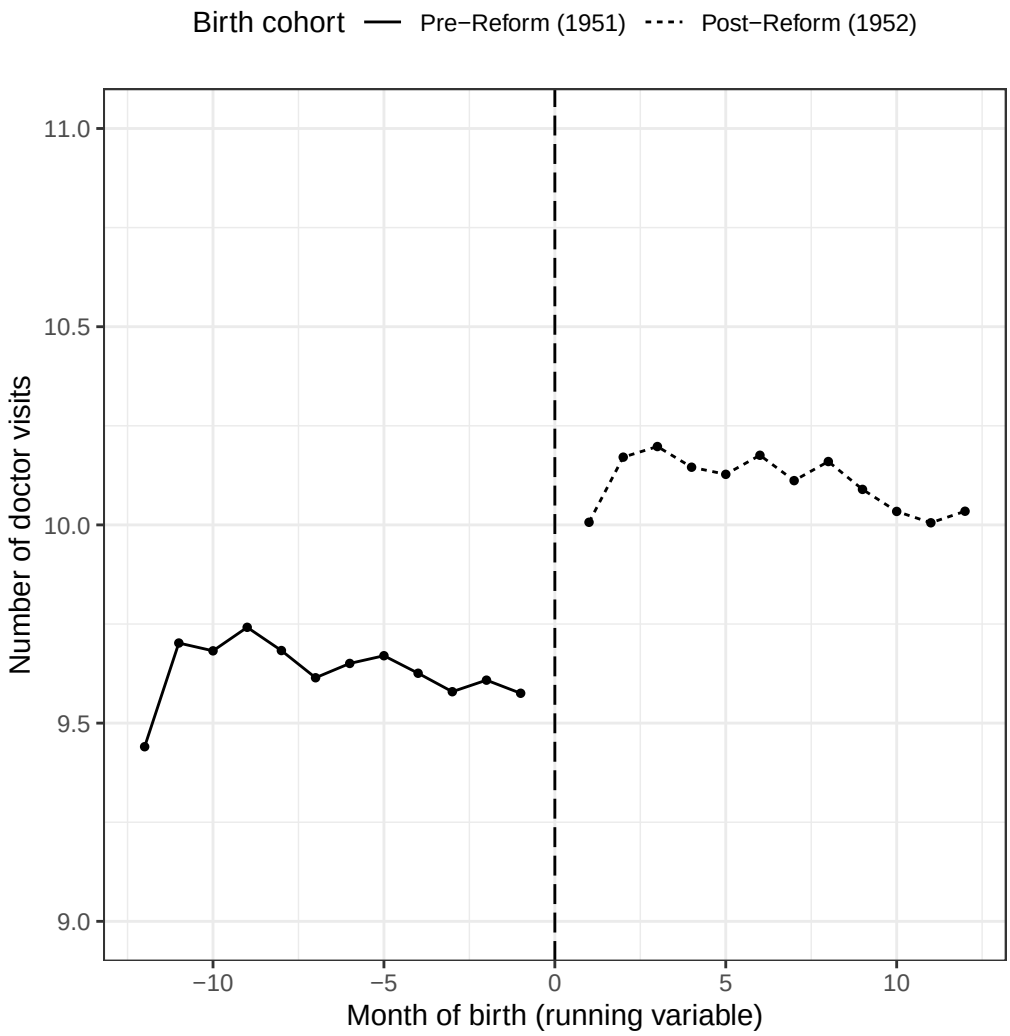
In the RDD, we find a marginally significant increase in doctor visits (Table 9). In relation to the 1951 cohort average, this effect amounts to about 2.6 percent. The effect in the DiD analysis is slightly smaller and also significant only at the 10% level (Table 10). Interestingly, the effect for women aged 59 is more than double the size of the main effect on 60–62 year old women and highly significant. The number of doctor visits increases due to the reform

[26] Since doctor cases are recorded this way in the data, we do not have the possibility to define the variable differently for our application.

[27] If she visits only one doctor but switches health insurance providers, she would also be assigned two doctor visits. However, since only 3% of women in our sample switch health insurance providers, this issue is negligible.

[28] This measure does not capture all doctor visits, thus the observed difference between the two birth cohorts is a lower-bound estimate of the effect of the reform on healthcare consumption.

Figure 6: Number of doctor visits by month of birth




*Notes:* The figure presents the average number of annual doctor visits of women between age 60 and 62 for each birth month. The vertical lines represent the cutoff date (01/1952).
*Source: KBV, own calculations*

by more than half a doctor visit (Table 9). In relative terms, this effect amounts to about 6.6 percent in relation to the cohort 1951 average. Using a DiD estimation approach, this age-59-effect stays at a similar magnitude.

For both age groups (59 and 60–62 year old women), the RDD placebo effects are considerably smaller and insignificant, as shown in Table A.27. Further, the causal effects are fairly robust to using a linear RDD specification and adding regional controls (Table A.28, A.29 and A.30 in the Appendix) as well. The point estimates slightly decrease and the effect on 60–62 year old women turns insignificant when using a linear RDD specification.

In sum, doctor visits increase due to the 1999 pension reform, in particular at age 59. The reasons for this significant age-59-effect can be manifold. One possibility is that women born in 1952 might try to enter retirement early via the disability/invalidity pension schemes in the absence of the old age pension scheme for women. Disability pension is only granted if a person has a reduced earnings capacity and the process is strict. Doctor visits might be indicative of cohort 1952 trying to prove reduced earnings capacity for medical reasons. However, Geyer and Welteke (2021) show that there is no effect of the 1999 pension reform on actual disability pension claims. Thus, despite a possible increase in applications and related doctor visits, the actual claiming behavior is not very different between cohort 1951 and 1952.

Another possible reason for differences in healthcare consumption between the cohorts could be different time budgets and time-use decisions in response to the reform. Eligible women born in cohort 1951 know that they can retire at age 60. Thus, they might delay time consuming activities, like (non-urgent) doctor visits from age 59 to their retirement a couple of months later, resulting in fewer doctor visits at age 59. In contrast, women born in 1952 expect to retire only years later, which means that they are less likely to shift time consuming activities from age 59 to age 60. Thus, women born 1952 could have more doctor visits at age 59 than women born in 1951.

Table 9: RDD results: Number of doctor visits

| | *Dependent variable: Doctor visits* | |
|---|---|---|
| | Main | Age-59 |
| $D_i$ | $0.252^{+}$ | $0.565^{***}$ |
| | (0.150) | (0.150) |
| $Birthmonths$ | $-0.055^{*}$ | $-0.087^{***}$ |
| | (0.024) | (0.024) |
| $(Birthmonths)^2$ | $-0.006^{**}$ | $-0.007^{**}$ |
| | (0.002) | (0.002) |
| $D_i \times (Birthmonths)$ | $0.141^{**}$ | $0.162^{***}$ |
| | (0.043) | (0.043) |
| $D_i \times (Birthmonths)^2$ | 0.0004 | 0.001 |
| | (0.002) | (0.002) |
| Pre-treatment mean | 9.631 | 8.606 |
| Age group included | 60-62 years | 59 years |
| Control for age | yes | no |
| Control for birth season | yes | yes |
| Observations | 3,429,155 | 1,235,612 |

$^{+}$p<0.1;$^{*}$p<0.05; $^{**}$p<0.01; $^{***}$p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) shows the RDD estimates for women aged 60–62 years and include age and birth quarter dummies as control variables. Column (2) shows the RDD estimates for women at age 59 and include birth quarter dummies as control variables. Both regressions include linear and quadratic cohort trends in the running variable on both sides of the policy cut-off.
*Source: KBV, own calculations*

Table 10: DiD: Number of doctor visits

| | *Dependent variable: Doctor visits* | |
|---|---|---|
| | Main | Age-59 |
| $Winter5152_i \times JanFebMar_i$ | $0.174^{+}$ | $0.509^{***}$ |
| | (0.093) | (0.091) |
| $Winter5152_i$ | $0.344^{***}$ | $0.108^{**}$ |
| | (0.027) | (0.038) |
| $JanFebMar_i$ | $0.363^{***}$ | $0.146^{*}$ |
| | (0.077) | (0.071) |
| Pre-treatment mean | 9.43 | 8.52 |
| Age group included | 60-62 years | 59 years |
| Control for age | yes | no |
| Observations | 1,738,083 | 627,391 |

$^{+}$p<0.1;$^{*}$p<0.05; $^{**}$p<0.01; $^{***}$p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) shows the DiD estimates for women aged 60–62 years and include age as control variable. Column (2) shows the DiD estimates for women at age 59.
*Source: KBV, own calculations*

### 5.5 Post-Employment Effects

In the final Section we analyze if the effective increase in the retirement age from 60 to 63 has an effect on health outcomes of women aged 63 and older. These results can be interpreted as indirect or medium run effects of the pension reform since at these ages women of both cohorts have access to retirement and are thus not directly affected by the pension reform. For the analysis we consider again the RDD and the DiD estimation. Results of the DiD are presented in the main text, the results of the RDD including the related Placebo estimations can be found in Tables A.31 and A.32 in the Appendix. The effect size and the significance of the estimators are quite similar in the main RDD specification and the Placebo specification. This suggests that the estimated health effects for this age groups are strongly affected by month of birth effects. For a more detailed interpretation and quantification we therefore turn to the effects DiD estimation which directly accounts for month of birth effects.

In line with the results for the RDD, these findings suggest that the increase of the retirement age has a smaller impact on medium run health outcomes of women (Table 11). We only find significant effects below the 5% level for mood disorders, arthrosis and dorsopathies and obesity. Recall, for women aged 60-62 years, we have documented significant and robust evidence for an increase in the prevalence of stress-related diseases, mood disorders, arthrosis and obesity. The effects on mood disorders, arthrosis and obesity seem to persist also in the medium run. However, effect sizes are smaller (2.4% vs. 4.8% for mood disorders, 2% vs. 2.1% for arthrosis and 4% vs. 7.4% for obesity). This pattern suggests that the detrimental health effects of the increase in retirement age are strongest for women directly affected by the pension reform. However, the majority of the effects persist at least until age 65. As the effect sizes decrease with age, our results indicate that the differences in health outcomes between the two cohorts fade out at older ages, i.e. in the long run. A formal analysis of the long run effects remains for future research when data are available.

Table 11: DiD results: 63-65 year olds

| | *Dependent variable:* | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Stress-related | Mood disorder | Arthrosis | Dorsopathies | Diabetes | Obesity | Hypertension | Heart | Strokes | Doc. visits |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| $Winter5152_i \times JanFebMar_i$ | 0.003 | $0.005^{**}$ | $0.006^{*}$ | $0.005^{*}$ | $0.005^{+}$ | $0.007^{***}$ | 0.005 | −0.001 | 0.001 | 0.110 |
| | (0.003) | (0.002) | (0.002) | (0.002) | (0.003) | (0.002) | (0.003) | (0.001) | (0.001) | (0.075) |
| $Winter5152_i$ | $0.012^{***}$ | $0.009^{***}$ | $0.011^{***}$ | $0.014^{***}$ | −0.0004 | $0.008^{***}$ | $0.012^{***}$ | 0.001 | $0.003^{***}$ | $0.345^{***}$ |
| | (0.002) | (0.001) | (0.001) | (0.001) | (0.002) | (0.001) | (0.002) | (0.001) | (0.0003) | (0.022) |
| $JanFebMar_i$ | $0.009^{**}$ | $0.012^{***}$ | $0.020^{***}$ | $0.014^{***}$ | $0.013^{***}$ | $0.016^{***}$ | $0.023^{***}$ | $0.008^{***}$ | $0.008^{***}$ | $0.419^{***}$ |
| | (0.003) | (0.001) | (0.002) | (0.002) | (0.002) | (0.001) | (0.003) | (0.001) | (0.0002) | (0.063) |
| Pre-treatment mean | 0.247 | 0.208 | 0.293 | 0.412 | 0.166 | 0.167 | 0.498 | 0.07 | 0.061 | 10.943 |
| Observations | 1,543,601 | 1,543,601 | 1,543,601 | 1,543,601 | 1,543,601 | 1,543,601 | 1,543,601 | 1,543,601 | 1,543,601 | 1,543,601 |

$^{+}$p<0.1;$^{*}$p<0.05; $^{**}$p<0.01; $^{***}$p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. All regressions include the cohort indicator, the reform indicator and their interaction term.
*Source: KBV, own calculations*

# 6 Conclusion

This paper provides novel insights about the causal effects of an increasing retirement age on a multi-dimensional and comprehensive set of health outcomes. For the identification, we exploit a large exogenous increase in the ERA for women in Germany. In particular, we focus on the 1999 pension reform that increases the ERA by three years for women born after December 1951.

Previous literature is inconclusive in terms of magnitude and direction of the overall effects of retirement on health. Earlier work often relies on survey data that often include subjective and broad health measures. However, health is multi-dimensional and the effects of retirement (reforms) on different health outcomes might, therefore, go into different directions.

Our analyses are based on administrative data from German health insurance funds that include health diagnoses of all publicly insured individuals. We use a sample of women born between 1950 and 1952 who are observed between 2009 and 2018. The data contain all diagnoses in outpatient care during the observation period. Specifically, we identify and consider relevant diagnoses and measures within three dimensions of health outcomes: mental health, physical health, and healthcare consumption.

The sharp discontinuity in the ERA by cohorts induced by the pension reform allows us to analyze the health effects using an RDD as well as a DiD approach. The combination of both approaches enables us to address month of birth effects in the prevalence of diseases as well as cohort trends.

The empirical findings reflect the multi-dimensionality of health outcomes but allow for deriving two broader conclusions. We provide evidence that the increase in the retirement age has a negative effect on health outcomes as the prevalence of several diagnoses, e.g., mental health, arthrosis, and obesity, increases. In contrast, we do not find support for an improvement in health related to a prolonged working life since there is no significant evidence of a reduction in the prevalence of any health outcome we consider. These findings hold for both identification strategies, are robust to sensitivity checks, and do not change when correcting for multiple hypothesis testing.

More precisely, we find that the pension reform increased the prevalence of both groups of mental diseases in 60–62 year old women. The effect sizes range from 3.6 to 4.8 percent

for stress-related diseases and from 4.8 to 8.3 percent for mood disorders relative to the respective pre-treatment means. The effects for 59 year old women are of similar magnitude and significance. Considering that only about 60% of the women were eligible for the old age pension for women (Geyer and Welteke, 2021), the reform effect on eligible women turns out even larger. For example, scaling the ITT effects with this eligibility rate in a back-of-the-envelope calculation, the effects on stress-related diseases for 60–62 year old women range between 6.1 and 8.1, the effects on mood-disorders between 8.1 and 14 percent.

Within the physical health dimension, our ITT estimates suggest that raising the retirement age increases the prevalence of arthrosis and obesity at age 60–62 years as well as 59 years. For other physical health outcomes our results are less clear but, as mentioned above, we do not find significant evidence for an improvement in physical health in response to the reform. Furthermore, we find a significant increase in healthcare consumption for 59 year olds following the reform.

Additional analyses on post-employment effects suggest that the effects on mood disorders, arthrosis and obesity persist also in the medium run. However, effect sizes are smaller for 63–65 year old women compared to 60–62 year old women suggesting that the detrimental health effects do last into retirement but at a lower level.

Increasing the retirement age is controversially discussed in politics and society. Our results inform this debate, as health implications are an important aspect. For future pension reforms, policy makers should keep in mind that a prolonged working life might have considerable negative health consequences, particularly for mental health. Further research is needed to identify the mechanisms behind our findings. The main treatment effect results from a prolonged duration in the labor market. Therefore, the reform effect operates through different channels which we cannot differentiate with the data at hand. The majority of treated women stays longer in employment which might affect health. However, the prolonged status in unemployment could as well impact health. Targeted health programs that support different groups in the labor market in dealing with stress or providing sport and exercise programs could counteract the negative effects. Another solution might be to extend old-age-part-time work to smooth the transition into retirement.

In future research, it would be important to assess whether these multi-dimensional health effects further differ by socioeconomic characteristics. The literature shows that such characteristics may matter for the health effects of retirement (see e.g., Etgeton and Hammer-

schmid, 2019, and references therein). The data we use only includes very limited individual characteristics beyond health. Thus, with the data at hand, assessing the socioeconomic gradient is not possible. Furthermore, it would be interesting to analyze the effects at ages older that 65 years to understand how persistent the effects are.

## References


Atalay, K. and Barrett, G. F. (2014), 'The causal effect of retirement on health: New evidence from Australian pension reform', *Economics Letters* **125**(3), 392–395.

Atalay, K., Barrett, G. F. and Staneva, A. (2019), 'The effect of retirement on elderly cognitive functioning', *Journal of health economics* **66**, 37–53.

Behncke, S. (2012), 'Does retirement trigger ill health?', *Health Economics* **21**(3), 282–300.

Belloni, M., Meschi, E. and Pasini, G. (2016), 'The effect on mental health of retiring during the economic crisis', *Health economics* **25**, 126–140.

Bloemen, H., Hochguertel, S. and Zweerink, J. (2017), 'The causal effect of retirement on mortality: Evidence from targeted incentives to retire early', *Health Economics* **26**(12), e204–e218.

Boland, M. R., Shahn, Z., Madigan, D., Hripcsak, G. and Tatonetti, N. P. (2015), 'Birth month affects lifetime disease risk: A phenome-wide method', *Journal of the American Medical Informatics Association* **22**(5), 1042–1053.

Brockmann, H., Müller, R. and Helmert, U. (2009), 'Time to retire–time to die? A prospective cohort study of the effects of early retirement on long-term survival', *Social Science & Medicine* **69**(2), 160–164.

Celidoni, M. and Rebba, V. (2017), 'Healthier lifestyles after retirement in europe? evidence from share', *The European Journal of Health Economics* **18**(7), 805–830.

Charles, K. K. (2004), 'Is retirement depressing? Labor force inactivity and psychological well-being in later life', *Research in Labor Economics* **23**(2004), 269–299.

Coe, N. B. and Zamarro, G. (2011), 'Retirement effects on health in Europe', *Journal of Health Economics* **30**(1), 77–86.

Doblhammer, G. and Vaupel, J. W. (2001), 'Lifespan depends on month of birth', *Proceedings of the National Academy of Sciences* **98**(5), 2934–2939.

Eibich, P. (2015), 'Understanding the effect of retirement on health: Mechanisms and heterogeneity', *Journal of Health Economics* **43**, 1–12.

Etgeton, S., Fischer, B. and Ye, H. (2021), The effect of increasing retirement age on

households' savings and consumption expenditures, CRC TR 224 Discussion Paper Series, University of Bonn and University of Mannheim, Germany.

Etgeton, S. and Hammerschmid, A. (2019), 'The effect of early retirement on health: Evidence from a German pension reform', *Mimeo* .

Fischer, B. and Müller, K.-U. (2020), 'Time to care? the effects of retirement on informal care provision', *Journal of Health Economics* **73**, 102350.
**URL:** *http://www.sciencedirect.com/science/article/pii/S0167629619306071*

Fitzpatrick, M. D. and Moore, T. J. (2018), 'The mortality effects of retirement: Evidence from social security eligibility at age 62', *Journal of Public Economics* **157**, 121–137.

Frimmel, W. and Pruckner, G. J. (2020), 'Retirement and healthcare utilization', *Journal of Public Economics* **184**, 104146.

Garrouste, C. and Perdrix, E. (2021), 'Is there a consensus on the health consequences of retirement? a literature review', *Journal of Economic Surveys* .

Geyer, J., Haan, P., Hammerschmid, A. and Peters, M. (2020), 'Labor market and distributional effects of an increase in the retirement age', *Labour Economics* pp. 101–117.

Geyer, J. and Welteke, C. (2021), 'Closing routes to retirement for women: How do they respond?', *Journal of Human Resources* **56**(1), 311–341.

Godard, M. (2016), 'Gaining weight through retirement? Results from the SHARE survey', *Journal of Health Economics* **45**, 27–46.

Gohl, N., Haan, P., Kurz, E. and Weinhardt, F. (2020), Working life and human capital investment: Causal evidence from pension reform, IZA Discussion Paper 12891.

Gorry, A., Gorry, D. and Slavov, S. N. (2018), 'Does retirement improve health and life satisfaction?', *Health economics* **27**(12), 2067–2086.

Grip, A. d., Lindeboom, M. and Montizaan, R. (2011), 'Shattered dreams: The effects of changing the pension system late in the game', *The Economic Journal* **122**(559), 1–25.

Hagen, J. (2018), 'The effects of increasing the normal retirement age on health care utilization and mortality', *Journal of Population Economics* **31**(1), 193–234.

Hallberg, D., Johansson, P. and Josephson, M. (2015), 'Is an early retirement offer good for

your health? Quasi-experimental evidence from the army', *Journal of Health Economics* **44**, 274–285.

Heller-Sahlgren, G. (2017), 'Retirement blues', *Journal of Health Economics* **54**, 66–78.

Hernaes, E., Markussen, S., Piggott, J. and Vestad, O. L. (2013), 'Does retirement age impact mortality?', *Journal of Health Economics* **32**(3), 586–598.

Horner, E. M. and Cullen, M. R. (2016), 'The impact of retirement on health: Quasi-experimental methods using administrative data', *BMC Health Services Research* **16**(68). https://doi.org/10.1186/s12913-016-1318-5.

Kahn, H. S., Morgan, T. M., Case, L. D., Dabelea, D., Mayer-Davis, E. J., Lawrence, J. M., Marcovina, S. M., Imperatore, G., for Diabetes in Youth Study Group, S. et al. (2009), 'Association of type 1 diabetes with month of birth among US youth: The SEARCH for diabetes in youth study', *Diabetes care* **32**(11), 2010–2015.

Kemptner, D., Jürges, H. and Reinhold, S. (2011), 'Changes in compulsory schooling and the causal effect of education on health: Evidence from germany', *Journal of health economics* **30**(2), 340–354.

Kuhn, A., Staubli, S., Wuellrich, J.-P. and Zweimüller, J. (2019), 'Fatal attraction? Extended unemployment benefits, labor force exits, and mortality', *Journal of Public Economics* **191**, 104087. https://doi.org/10.1016/j.jpubeco.2019.104087.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0047272719301483*

Kuusi, T., Martikainen, P. and Valkonen, T. (2020), 'The influence of old-age retirement on health: Causal evidence from the Finnish register data', *The Journal of the Economics of Ageing* p. 100257.

Leimer, B. (2017), 'No 'honeymoon phase'– whose health benefits from retirement and when', *Gutenberg School of Management and Economics & Research Unit "Interdisciplinary Public Policy" Discussion Paper Series* (1718).

Mazzonna, F. and Peracchi, F. (2017), 'Unhealthy retirement?', *Journal of Human Resources* **52**(1), 128–151.

Nielsen, N. F. (2019), 'Sick of retirement?', *Journal of Health Economics* **65**, 133–152.
**URL:** *https://www.sciencedirect.com/science/article/pii/S0167629618300080*

Nishimura, Y., Oikawa, M. and Motegi, H. (2018), 'What explains the difference in the effect of retirement on health? Evidence from global aging data', *Journal of Economic Surveys* **32**(3), 792–847.

OECD (2019), Germany: Country Health Profile 2019, State of Health in the EU, Technical report, OECD Publishing, Paris/European Observatory on Health Systems and Policies, Brussels.

Phillips, D. and Young, J. B. (2000), 'Birth weight, climate at birth and the risk of obesity in adult life', *International journal of obesity* **24**(3), 281–287.

Pilipiec, P., Groot, W. and Pavlova, M. (2020), 'The effect of an increase of the retirement age on the health, well-being, and labor force participation of older workers: a systematic literature review', *Journal of Population Ageing* pp. 1–45.

RKI (2010), Faktenblätter – Inanspruchnahme von Leistungen des Gesundheitssystems: Arztbesuch, Technical report, Robert-Koch-Institut, Epidemiologie und Gesundheitsberichterstattung.

Rohwedder, S. and Willis, R. J. (2010), 'Mental retirement', *Journal of Economic Perspectives* **24**(1), 119–138.

Rose, L. (2020), 'Retirement and health: Evidence from england', *Journal of Health Economics* **73**, 102352.

Schönberg, U. and Ludsteck, J. (2014), 'Expansions in maternity leave coverage and mothers' labor market outcomes after childbirth', *Journal of Labor Economics* **32**(3), 469–505.

Shai, O. (2018), 'Is retirement good for men's health? evidence using a change in the retirement age in israel', *Journal of health economics* **57**, 15–30.

Vaiserman, A. and Khalangot, M. (2008), 'Similar seasonality of birth in type 1 and type 2 diabetes patients: a sign for common etiology?', *Medical hypotheses* **71**(4), 604–605.

Van den Berg, G. J. and Siflinger, B. M. (2020), The effects of day care on health during childhood: evidence by age, IZA Discussion Paper 12891.

van Ours, J. C. and Picchio, M. (2020), 'Mental health effects of retirement', *De Economist* **168**(3), 419–452.

Wattie, N., Ardern, C. I. and Baker, J. (2008), 'Season of birth and prevalence of overweight and obesity in Canada', *Early human development* **84**(8), 539–547.

Zhang, Y., Salm, M. and van Soest, A. (2018), 'The effect of retirement on healthcare utilization: Evidence from China', *Journal of Health Economics* **62**, 165–177.

# Appendices

## A Additional results

### A.1 Mental health

#### A.1.1 Placebo-Reform (01/1951)

Table A.1: Placebo-Reform (01/1951): Mental diagnoses

| | *Stress-related diagnosis* | | *Mood disorder* | |
|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 |
| $D_i$ | 0.005 | 0.0004 | 0.004 | −0.004 |
| | (0.007) | (0.009) | (0.005) | (0.007) |
| $Birthmonths$ | −0.004* | −0.004* | −0.003** | −0.003* |
| | (0.001) | (0.002) | (0.001) | (0.001) |
| $(Birthmonths)^2$ | −0.0004** | −0.0004* | −0.0002*** | −0.0003*** |
| | (0.0001) | (0.0002) | (0.0001) | (0.0001) |
| $D_i \times (Birthmonths)$ | 0.008*** | 0.009*** | 0.007*** | 0.007*** |
| | (0.002) | (0.003) | (0.001) | (0.002) |
| $D_i \times (Birthmonths)^2$ | 0.0001 | 0.00001 | −0.00004 | −0.00003 |
| | (0.0001) | (0.0001) | (0.0001) | (0.0001) |
| Pre-treatment mean | 0.218 | 0.204 | 0.182 | 0.169 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no |
| Control for birth season | yes | yes | yes | yes |
| Observations | 3,524,843 | 1,275,284 | 3,524,843 | 1,275,284 |

$^{+}$p<0.1;$^{*}$p<0.05; $^{**}$p<0.01; $^{***}$p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) and (3) show the RDD estimates for women aged 60–62 years and include age and birth quarter dummies as control variables. Column (2) and (4) show the RDD estimates for women at age 59 and include birth quarter dummies as control variables. All regressions include linear and quadratic cohort trends in the running variable on both sides of the policy cut-off. The cut-off date is 01/1951.
*Source: KBV, own calculations*

### A.1.2 Linear cohort trend

Table A.2: Linear cohort trend results: Mental diagnoses

| | *Stress-related* | | *Mood disorder* | |
|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 |
| $D_i$ | 0.011* | 0.010 | 0.016* | 0.013 |
| | (0.005) | (0.007) | (0.007) | (0.009) |
| *Birthmonths* | 0.001+ | 0.001 | 0.0002 | 0.0002 |
| | (0.0004) | (0.001) | (0.001) | (0.001) |
| $D_i \times (Birthmonths)$ | 0.00003 | −0.0003 | −0.0003 | 0.0001 |
| | (0.0002) | (0.0003) | (0.0002) | (0.0003) |
| Pre-treatment mean | 0.231 | 0.212 | 0.192 | 0.173 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no |
| Control for birth season | yes | yes | yes | yes |
| Observations | 3,429,155 | 1,235,612 | 3,429,155 | 1,235,612 |

+p<0.1;*p<0.05; **p<0.01; ***p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) and (3) show the RDD estimates for women aged 60–62 years and include age and birth quarter dummies as control variables. Column (2) and (4) show the RDD estimates for women at age 59 and include birth quarter dummies as control variables. All regressions include linear cohort trends in the running variable on both sides of the policy cut-off.
*Source: KBV, own calculations*

### A.1.3 Control for living in West-Germany

Table A.3: Control for west RDD-results: Mental diagnoses

| | *Stress-related* | | *Mood disorder* | |
|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 |
| $D_i$ | 0.012*** | 0.014*** | 0.016*** | 0.014*** |
| | (0.004) | (0.004) | (0.005) | (0.004) |
| *Birthmonths* | −0.001* | −0.002*** | −0.002*** | −0.004*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| $(Birthmonths)^2$ | −0.0001*** | −0.0002*** | −0.0002*** | −0.0003*** |
| | (0.00004) | (0.00004) | (0.00005) | (0.0001) |
| $D_i \times (Birthmonths)$ | 0.004*** | 0.005*** | 0.005*** | 0.009*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| $D_i \times (Birthmonths)^2$ | −0.00005 | 0.00003 | −0.00002 | −0.00000 |
| | (0.0001) | (0.0001) | (0.0001) | (0.0001) |
| Pre-treatment mean | 0.231 | 0.212 | 0.192 | 0.173 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no |
| Control for birth season | yes | yes | yes | yes |
| Control for west | yes | yes | yes | yes |
| Observations | 3,429,155 | 1,235,612 | 3,429,155 | 1,235,612 |

+p<0.1;*p<0.05; **p<0.01; ***p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) and (3) show the RDD estimates for women aged 60–62 years and include age, birth quarter dummies and a West-Germany dummy as control variables. Column (2) and (4) show the RDD estimates for women at age 59 and include birth quarter dummies and a West-Germany dummy as control variables as control variables. All regressions include linear and quadratic cohort trends in the running variable on both sides of the policy cut-off.
*Source: KBV, own calculations*

Table A.4: Control for west DiD-results: Mental diagnoses

| | *Stress-related* | | *Mood disorder* | |
|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 |
| $Winter5152_i \times JanFebMar_i$ | $0.008^{**}$ | $0.015^{***}$ | $0.009^{***}$ | $0.014^{***}$ |
| | (0.003) | (0.003) | (0.002) | (0.003) |
| $Winter5152_i$ | $0.012^{***}$ | $0.006^{***}$ | $0.008^{***}$ | 0.002 |
| | (0.002) | (0.001) | (0.002) | (0.002) |
| $JanFebMar_i$ | $0.008^{***}$ | $0.004^{+}$ | $0.009^{***}$ | $0.006^{*}$ |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Pre-treatment mean | 0.222 | 0.206 | 0.186 | 0.17 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no |
| Control for west | yes | yes | yes | yes |
| Observations | 1,738,083 | 627,391 | 1,738,083 | 627,391 |

$^{+}$p<0.1;$^{*}$p<0.05; $^{**}$p<0.01; $^{***}$p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) and (3) show the DiD estimates for women aged 60–62 years and include age and and a West-Germany dummy as control variables. Column (2) and (4) show the DiD estimates for women at age 59 and include a West-Germany dummy as control variable. All regressions include the cohort indicator, the reform indicator and their interaction term.
*Source: KBV, own calculations*

### A.1.4 M2Q criterion

Table A.5: M2Q RDD-results: Mental diagnoses

| | *Stress-related disease* | | *Mood disorder* | |
|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 |
| $D_i$ | 0.006 | $0.010^{**}$ | $0.011^{*}$ | $0.018^{***}$ |
| | (0.003) | (0.003) | (0.004) | (0.004) |
| $Birthmonths$ | −0.0004 | $-0.002^{**}$ | $-0.002^{**}$ | $-0.004^{***}$ |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| $(Birthmonths)^2$ | $-0.0001^{**}$ | $-0.0002^{***}$ | $-0.0002^{***}$ | $-0.0002^{***}$ |
| | (0.00003) | (0.00004) | (0.00004) | (0.00005) |
| $D_i \times (Birthmonths)$ | $0.003^{**}$ | $0.004^{***}$ | $0.005^{***}$ | $0.007^{***}$ |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| $D_i \times (Birthmonths)^2$ | −0.00003 | 0.00002 | $-0.0001^{*}$ | −0.00002 |
| | (0.00005) | (0.0001) | (0.00005) | (0.0001) |
| Pre-treatment mean | 0.148 | 0.132 | 0.148 | 0.128 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no |
| Control for birth season | yes | yes | yes | yes |
| Observations | 3,429,155 | 1,235,612 | 3,429,155 | 1,235,612 |

$^{+}$p<0.1;$^{*}$p<0.05; $^{**}$p<0.01; $^{***}$p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) and (3) show the RDD estimates for women aged 60–62 years and include age and birth quarter dummies as control variables. Column (2) and (4) show the RDD estimates for women at age 59 and include birth quarter dummies as control variables. All regressions include linear and quadratic cohort trends in the running variable on both sides of the policy cut-off. The outcome variables are defined according to the M2Q-criterion.
*Source: KBV, own calculations*

Table A.6: M2Q DiD-results: Mental diagnoses

| | *Stress-related disease* | | *Mood disorder* | |
|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 |
| $Winter5152_i \times JanFebMar_i$ | 0.007** | 0.006** | 0.007** | 0.010*** |
| | (0.002) | (0.002) | (0.002) | (0.003) |
| $Winter5152_i$ | 0.010*** | 0.007*** | 0.008*** | 0.003$^+$ |
| | (0.002) | (0.001) | (0.001) | (0.002) |
| $JanFebMar_i$ | 0.007*** | 0.007*** | 0.010*** | 0.007*** |
| | (0.002) | (0.001) | (0.001) | (0.002) |
| Pre-treatment mean | 0.143 | 0.128 | 0.144 | 0.125 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no |
| Observations | 1,738,083 | 627,391 | 1,738,083 | 627,391 |

$^+$p<0.1;*p<0.05; **p<0.01; ***p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) and (3) show the DiD estimates for women aged 60–62 years and include age as control variable. Column (2) and (4) show the DiD estimates for women at age 59. All regressions include the cohort indicator, the reform indicator and their interaction term. The outcome variables are defined according to the M2Q-criterion.
*Source: KBV, own calculations*

## A.2 Physical health

### A.2.1 Placebo-Reform (01/1951)

Table A.7: Placebo-Reform (01/1951): Results metabolic/nutritional diseases

| | *Diabetes* | | *Obesity* | |
|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 |
| $D_i$ | 0.021*** | 0.016*** | 0.009$^+$ | −0.0003 |
| | (0.005) | (0.004) | (0.005) | (0.004) |
| $Birthmonths$ | 0.0004 | −0.0003 | −0.001 | −0.001 |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| $(Birthmonths)^2$ | 0.0002** | 0.0001$^+$ | 0.00001 | −0.00003 |
| | (0.0001) | (0.00005) | (0.0001) | (0.0001) |
| $D_i \times (Birthmonths)$ | −0.003* | −0.002 | 0.001 | 0.002 |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| $D_i \times (Birthmonths)^2$ | −0.0001 | −0.00002 | −0.0001 | −0.0001 |
| | (0.00005) | (0.00004) | (0.0001) | (0.0001) |
| Pre-treatment mean | 0.123 | 0.097 | 0.137 | 0.128 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no |
| Control for birth season | yes | yes | yes | yes |
| Observations | 3,524,843 | 1,275,284 | 3,524,843 | 1,275,284 |

$^+$p<0.1;*p<0.05; **p<0.01; ***p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) and (3) show the RDD estimates for women aged 60–62 years and include age and birth quarter dummies as control variables. Column (2) and (4) show the RDD estimates for women at age 59 and include birth quarter dummies as control variables. All regressions include linear and quadratic cohort trends in the running variable on both sides of the policy cut-off. The cut-off date is 01/1951.
*Source: KBV, own calculations*

Table A.8: Placebo-Reform (01/1951): Results circulatory/heart diagnoses

| | *Hypertension* | | *Heart diagnosis* | | *Stroke* | |
|---|---|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 | Main | Age-59 |
| $D_i$ | 0.007 | 0.0002 | $0.007^{*}$ | $0.005^{+}$ | $0.004^{*}$ | 0.002 |
| | (0.009) | (0.009) | (0.004) | (0.003) | (0.002) | (0.002) |
| *Birthmonths* | $-0.005^{***}$ | $-0.007^{***}$ | −0.001 | $-0.001^{+}$ | $-0.001^{***}$ | $-0.001^{+}$ |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.0003) | (0.0003) |
| $(Birthmonths)^2$ | $-0.0004^{**}$ | $-0.001^{***}$ | −0.00002 | −0.00004 | $-0.0001^{***}$ | $-0.0001^{*}$ |
| | (0.0001) | (0.0001) | (0.00003) | (0.00003) | (0.00002) | (0.00002) |
| $D_i \times (Birthmonths)$ | $0.011^{***}$ | $0.013^{***}$ | 0.001 | 0.001 | $0.002^{***}$ | $0.001^{*}$ |
| | (0.002) | (0.002) | (0.001) | (0.001) | (0.0005) | (0.0005) |
| $D_i \times (Birthmonths)^2$ | 0.00001 | 0.0001 | −0.00000 | −0.00001 | 0.00001 | 0.00002 |
| | (0.0001) | (0.0001) | (0.00003) | (0.00003) | (0.00002) | (0.00002) |
| Pre-treatment mean | 0.401 | 0.348 | 0.051 | 0.041 | 0.037 | 0.026 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no | yes | no |
| Control for birth season | yes | yes | yes | yes | yes | yes |
| Observations | 3,524,843 | 1,275,284 | 3,524,843 | 1,275,284 | 3,524,843 | 1,275,284 |

$^{+}$p<0.1;$^{*}$p<0.05; $^{**}$p<0.01; $^{***}$p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1),(3) and (5) show the RDD estimates for women aged 60–62 years and include age and birth quarter dummies as control variables. Column (2), (4) and (6) show the RDD estimates for women at age 59 and include birth quarter dummies as control variables. All regressions include linear and quadratic cohort trends in the running variable on both sides of the policy cut-off. The cut-off date is 01/1951.
*Source: KBV, own calculations*

Table A.9: Placebo-Reform (01/1951): Results musculoskeletal diagnoses

| | *Arthrosis* | | *Dorsopathies* | |
|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 |
| $D_i$ | $0.012^{+}$ | $0.013^{+}$ | −0.004 | −0.012 |
| | (0.008) | (0.008) | (0.008) | (0.008) |
| *Birthmonths* | $-0.003^{*}$ | $-0.004^{***}$ | −0.002 | $-0.005^{***}$ |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| $(Birthmonths)^2$ | $-0.0002^{*}$ | $-0.0003^{***}$ | $-0.0003^{*}$ | $-0.0005^{***}$ |
| | (0.0001) | (0.0001) | (0.0001) | (0.0001) |
| $D_i \times (Birthmonths)$ | $0.006^{**}$ | $0.008^{***}$ | $0.006^{**}$ | $0.011^{***}$ |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| $D_i \times (Birthmonths)^2$ | −0.00005 | −0.00002 | 0.0001 | 0.0001 |
| | (0.0001) | (0.0001) | (0.0001) | (0.0001) |
| Pre-treatment mean | 0.23 | 0.197 | 0.369 | 0.352 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no |
| Control for birth season | yes | yes | yes | yes |
| Observations | 3,524,843 | 1,275,284 | 3,524,843 | 1,275,284 |

$^{+}$p<0.1;$^{*}$p<0.05; $^{**}$p<0.01; $^{***}$p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) and (3) show the RDD estimates for women aged 60–62 years and include age and birth quarter dummies as control variables. Column (2) and (4) show the RDD estimates for women at age 59 and include birth quarter dummies as control variables. All regressions include linear and quadratic cohort trends in the running variable on both sides of the policy cut-off. The cut-off date is 01/1951.
*Source: KBV, own calculations*

### A.2.2 Linear cohort trend

Table A.10: Linear cohort trend results: Metabolic/nutritional diagnoses

| | *Diabetes* | | *Obesity* | |
|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 |
| $D_i$ | $0.018^{***}$ | $0.017^{***}$ | $0.015^{**}$ | $0.014^{***}$ |
| | (0.006) | (0.004) | (0.005) | (0.004) |
| *Birthmonths* | $-0.001^{*}$ | $-0.001^{*}$ | $-0.0002$ | $-0.001^{+}$ |
| | (0.0004) | (0.0003) | (0.0004) | (0.0003) |
| $D_i \times (Birthmonths)$ | $-0.0004^{*}$ | $-0.0003^{+}$ | $-0.001^{***}$ | $-0.001^{***}$ |
| | (0.0002) | (0.0001) | (0.0002) | (0.0002) |
| Pre-treatment mean | 0.124 | 0.098 | 0.138 | 0.125 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no |
| Control for birth season | yes | yes | yes | yes |
| Observations | 3,429,155 | 1,235,612 | 3,429,155 | 1,235,612 |

$^{+}$p<0.1;$^{*}$p<0.05; $^{**}$p<0.01; $^{***}$p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) and (3) show the RDD estimates for women aged 60–62 years and include age and birth quarter dummies as control variables. Column (2) and (4) show the RDD estimates for women at age 59 and include birth quarter dummies as control variables. All regressions include linear cohort trends in the running variable on both sides of the policy cut-off.
*Source: KBV, own calculations*

Table A.11: Linear cohort trend results: Circulatory/heart diagnoses

| | *Hypertension* | | *Heart diagnosis* | | *Stroke* | |
|---|---|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 | Main | Age-59 |
| $D_i$ | 0.008 | 0.013 | $0.006^{*}$ | $0.006^{*}$ | $0.005^{**}$ | $0.006^{**}$ |
| | (0.011) | (0.011) | (0.003) | (0.003) | (0.002) | (0.002) |
| *Birthmonths* | 0.001 | 0.001 | $-0.0004$ | $-0.0005^{+}$ | $-0.0001$ | $-0.0001$ |
| | (0.001) | (0.001) | (0.0003) | (0.0002) | (0.0002) | (0.0002) |
| $D_i \times (Birthmonths)$ | $-0.001^{+}$ | $-0.001^{+}$ | 0.0001 | 0.0001 | $-0.0001$ | $-0.0001^{+}$ |
| | (0.0004) | (0.0004) | (0.0001) | (0.0001) | (0.0001) | (0.0001) |
| Pre-treatment mean | 0.412 | 0.347 | 0.052 | 0.041 | 0.04 | 0.028 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no | yes | no |
| Control for birth season | yes | yes | yes | yes | yes | yes |
| Observations | 3,429,155 | 1,235,612 | 3,429,155 | 1,235,612 | 3,429,155 | 1,235,612 |

$^{+}$p<0.1;$^{*}$p<0.05; $^{**}$p<0.01; $^{***}$p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1),(3) and (5) show the RDD estimates for women aged 60–62 years and include age and birth quarter dummies as control variables. Column (2), (4) and (6) show the RDD estimates for women at age 59 and include birth quarter dummies as control variables. All regressions include linear cohort trends in the running variable on both sides of the policy cut-off.
*Source: KBV, own calculations*

Table A.12: Linear cohort trend results: Musculoskeletal diagnoses

| | *Arthrosis* | | *Dorsopathies* | |
|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 |
| $D_i$ | 0.017* | 0.019* | 0.010 | 0.009 |
| | (0.007) | (0.008) | (0.007) | (0.009) |
| *Birthmonths* | −0.0001 | −0.001 | 0.001 | 0.001 |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| $D_i \times (Birthmonths)$ | −0.0002 | −0.0001 | −0.001* | −0.0004 |
| | (0.0003) | (0.0004) | (0.0003) | (0.0004) |
| Pre-treatment mean | 0.239 | 0.203 | 0.382 | 0.354 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no |
| Control for birth season | yes | yes | yes | yes |
| Observations | 3,429,155 | 1,235,612 | 3,429,155 | 1,235,612 |

$^{+}$p<0.1;$^{*}$p<0.05; $^{**}$p<0.01; $^{***}$p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) and (3) show the RDD estimates for women aged 60–62 years and include age and birth quarter dummies as control variables. Column (2) and (4) show the RDD estimates for women at age 59 and include birth quarter dummies as control variables. All regressions include linear cohort trends in the running variable on both sides of the policy cut-off.
*Source: KBV, own calculations*

### A.2.3 Control for living in West-Germany

Table A.13: Control for west RDD-results: Metabolic/nutritional diagnoses

| | *Diabetes* | | *Obesity* | |
|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 |
| $D_i$ | 0.021*** | 0.019*** | 0.020*** | 0.018*** |
| | (0.005) | (0.005) | (0.005) | (0.003) |
| *Birthmonths* | −0.0004 | −0.0004 | −0.001 | −0.001$^{+}$ |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| $(Birthmonths)^2$ | 0.0001 | 0.00004 | −0.00003 | −0.00003 |
| | (0.00005) | (0.00005) | (0.00005) | (0.00004) |
| $D_i \times (Birthmonths)$ | −0.002* | −0.002 | −0.001 | −0.001 |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| $D_i \times (Birthmonths)^2$ | 0.0001 | 0.00002 | 0.0001* | 0.0001* |
| | (0.00005) | (0.00005) | (0.00004) | (0.00004) |
| Pre-treatment mean | 0.124 | 0.098 | 0.138 | 0.125 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no |
| Control for birth season | yes | yes | yes | yes |
| Control for west | yes | yes | yes | yes |
| Observations | 3,429,155 | 1,235,612 | 3,429,155 | 1,235,612 |

$^{+}$p<0.1;$^{*}$p<0.05; $^{**}$p<0.01; $^{***}$p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) and (3) show the RDD estimates for women aged 60–62 years and include age, birth quarter dummies and a West-Germany dummy as control variables. Column (2) and (4) show the RDD estimates for women at age 59 and include birth quarter dummies and a West-Germany dummy as control variables as control variables. All regressions include linear and quadratic cohort trends in the running variable on both sides of the policy cut-off.
*Source: KBV, own calculations*

Table A.14: Control for west DiD-results: Metabolic/nutritional diagnoses

| | *Diabetes* | | *Obesity* | |
|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 |
| $Winter5152_i \times JanFebMar_i$ | $0.003^{*}$ | $0.005^{**}$ | $0.010^{***}$ | $0.009^{***}$ |
| | (0.002) | (0.002) | (0.001) | (0.001) |
| $Winter5152_i$ | $0.002^{+}$ | 0.001 | $0.003^{**}$ | $-0.002^{*}$ |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| $JanFebMar_i$ | $0.013^{***}$ | $0.011^{***}$ | $0.010^{***}$ | $0.005^{***}$ |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Pre-treatment mean | 0.123 | 0.097 | 0.135 | 0.123 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no |
| Control for west | yes | yes | yes | yes |
| Observations | 1,738,083 | 627,391 | 1,738,083 | 627,391 |

$^{+}$p<0.1;$^{*}$p<0.05; $^{**}$p<0.01; $^{***}$p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) and (3) show the DiD estimates for women aged 60–62 years and include age and and a West-Germany dummy as control variables. Column (2) and (4) show the DiD estimates for women at age 59 and include a West-Germany dummy as control variable. All regressions include the cohort indicator, the reform indicator and their interaction term.
*Source: KBV, own calculations*

Table A.15: Control for west RDD-results: Circulatory/heart diagnoses

| | *Hypertension* | | *Heart diagnosis* | | *Stroke* | |
|---|---|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 | Main | Age-59 |
| $D_i$ | $0.016^{*}$ | $0.020^{**}$ | $0.007^{*}$ | $0.008^{**}$ | $0.006^{***}$ | $0.006^{***}$ |
| | (0.008) | (0.008) | (0.003) | (0.003) | (0.002) | (0.002) |
| $Birthmonths$ | $-0.004^{**}$ | $-0.004^{**}$ | $-0.001$ | $-0.001^{+}$ | $-0.001^{*}$ | $-0.001^{*}$ |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.0003) | (0.0003) |
| $(Birthmonths)^2$ | $-0.0004^{***}$ | $-0.0003^{***}$ | $-0.00000$ | $-0.00004$ | $-0.0001^{*}$ | $-0.0001^{*}$ |
| | (0.0001) | (0.0001) | (0.00005) | (0.00004) | (0.00002) | (0.00002) |
| $D_i \times (Birthmonths)$ | $0.008^{***}$ | $0.008^{***}$ | 0.0003 | 0.001 | $0.001^{*}$ | $0.001^{*}$ |
| | (0.002) | (0.002) | (0.001) | (0.001) | (0.001) | (0.001) |
| $D_i \times (Birthmonths)^2$ | 0.0001 | 0.00003 | $-0.00001$ | 0.00001 | 0.00001 | 0.00001 |
| | (0.0001) | (0.0001) | (0.00003) | (0.00003) | (0.00002) | (0.00002) |
| Pre-treatment mean | 0.412 | 0.347 | 0.052 | 0.041 | 0.04 | 0.028 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no | yes | no |
| Control for birth season | yes | yes | yes | yes | yes | yes |
| Control for west | yes | yes | yes | yes | yes | yes |
| Observations | 3,429,155 | 1,235,612 | 3,429,155 | 1,235,612 | 3,429,155 | 3,429,155 |

$^{+}$p<0.1;$^{*}$p<0.05; $^{**}$p<0.01; $^{***}$p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1),(3) and (5) show the RDD estimates for women aged 60–62 years and include age, birth quarter dummies and a West-Germany dummy as control variables. Column (2), (4) and (6) show the RDD estimates for women at age 59 and include birth quarter dummies and a West-Germany dummy as control variables. All regressions include linear and quadratic cohort trends in the running variable on both sides of the policy cut-off.
*Source: KBV, own calculations*

Table A.16: Control for west DiD-results: Circulatory/heart diagnoses

| | *Hypertension* | | *Heart diagnosis* | | *Stroke* | |
|---|---|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 | Main | Age-59 |
| $Winter5152_i \times JanFebMar_i$ | $0.007^{+}$ | $0.024^{***}$ | 0.0004 | −0.0002 | 0.001 | $0.002^{***}$ |
| | (0.004) | (0.004) | (0.001) | (0.001) | (0.001) | (0.001) |
| $Winter5152_i$ | $0.011^{***}$ | −0.0005 | 0.0004 | 0.0005 | $0.003^{***}$ | $0.001^{***}$ |
| | (0.002) | (0.002) | (0.001) | (0.0004) | (0.001) | (0.0004) |
| $JanFebMar_i$ | $0.023^{***}$ | $0.010^{***}$ | $0.006^{***}$ | $0.005^{***}$ | $0.006^{***}$ | $0.004^{***}$ |
| | (0.003) | (0.003) | (0.001) | (0.001) | (0.0005) | (0.0004) |
| Pre-treatment mean | 0.403 | 0.342 | 0.05 | 0.04 | 0.038 | 0.027 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no | yes | no |
| Control for west | yes | yes | yes | yes | yes | yes |
| Observations | 1,738,083 | 627,391 | 1,738,083 | 627,391 | 1,738,083 | 627,391 |

$^{+}$p<0.1;$^{*}$p<0.05; $^{**}$p<0.01; $^{***}$p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1), (3) and (5) show the DiD estimates for women aged 60–62 years and include age and a West-Germany dummy as control variables. Column (2), (4) and (6) show the DiD estimates for women at age 59 and include a West-Germany dummy as control variable. All regressions include the cohort indicator, the reform indicator and their interaction term.
*Source: KBV, own calculations*

Table A.17: Control for west RDD-results: Musculoskeletal diagnoses

| | *Arthrosis* | | *Dorsopathies* | |
|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 |
| $D_i$ | $0.022^{***}$ | $0.022^{***}$ | $0.012^{*}$ | $0.012^{*}$ |
| | (0.005) | (0.005) | (0.006) | (0.006) |
| $Birthmonths$ | $-0.003^{***}$ | $-0.003^{***}$ | $-0.001^{+}$ | $-0.001^{+}$ |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| $(Birthmonths)^2$ | $-0.0002^{**}$ | $-0.0002^{**}$ | $-0.0002^{**}$ | $-0.0002^{**}$ |
| | (0.0001) | (0.0001) | (0.0001) | (0.0001) |
| $D_i \times (Birthmonths)$ | $0.004^{**}$ | $0.004^{**}$ | $0.004^{***}$ | $0.004^{***}$ |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| $D_i \times (Birthmonths)^2$ | 0.0001 | 0.0001 | −0.0001 | −0.0001 |
| | (0.0001) | (0.0001) | (0.0001) | (0.0001) |
| Pre-treatment mean | | 0.203 | | 0.354 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no |
| Control for birth season | yes | yes | yes | yes |
| Control for west | yes | yes | yes | yes |
| Observations | 3,429,155 | 3,429,155 | 3,429,155 | 3,429,155 |

$^{+}$p<0.1;$^{*}$p<0.05; $^{**}$p<0.01; $^{***}$p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) and (3) show the RDD estimates for women aged 60–62 years and include age, birth quarter dummies and a West-Germany dummy as control variables. Column (2) and (4) show the RDD estimates for women at age 59 and include birth quarter dummies and a West-Germany dummy as control variables as control variables. All regressions include linear and quadratic cohort trends in the running variable on both sides of the policy cut-off.
*Source: KBV, own calculations*

Table A.18: Control for west DiD-results: Musculoskeletal diagnoses

| | *Arthrosis* | | *Dorsopathies* | |
|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 |
| $Winter5152_i \times JanFebMar_i$ | 0.008** | 0.007* | 0.008** | 0.021*** |
| | (0.003) | (0.003) | (0.003) | (0.004) |
| $Winter5152_i$ | 0.008*** | 0.004+ | 0.013*** | 0.001 |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| $JanFebMar_i$ | 0.017*** | 0.014*** | 0.012*** | 0.001 |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| Pre-treatment mean | 0.235 | 0.201 | 0.374 | 0.352 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no |
| Control for west | yes | yes | yes | yes |
| Observations | 1,738,083 | 627,391 | 1,738,083 | 627,391 |

+p<0.1;*p<0.05; **p<0.01; ***p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) and (3) show the DiD estimates for women aged 60–62 years and include age and and a West-Germany dummy as control variables. Column (2) and (4) show the DiD estimates for women at age 59 and include a West-Germany dummy as control variable. All regressions include the cohort indicator, the reform indicator and their interaction term.
*Source: KBV, own calculations*

### A.2.4 M2Q criterion

Table A.19: M2Q RDD-results: Metabolic/nutritional diagnoses

| | *Diabetes* | | *Obesity* | |
|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 |
| $D_i$ | 0.017*** | 0.016*** | 0.016*** | 0.008* |
| | (0.005) | (0.004) | (0.005) | (0.003) |
| $Birthmonths$ | −0.0001 | −0.0004 | −0.0005 | −0.0004 |
| | (0.001) | (0.001) | (0.001) | (0.0004) |
| $(Birthmonths)^2$ | 0.00004 | 0.00002 | −0.00002 | −0.00003 |
| | (0.00005) | (0.00004) | (0.00004) | (0.00003) |
| $D_i \times (Birthmonths)$ | −0.002+ | −0.001 | −0.001 | −0.001 |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| $D_i \times (Birthmonths)^2$ | 0.0001 | 0.00003 | 0.0001* | 0.0001* |
| | (0.00005) | (0.00004) | (0.00004) | (0.00003) |
| Pre-treatment mean | 0.113 | 0.087 | 0.1 | 0.089 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no |
| Control for birth season | yes | yes | yes | yes |
| Observations | 3,429,155 | 1,235,612 | 3,429,155 | 1,235,612 |

+p<0.1;*p<0.05; **p<0.01; ***p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) and (3) show the RDD estimates for women aged 60–62 years and include age and birth quarter dummies as control variables. Column (2) and (4) show the RDD estimates for women at age 59 and include birth quarter dummies as control variables. All regressions include linear and quadratic cohort trends in the running variable on both sides of the policy cut-off. The outcome variables are defined according to the M2Q-criterion.
*Source: KBV, own calculations*

Table A.20: M2Q DiD-results: Metabolic/nutritional diagnoses

| | *Diabetes* | | *Obesity* | |
|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 |
| $Winter5152_i \times JanFebMar_i$ | 0.003* | 0.005** | 0.008*** | 0.006*** |
| | (0.002) | (0.002) | (0.001) | (0.001) |
| $Winter5152_i$ | 0.003* | 0.002 | 0.004*** | −0.0001 |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| $JanFebMar_i$ | 0.012*** | 0.009*** | 0.008*** | 0.004*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Pre-treatment mean | 0.111 | 0.111 | 0.097 | 0.097 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no |
| Observations | 1,738,083 | 627,391 | 1,738,083 | 627,391 |

$^{+}$p<0.1;*p<0.05; **p<0.01; ***p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) and (3) show the DiD estimates for women aged 60–62 years and include age as control variable. Column (2) and (4) show the DiD estimates for women at age 59. All regressions include the cohort indicator, the reform indicator and their interaction term. The outcome variables are defined according to the M2Q-criterion.
*Source: KBV, own calculations*

Table A.21: M2Q RDD-results: Circulatory/heart diagnoses

| | *Hypertension* | | *Heart diagnosis* | | *Stroke* | |
|---|---|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 | Main | Age-59 |
| $D_i$ | 0.009 | 0.021** | 0.006* | 0.007** | 0.001 | 0.001 |
| | (0.007) | (0.008) | (0.003) | (0.002) | (0.002) | (0.002) |
| $Birthmonths$ | −0.004*** | −0.005*** | −0.001 | −0.001 | −0.0001 | 0.00001 |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.0003) | (0.0003) |
| $(Birthmonths)^2$ | −0.0004*** | −0.0004*** | −0.00001 | −0.00002 | −0.00001 | 0.00000 |
| | (0.0001) | (0.0001) | (0.00005) | (0.00004) | (0.00002) | (0.00002) |
| $D_i \times (Birthmonths)$ | 0.008*** | 0.008*** | 0.0003 | 0.001 | 0.001 | 0.0003 |
| | (0.002) | (0.002) | (0.001) | (0.001) | (0.0005) | (0.0004) |
| $D_i \times (Birthmonths)^2$ | 0.0001 | 0.0001 | 0.00001 | 0.00000 | 0.00000 | −0.00002 |
| | (0.0001) | (0.0001) | (0.00003) | (0.00003) | (0.00003) | (0.00002) |
| Pre-treatment mean | 0.366 | 0.302 | 0.038 | 0.029 | 0.031 | 0.022 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no | yes | no |
| Control for birth season | yes | yes | yes | yes | yes | yes |
| Observations | 3,429,155 | 1,235,612 | 3,429,155 | 1,235,612 | 3,429,155 | 1,235,612 |

$^{+}$p<0.1;*p<0.05; **p<0.01; ***p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1),(3) and (5) show the RDD estimates for women aged 60–62 years and include age and birth quarter dummies as control variables. Column (2), (4) and (6) show the RDD estimates for women at age 59 and include birth quarter dummies as control variables. All regressions include linear and quadratic cohort trends in the running variable on both sides of the policy cut-off. The outcome variables are defined according to the M2Q-criterion.
*Source: KBV, own calculations*

Table A.22: M2Q DiD-results: Circulatory/heart diagnoses

| | *Hypertension* | | *Heart diagnosis* | | *Stroke* | |
|---|---|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 | Main | Age-59 |
| $Winter5152_i \times JanFebMar_i$ | 0.006 | 0.024*** | 0.0003 | 0.0005 | 0.001* | 0.002** |
| | (0.004) | (0.004) | (0.001) | (0.001) | (0.0004) | (0.001) |
| $Winter5152_i$ | 0.013*** | 0.002 | 0.001 | 0.001 | 0.002*** | 0.001* |
| | (0.002) | (0.002) | (0.001) | (0.0004) | (0.0003) | (0.0004) |
| $JanFebMar_i$ | 0.022*** | 0.010** | 0.005*** | 0.003*** | 0.004*** | 0.002*** |
| | (0.003) | (0.003) | (0.001) | (0.001) | (0.0003) | (0.0004) |
| Pre-treatment mean | 0.357 | 0.296 | 0.037 | 0.028 | 0.025 | 0.017 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no | yes | no |
| Observations | 1,738,083 | 627,391 | 1,738,083 | 627,391 | 1,738,083 | 627,391 |

$^{+}$p<0.1;$^{*}$p<0.05; $^{**}$p<0.01; $^{***}$p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1), (3) and (5) show the DiD estimates for women aged 60–62 years and include age as control variable. Column (2), (4) and (6) show the DiD estimates for women at age 59. All regressions include the cohort indicator, the reform indicator and their interaction term. The outcome variables are defined according to the M2Q-criterion.
*Source: KBV, own calculations*

Table A.23: M2Q RDD-results: Musculoskeletal diagnoses

| | *Arthrosis* | | *Dorsopathies* | |
|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 |
| $D_i$ | 0.017** | 0.015*** | 0.013* | 0.020*** |
| | (0.005) | (0.003) | (0.005) | (0.005) |
| $Birthmonths$ | −0.001$^{+}$ | −0.002*** | −0.001 | −0.003*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| $(Birthmonths)^2$ | −0.0001 | −0.0002*** | −0.0001* | −0.0002*** |
| | (0.0001) | (0.00005) | (0.00005) | (0.00005) |
| $D_i \times (Birthmonths)$ | 0.002 | 0.004*** | 0.002$^{+}$ | 0.004*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| $D_i \times (Birthmonths)^2$ | 0.00003 | 0.00002 | −0.00001 | 0.0001$^{+}$ |
| | (0.0001) | (0.0001) | (0.00005) | (0.0001) |
| Pre-treatment mean | 0.17 | 0.14 | 0.266 | 0.242 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no |
| Control for birth season | yes | yes | yes | yes |
| Observations | 3,429,155 | 1,235,612 | 3,429,155 | 1,235,612 |

$^{+}$p<0.1;$^{*}$p<0.05; $^{**}$p<0.01; $^{***}$p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) and (3) show the RDD estimates for women aged 60–62 years and include age and birth quarter dummies as control variables. Column (2) and (4) show the RDD estimates for women at age 59 and include birth quarter dummies as control variables. All regressions include linear and quadratic cohort trends in the running variable on both sides of the policy cut-off. The outcome variables are defined according to the M2Q-criterion.
*Source: KBV, own calculations*

Table A.24: M2Q DiD-results: Musculoskeletal diagnoses

| | *Arthrosis* | | *Dorsopathies* | |
|---|---|---|---|---|
| | Main | Age-59 | Main | Age-59 |
| $Winter5152_i \times JanFebMar_i$ | 0.007*** | 0.005* | 0.009*** | 0.013*** |
| | (0.002) | (0.002) | (0.002) | (0.002) |
| $Winter5152_i$ | 0.007*** | 0.003+ | 0.010*** | 0.002 |
| | (0.001) | (0.002) | (0.001) | (0.002) |
| $JanFebMar_i$ | 0.015*** | 0.011*** | 0.011*** | 0.006*** |
| | (0.002) | (0.002) | (0.001) | (0.001) |
| Pre-treatment mean | 0.166 | 0.137 | 0.261 | 0.24 |
| Age group included | 60-62 years | 59 years | 60-62 years | 59 years |
| Control for age | yes | no | yes | no |
| Control for west | yes | yes | yes | yes |
| Observations | 1,738,083 | 627,391 | 1,738,083 | 627,391 |

+p<0.1;*p<0.05; **p<0.01; ***p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) and (3) show the DiD estimates for women aged 60–62 years and include age as control variable. Column (2) and (4) show the DiD estimates for women at age 59. All regressions include the cohort indicator, the reform indicator and their interaction term. The outcome variables are defined according to the M2Q-criterion.
*Source: KBV, own calculations*

## A.3 Multiple hypothesis testing

Table A.25: Bonferroni correction for multiple hypothesis testing in RDD - P-values

| | *60-62 years* | | *59 years* | |
|---|---|---|---|---|
| | without correction | Bonferroni | without correction | Bonferroni |
| Stress-related diseases | 0.002** | 0.0179* | 0.0022** | 0.0199* |
| Mood disorder | 0.0005*** | 0.0045** | 0.0004*** | 0.0034** |
| Diabetes | 0.0001*** | 0.0012** | 0.0000*** | 0.0001*** |
| Obesity | 0.0001*** | 0.0008*** | 0.0000*** | 0.0000*** |
| Hypertension | 0.1091 | 0.9815 | 0.0564+ | 0.5077 |
| Ischaemic heart diseases | 0.0503+ | 0.4531 | 0.0167* | 0.1506 |
| Stroke | 0.0003*** | 0.0023** | 0.0364* | 0.3280 |
| Arthrosis | 0.0000*** | 0.0002*** | 0.0000*** | 0.0004*** |
| Other dorsopathies | 0.0908+ | 0.8172 | 0.0289* | 0.2605 |

+p<0.1;*p<0.05; **p<0.01; ***p<0.001

*Note:* Column (1) and (3) show the p-values retrieved from the baseline RDD estimation. The underlying standard errors are clustered on month of birth (running variable) and robust. Column (2) and (4) show the Bonferroni-corrected p-values.
*Source: KBV, own calculations*

Table A.26: Bonferroni correction for multiple hypothesis testing in DiD - P-values

| | *60-62 years* | | *59 years* | |
|---|---|---|---|---|
| | without correction | Bonferroni | without correction | Bonferroni |
| Stress-related diseases | 0.0071** | 0.0637+ | 0.0000*** | 0.0000*** |
| Mood disorder | 0.0003*** | 0.0026** | 0.0000*** | 0.0003*** |
| Diabetes | 0.0491* | 0.4421 | 0.0084** | 0.0759+ |
| Obesity | 0.0000*** | 0.0000*** | 0.0000*** | 0.0000*** |
| Hypertension | 0.0994+ | 0.8945 | 0.0000*** | 0.0000*** |
| Ischaemic heart diseases | 0.7339 | 1.0000 | 0.6758 | 1.0000 |
| Stroke | 0.1589 | 1.0000 | 0.0008*** | 0.0074** |
| Arthrosis | 0.0023** | 0.0211* | 0.0238* | 0.2142 |
| Other dorsopathies | 0.0072** | 0.0652+ | 0.0000*** | 0.0000*** |

+p<0.1;*p<0.05; **p<0.01; ***p<0.001

*Note:* Column (1) and (3) show the p-values retrieved from the baseline DiD estimation. The underlying standard errors are clustered on month of birth (running variable) and robust. Column (2) and (4) show the Bonferroni-corrected p-values.
*Source: KBV, own calculations*

## A.4 Healthcare consumption

### A.4.1 Placebo-Reform (01/1951)

Table A.27: Placebo-Reform (01/1951): Number of doctor visits

| | *Dependent variable: Doctor visits* | |
|---|---|---|
| | Main | Age-59 |
| $D_i$ | 0.088 | −0.107 |
| | (0.239) | (0.251) |
| $Birthmonths$ | −0.092* | −0.143** |
| | (0.041) | (0.049) |
| $(Birthmonths)^2$ | −0.009** | −0.013** |
| | (0.003) | (0.004) |
| $D_i \times (Birthmonths)$ | 0.209** | 0.290*** |
| | (0.067) | (0.078) |
| $D_i \times (Birthmonths)^2$ | 0.002 | 0.004 |
| | (0.003) | (0.004) |
| Pre-treatment mean | 9.631 | 8.606 |
| Age group included | 60-62 years | 59 years |
| Control for age | yes | no |
| Control for birth season | yes | yes |
| Observations | 3,429,155 | 1,235,612 |

+p<0.1;*p<0.05; **p<0.01; ***p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) shows the RDD estimates for women aged 60–62 years and include age and birth quarter dummies as control variables. Column (2) shows the RDD estimates for women at age 59 and include birth quarter dummies as control variables. Both regressions include linear and quadratic cohort trends in the running variable on both sides of the policy cut-off. The cut-off date is 01/1951.
*Source: KBV, own calculations*

### A.4.2 Linear cohort trend

Table A.28: Linear cohort trend results: Number of doctor visits

| | *Dependent variable: Doctor visits* | |
|---|---|---|
| | Main | Age-59 |
| $D_i$ | 0.215 | 0.494* |
| | (0.207) | (0.213) |
| $Birthmonths$ | 0.023 | 0.009 |
| | (0.018) | (0.019) |
| $D_i \times (Birthmonths)$ | −0.007 | −0.007 |
| | (0.009) | (0.009) |
| Pre-treatment mean | 9.631 | 8.606 |
| Age group included | 60-62 years | 59 years |
| Control for age | yes | no |
| Control for birth season | yes | yes |
| Observations | 3,429,155 | 1,235,612 |

$^{+}$p<0.1;$^{*}$p<0.05; $^{**}$p<0.01; $^{***}$p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) shows the RDD estimates for women aged 60–62 years and include age and birth quarter dummies as control variables. Column (2) shows the RDD estimates for women at age 59 and include birth quarter dummies as control variables. Both regressions include linear cohort trends in the running variable on both sides of the policy cut-off.

*Source: KBV, own calculations*

### A.4.3 Control for living in West-Germany

Table A.29: Control for west RDD-results: Number of doctor visits

| | *Dependent variable: Doctor visits* | |
|---|---|---|
| | Main | Age-59 |
| $D_i$ | $0.297^{*}$ | $0.610^{***}$ |
| | (0.151) | (0.151) |
| $Birthmonths$ | $-0.054^{*}$ | $-0.085^{***}$ |
| | (0.024) | (0.024) |
| $(Birthmonths)^2$ | $-0.005^{*}$ | $-0.007^{**}$ |
| | (0.002) | (0.002) |
| $D_i \times (Birthmonths)$ | $0.135^{**}$ | $0.155^{***}$ |
| | (0.043) | (0.043) |
| $D_i \times (Birthmonths)^2$ | 0.00005 | 0.001 |
| | (0.002) | (0.002) |
| Pre-treatment mean | 9.631 | 8.606 |
| Age group included | 60-62 years | 59 years |
| Control for age | yes | no |
| Control for birth season | yes | yes |
| Control for west | yes | yes |
| Observations | 3,429,155 | 1,235,612 |

$^{+}$p<0.1;$^{*}$p<0.05; $^{**}$p<0.01; $^{***}$p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) shows the RDD estimates for women aged 60–62 years and include age, birth quarter dummies and a West-Germany dummy as control variables. Column (2) shows the RDD estimates for women at age 59 and include birth quarter dummies and a West-Germany dummy as control variables. Both regressions include linear and quadratic cohort trends in the running variable on both sides of the policy cut-off.
*Source: KBV, own calculations*

Table A.30: Control for west DiD-results: Number of doctor visits

| | *Dependent variable: Doctor visits* | |
|---|---|---|
| | Main | Age-59 |
| $Winter5152_i \times JanFebMar_i$ | 0.181* | 0.516*** |
| | (0.087) | (0.087) |
| $Winter5152_i$ | 0.328*** | 0.092* |
| | (0.026) | (0.039) |
| $JanFebMar_i$ | 0.377*** | 0.161* |
| | (0.071) | (0.066) |
| Pre-treatment mean | 9.43 | 8.52 |
| Age group included | 60-62 years | 59 years |
| Control for age | yes | no |
| Control for west | yes | yes |
| Observations | 1,738,083 | 627,391 |

$^{+}$p<0.1;*p<0.05; **p<0.01; ***p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. Column (1) shows the DiD estimates for women aged 60–62 years and include age and a West-Germany dummy as control variable. Column (2) shows the DiD estimates for women at age 59 and includes a West-Germany dummy as control variable. All regressions include the cohort indicator, the reform indicator and their interaction term.
*Source: KBV, own calculations*

## A.5 Post-employment effects

Table A.31: Main results: 63-65 year olds

| | *Dependent variable:* | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Stress-related | Mood disorder | Arthrosis | Dorsopathies | Diabetes | Obesity | Hypertension | Heart | Strokes | Doc. visits |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| $D_i$ | 0.008* | 0.013* | 0.028*** | 0.016** | 0.025*** | 0.026*** | 0.018** | 0.009* | 0.005** | 0.342* |
| | (0.004) | (0.005) | (0.007) | (0.006) | (0.007) | (0.005) | (0.007) | (0.004) | (0.002) | (0.134) |
| $Birthmonths$ | −0.003*** | −0.002+ | −0.003* | −0.003*** | −0.0000 | −0.001+ | −0.004** | −0.001 | −0.001* | −0.058* |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.0003) | (0.024) |
| $(Birthmonths)^2$ | −0.0002*** | −0.0001* | −0.0002+ | −0.0002*** | 0.0001 | −0.0001 | −0.0003*** | −0.00002 | −0.00003+ | −0.005* |
| | (0.00005) | (0.0001) | (0.0001) | (0.00005) | (0.0001) | (0.00004) | (0.0001) | (0.0001) | (0.00002) | (0.002) |
| $D_i \times (Birthmonths)$ | 0.007*** | 0.005*** | 0.004* | 0.006*** | −0.004** | −0.001 | 0.008*** | 0.001 | 0.001*** | 0.128*** |
| | (0.001) | (0.001) | (0.002) | (0.001) | (0.001) | (0.001) | (0.002) | (0.001) | (0.004) | (0.039) |
| $D_i \times (Birthmonths)^2$ | −0.0001 | −0.0001* | 0.00005 | −0.00003 | 0.0001 | 0.0002** | 0.00002 | 0.00000 | −0.00004 | 0.0001 |
| | (0.00005) | (0.0001) | (0.0001) | (0.0001) | (0.0001) | (0.0001) | (0.0001) | (0.00004) | (0.00003) | (0.002) |
| Pre-treatment mean | 0.255 | 0.214 | 0.298 | 0.421 | 0.164 | 0.172 | 0.507 | 0.071 | 0.063 | 11.145 |
| Age group included | 63-65 years | 63-65 years | 63-65 years | 63-65 years | 63-65 years | 63-65 years | 63-65 years | 63-65 years | 63-65 years | 63-65 years |
| Control for age | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| Control for birth season | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| Observations | 3,047,412 | 3,047,412 | 3,047,412 | 3,047,412 | 3,047,412 | 3,047,412 | 3,047,412 | 3,047,412 | 3,047,412 | 3,047,412 |

+p<0.1;*p<0.05; **p<0.01; ***p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. All regressions include linear and quadratic cohort trends in the running variable on both sides of the policy cut-off.
*Source: KBV, own calculations*

Table A.32: Placebo results: 63-65 year olds

| | *Dependent variable:* | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Stress-related | Mood disorder | Arthrosis | Dorsopathies | Diabetes | Obesity | Hypertension | Heart | Strokes | Doc visits |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| $D_i$ | 0.011* | 0.012* | 0.028** | 0.013* | 0.030*** | 0.016** | 0.016+ | 0.008+ | 0.005** | 0.304 |
| | (0.005) | (0.005) | (0.008) | (0.006) | (0.006) | (0.006) | (0.009) | (0.004) | (0.002) | (0.228) |
| $Birthmonths$ | −0.005*** | −0.002* | −0.004*** | −0.003*** | 0.0005 | −0.0003 | −0.005*** | −0.001 | −0.001* | −0.097* |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.0003) | (0.041) |
| $(Birthmonths)^2$ | −0.0004*** | −0.0002** | −0.0002* | −0.0003*** | 0.0003*** | 0.0001 | −0.0003** | −0.00002 | −0.00003+ | −0.008** |
| | (0.0001) | (0.0001) | (0.0001) | (0.0001) | (0.0001) | (0.0001) | (0.0001) | (0.00004) | (0.00002) | (0.003) |
| $D_i \times (Birthmonths)$ | 0.009*** | 0.005*** | 0.005** | 0.006*** | −0.005** | 0.0001 | 0.009*** | 0.001 | 0.001*** | 0.189** |
| | (0.001) | (0.001) | (0.002) | (0.001) | (0.002) | (0.001) | (0.002) | (0.001) | (0.0004) | (0.063) |
| $D_i \times (Birthmonths)^2$ | 0.0001 | −0.00005 | −0.00003 | 0.0001 | −0.0001+ | −0.0001 | 0.00002 | −0.00003 | −0.00004 | 0.002 |
| | (0.0001) | (0.0001) | (0.0001) | (0.0001) | (0.0001) | (0.0001) | (0.0001) | (0.00004) | (0.00003) | (0.003) |
| Pre-treatment mean | 0.241 | 0.204 | 0.287 | 0.407 | 0.165 | 0.165 | 0.498 | 0.07 | 0.059 | 10.772 |
| Age group included | 63-65 years | 63-65 years | 63-65 years | 63-65 years | 63-65 years | 63-65 years | 63-65 years | 63-65 years | 63-65 years | 63-65 years |
| Control for age | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| Control for birth season | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| Observations | 3,126,981 | 3,126,981 | 3,126,981 | 3,126,981 | 3,126,981 | 3,126,981 | 3,126,981 | 3,126,981 | 3,126,981 | 3,126,981 |

+p<0.1;*p<0.05; **p<0.01; ***p<0.001

*Note:* Standard errors are clustered on month of birth (running variable) and robust. All regressions include linear and quadratic cohort trends in the running variable on both sides of the policy cut-off. The cut-off date is 01/1951.
*Source: KBV, own calculations*