I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17999

The Labour Market and Health Effects of a Diabetes Warning: Evidence of Gender and Age Differences from the Lifelines Cohort Study

Claudio Annibali
Annette Bergemann
Rob Alessie

JULY 2025

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17999

The Labour Market and Health Effects of a Diabetes Warning: Evidence of Gender and Age Differences from the Lifelines Cohort Study

Claudio Annibali
University of Groningen and Centre for Public Health in Economics and Business

Annette Bergemann
University of Groningen, IZA, Aletta Jacobs School of Public Health and IFAU Uppsala

Rob Alessie
University of Groningen

JULY 2025

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Labour Market and Health Effects of a Diabetes Warning: Evidence of Gender and Age Differences from the Lifelines Cohort Study

To promote early detection of diabetes and ameliorate the negative consequences of diabetes, some governments provide diabetes screenings. This paper contributes to the literature by being the first to investigate whether an issued warning affects the individual's employment status. Additionally, our analysis also explores health effects, stratified by gender, age, and education , in order to receive indications for potential pathways of the employment effects. By doing so, we present the first results in the literature for individuals under 40. Using a multidimensional regression discontinuity design, we investigate the short- and long-run effects of a diabetes risk warning issued by Lifelines, a Dutch cohort study. In particular, low-educated individuals below 40 increase their labour market activities after a warning, which is generally more pronounced and also persistent for women. Surprisingly, this is not matched by similar strong effects on health outcomes by either gender. Health effects are very heterogeneous by gender, age and educational group. Older, highly educated women seem to benefit particularly strongly from a warning, as a significant reduction in the 4-year mortality rate indicates.

JEL Classification: I12, J16, I10

Keywords: diabetes, diabetes screening, employment, undiagnosed diabetes, multidimensional regression discontinuity design

Corresponding author:
Annette Bergemann
University of Groningen
Grote Kruisstraat 2/1
9712 TS Groningen
The Netherlands

E-mail: a.h.bergemann@rug.nl

1 Introduction

Diabetes mellitus (DM) type 2 is an expanding chronic disease that in 2024 affected 1 in 10 adults worldwide, as estimated by the International Diabetes Federation [1].[1] However, about 4 in 10 diabetic adults are unaware of their disease [1]. This is an alarming number, particularly in light of descriptive finding suggesting that undiagnosed individuals compared to diagnosed are more likely to develop health complications, such as heart disease [4], and experience larger mortality rates [5]. To promote early DM detection and prevent such negative consequences, some governments (i.e. Japan and Korea) provide diabetes screenings, which issue warnings if individuals are at risk of DM, targeted at the whole population aged 40 and over. In a comprehensive evaluation, we investigate whether receiving a DM warning initializes labour market and health effects in the Netherlands, a country where no nationwide DM screening exists.

To date, the effects of diabetes mellitus warnings on labour market behaviour have not been examined. This is an important gap, particularly in light of a growing body of evidence documenting the adverse labour market consequences of a formal DM diagnosis (as surveyed by [6]). A range of studies has shown a negative relationship between diagnosed DM and labour market outcomes. Moreover, there is some evidence that following a diabetes diagnosis, additional shocks might lead to a worsening of the disease [7]. These findings suggest that interventions at an earlier stage - such as DM warnings issued when individuals are identified as prediabetic - could, in principle, contribute to preventing these negative outcomes, for example by altering health trajectories or behavioural responses. Whether such warnings are effective in doing so, however, remains an open empirical question.

While a warning can positively affect health behaviours and outcomes, its effect on labour market outcomes is ex-ante unclear. Regarding health behaviours and outcomes, there is some evidence that issued warnings are effective in the short run [8, 9].[2] We add additional insights with regard to this pathway, particularly by taking the heterogeneity of different population groups into account. Following [12], these potential health changes may influence labour market outcomes by way of three principal channels: productivity, preferences and expectations, and financial incentives. If a warning, for example, prevents the onset of diabetes or mitigates the severity of its

[1] We consider only type 2 diabetes as it is the most widespread type and is usually developed in later stages of life [2]. In the Netherlands, it affects 90% of the diabetic population [3]. Type 2 DM causes higher insulin resistance and relative insulin deficiency [2].

[2] Studies on the effects of recent diabetes diagnoses also reached similar conclusions [10, 11].

complications, the associated decline in productivity may be attenuated. Regarding preferences and expectations, the effect is ambiguous. On the one hand, compared to individuals who experience more severe illness, those who receive an early warning may place lower relative value on leisure, exhibit greater willingness to invest in work capacity, and postpone retirement. On the other hand, they might expect a longer working life and therefore either choose to invest more time in education or exhibit a temporarily weaker attachment to the labour market, reflecting a reduced urgency to work intensively. With regards to financial incentive, improvements in health may reduce reliance on disability insurance or other forms of income support.

However, as our treatment is a warning, we also need to consider that it contains a potential information shock. Therefore, the direction of the labour market effects is sensitive to the mechanism through which the warning operates. In particular, the warning may increase awareness of an underlying health condition, potentially influencing behaviour in different ways. If the information confirms the presence of a manageable condition — thereby explaining existing symptoms without implying severe consequences — the warning may be perceived as relatively positive. In such cases, leisure might be given less weight compared to work. Conversely, if the news is completely unexpected and perceived as negative, individuals might place greater value on leisure or re-evaluate their eligibility for disability benefits or early retirement. At the same time, in the negative scenario, the information shock may prompt an increase in labour supply, as individuals perceive a shortened time horizon for work.

In total, the overall impact of DM warnings on labour market outcomes is theoretically ambiguous and cannot be determined ex ante. Particularly in light of the undetermined direction of the labour market effects of an issued warning, we are able to make an important contribution to the cost-effectiveness analysis (CEA) of screenings. In fact, [13] and [14] report that more than 90%, of the CEA studies only include direct benefits and costs, that is, health outcomes and costs, disregarding important indirect costs and benefits, such as labour market effects. These omissions neglect the recommendations of the Second US Panel on Cost-Effectiveness, which advocates for the inclusion of indirect costs in CEAs [15].

Our focus on labour market effects of DM warnings also fills a gap in the broader literature on the interaction of health and labour market. Despite the wide-spread interest in this topic[3] little is known about the effects of health interventions on labour

[3]See for example [16] for an early review of the papers, and [17] for the heterogeneous effects of health level on employment

market outcomes in developed economies. An exception is [18], who evaluate the labour market effects of a health intervention to reduce the risk of coronary heart disease mortality in the United States.

We employ Dutch data from Lifelines, a collection of longitudinal health measurements of more than 160,000 participants, covering the years between 2006 and 2018. At baseline, Lifelines performed blood sample tests and sent a warning to participants who had glucose or HbA1c[4] values above the predetermined threshold. This paper estimates the labour market and health effects of receiving a warning with a multidimensional regression discontinuity design (RDD). By using this large survey sample, we are able to estimate warning effects for a country that has not (yet) introduced a screening programme. Furthermore, we do so for all ages, i.e. not only for older age groups that are usually targeted.

The diabetes screening literature has overlooked gender[5] differences, although descriptive studies document remarkable gender differences in diabetes incidence, diabetes treatment and health outcomes. For instance, diabetic women suffer from higher mortality rates than men due to cardiovascular complications caused by DM and receive less aggressive medical treatment [19, 20, 21]. Furthermore, men and women might also differ in labour market responses, particularly in the Netherlands, where women are highly likely to work part-time [22]. We therefore carefully differentiate between men and women.

We further leverage the richness of the dataset at hand in terms of demographic and personal characteristics by stratifying our results by age and educational level. Available studies on large-scale diabetes screenings, like the Korean case of [8] and the Japanese one of [9], only have information on individuals aged over 40. Thanks to the structure of Lifelines, we also provide the first results for individuals under 40. Understanding the effects on younger patients is crucial, as particularly the prevalence among younger individuals is increasing [23] and those who develop diabetes early in life experience more complications over time, i.e. higher cardiovascular risk factor [24], and even all-cause mortality [25]. Educational attainment is used as a proxy for socioeconomic

[4]HbA1c stands for glycosylated haemoglobin, it recovers the average blood sugar level over the past 2 to 3 months [2]. To obtain information on HbA1c, a blood sample test is needed.

[5]We use the term 'gender' to align with the administrative data being analyzed, as these records often reflect individuals' legally recognized gender markers; this is particularly relevant for transgender individuals who have updated their documentation to reflect their preferred gender. However, we acknowledge that this approach may still misclassify other gender identities beyond the binary categories of men and women, as such information is not captured in the administrative data and is therefore beyond our control.

status (SES) since studies on diabetes interventions[6] show that low-SES patients tend to benefit the least from them when their specific needs are not considered [26].

We find that low-educated, younger individuals, after receiving a marginal warning, increase their labour market activities. This effect is particularly strong and long-lasting for women. However, this is not accompanied by a similar remarkable change in our health indicators. Moreover, the health reaction by gender is quite heterogeneous. Concerning young individuals of lower educational attainment, for example, men reduce smoking, whereas women report a lower probability of being in very good health. Turning to other demographic groups, it should be noted that higher educated older women seem to gain the most in terms of health from receiving a warning. Their mortality rate declines significantly in the long term. This could be related to a higher probability of consulting a general practitioner and a higher rate of diabetes diagnosis. Remarkably, at the same time, we do not observe changes in labour market activities.

The rest of the paper is organised as follows: Section 2 describes Lifelines. In Section 3, we present our econometric model and in Section 4 we discuss descriptive statistics. Our empirical findings are presented in Section 5 and Section 6 discusses sensitivity analyses. Lastly, Section 7 concludes.

2 Data source, warning structure and typical treatment

Lifelines is a multi-disciplinary prospective population-based cohort study examining in a unique three-generation design the health and health-related behaviours of 167,729 persons living in the North of the Netherlands that started in 2007. It employs a broad range of investigative procedures in assessing the biomedical, socio-demographic, behavioural, physical and psychological factors which contribute to the health and disease of the general population, with a special focus on multi-morbidity and complex genetics.[7]

Lifelines collaborated with more than 70% of all general practitioners (GPs) in the provinces of Groningen, Friesland and Drenthe to recruit participants [27]. GPs invited

[6] Diabetes interventions usually target diagnosed diabetics and teach them to better manage their disease.

[7] The Lifelines initiative has been made possible by a subsidy from the Dutch Ministry of Health, Welfare and Sport, the Dutch Ministry of Economic Affairs, the University Medical Center Groningen (UMCG), Groningen University and the Provinces in the North of the Netherlands (Drenthe, Friesland, Groningen).

all their patients aged 25-50 who satisfied the following requirements: able to (1) understand Dutch, (2) fill in a questionnaire, (3) visit the GP. Moreover, they did not suffer from severe mental illness and did not have a terminal illness, meaning that they had a life expectancy of at least 5 years. In this way, almost 81,500 individuals were recruited. Then, Lifelines asked recruited individuals to invite their relatives, resulting in 64,500 additional participants. Additionally, around 21,500 individuals signed up via Lifelines' website.

Lifelines is organised in waves that differ in the content they provide. Figure 1 shows the time span of the waves and their content. Waves with the letter A in their names consist of two questionnaires and two in-person visits and they all followed the same procedure. Questionnaire 1 surveys socioeconomic variables, such as employment and education, and information on diabetes and other diseases. Questionnaire 2 contains data on other variables not surveyed in Questionnaire 1 (i.e. some health behaviours). Wave 1B only consists of Questionnaire 1. Notably, no in-person visit took place.

The timing of the questionnaires and in-person visits in waves A was as follows: participants received Questionnaire 1 at home around 21 days before the first visit. On the day of the appointment, among others, the participant's height and weight were measured. During the visit, Lifelines personnel checked whether Questionnaire 1 was adequately completed. The second visit happened on average 40 days after the first one. Questionnaire 2 was sent to the participants' homes and was completed about 7 days before the visit. During this second appointment, participants brought their urine samples to the centre and blood samples were collected. Visits took place before 10:00 AM and lasted 20 minutes. Participants were required to fast before the visit and we observe that only 2% of the participants did not comply. We exclude them from the sample.

A letter containing the results of the physical examination was first sent to their GPs 2 weeks after the examination and then to the participants after an additional 2 weeks. The letter contained results on about 15 physical measurements, of which HbA1c and fasting glucose were the blood sugar measures. When results were above the thresholds that Lifelines established as a healthy amount of blood sugar, the specific measurement was flagged and a written warning was reported on the results' sheet.

Participants received slightly different worded warnings based on their HbA1c and fasting glucose levels. The thresholds used by Lifelines are summarised in Table 1 and are graphically displayed in Figure 2.A. However, since the content of the warnings

is very similar, we group them in our analysis, as shown in Figure 2.B. According to the American Diabetes Association's (ADA) diagnostic criteria [28], individuals who received warnings 2 to 4 are undiagnosed diabetics. Those who received warning 1 are to a large extent considered prediabetics by the ADA [28]. However, the ADA considers glucose greater than or equal to 7 mmol/L the diagnostic threshold for diabetes, while Lifelines uses the threshold glucose strictly larger than 7 mmol/L.[8] Nonetheless, the discrepancy between the thresholds is small and the warnings have very similar wording. We further address this issue in a sensitivity analysis in Appendix B. Lastly, as will be explained in the next Section, our final estimation sample mostly consists of individuals who are in fact prediabetics and received warning 1.

We focus on the potential warning given in wave 1A, which took place around the year 2008. In case a general practitioner (huisarts) in the Netherlands identified a patient with prediabetes during this period, would follow a management plan focused primarily on lifestyle intervention and monitoring, rather than immediate medication. Official guidance at the time – for example, the NHG-Standaard (see [29]) and [30] emphasized diet and exercise changes, regular follow-up of blood glucose, and addressing cardiovascular risk factors. After identifying prediabetes, the GP would implement a plan for regular monitoring to catch progression to diabetes early and to guide management. Follow-up testing and risk monitoring were important aspects of care. Additionally, GPs or their practice nurses would provide education and motivation for lifestyle changes (weight management, regular exercise, smoking cessation). In practice, a patient with prediabetes around 2008 might see the GP every few months initially to support lifestyle change, then about once a year for blood tests if stable.

In case a patient was diagnosed with DM, a structured, stepwise approach was followed. Management began with lifestyle modification — emphasising dietary adjustment, increased physical activity, and weight reduction — similar to the one for prediabetes. Pharmacological intervention commenced if lifestyle changes failed after 3–6 months or if HbA1c was already significantly elevated (typically HbA1c $>$ 7.0–7.5 %), with metformin as the first-line agent due to its efficacy, favourable side-effect profile, and cardiovascular benefits. Sulfonylureas were commonly used as second-line therapy, while insulin was introduced in cases of inadequate glycemic control despite oral agents. Treatment goals targeted an HbA1c below 7%. Care delivery was integrated within primary care settings, where protocolized multidisciplinary management

[8]We note that to formally diagnose diabetes based on fasting glucose levels, an individual must exhibit glucose levels exceeding the diagnostic threshold on at least two separate occasions according to the ADA.

and annual complication screening became standard practice.

Our sample encompasses data from waves 1A, 1B and 2A of Lifelines, following individuals for an average of 4 years. On the basis of the blood sample results, we know who received a warning at baseline (wave 1A). Since we are interested in the effect of the warning on individuals previously unaware of their diabetes risk, we exclude from the sample those who self-reported either a diabetes diagnosis or were using diabetes medication at baseline (12,707 individuals dropped). We keep individuals older than 28 at baseline (26,207 individuals dropped). These individuals likely completed their studies and, without a prior diabetes diagnosis, are unlikely to have type 1 diabetes at the time of warning.[9] Furthermore, we exclude those older than 60 at baseline (36,535 individuals dropped) to ensure that participants did not reach the statutory retirement age by wave 2A. Nonetheless, some participants took early retirement. Lastly, we exclude from the sample women who were pregnant at baseline (254 individuals dropped) as pregnant women with high blood sugar values might not suffer from type 2 DM but gestational diabetes. Unfortunately, we do not have information on gestational diabetes.

Regarding the outcome variables, participants self-reported employment, working hours, disability insurance, and (early) retirement. We assign a value of 1 to employment if they reported being employed irrespective of hours worked, and a value of 0 to individuals who were (early) retired, homemakers or unemployed. Individuals under disability insurance can also report being employed. Thus, in some cases, both dummies for employment and disability insurance can equal 1. We unfortunately do not have access to earnings information, and, therefore, cannot analyse the effects on income.

With regard to the health variables, we use information on health behaviour and objective health measures that are available in our data set and common in the literature: The GP visit variable equals 1 when participants reported visiting their GP at least once last year. We classify participants as smokers if they self-reported smoking in the last month and as alcohol drinkers if they reported consuming alcohol at least once in the same period. The dummy variable for follow-up physical activity equals 1 for individuals who reported being at least slightly more physically active compared to their baseline level. As objective health measures we classify the following: The DM diagnosis dummy equals 1 for those who self-report a new diagnosis. BMI was self-reported

[9] This type of DM is developed in early life stages and those affected by it are unlikely to survive until 30 without a proper diagnosis and treatment, as this type of diabetes often requires insulin treatment. To give an idea of how fundamental access to cures is for type 1 diabetics, the life expectancy of children with type 1 diabetes in rural Mozambique, where diabetes care is of poor quality, is 7 months [31].

in wave 1B but measured by Lifelines during the visits in wave 2A. HbA1c is retrieved from the blood sample test data. The mortality dummy equals 1 for those who died before 2018, which is the last year of wave 2A. In addition, our data allow us to present results for self-reported health, which have not been explored in previous studies that rely primarily on administrative data lacking this type of information. Here we create a dummy variable for very good health equals 1 for individuals who self-reported at least very good health. Note, individuals were asked to rate their health from 1 "poor" to 5 "excellent". We code very good health answers equal to or higher than 4.

3 Model

To investigate the effect of the diabetes risk warning received at baseline on labour market and health outcomes, we employ a regression discontinuity design (RDD), which estimates a local treatment effect. This method addresses endogeneity concerns, such as bias from omitted variables like unobserved health status, that would arise if employment were regressed directly on the warning indicator. The RDD circumvents this issue by exploiting the existence of an arbitrary threshold used to assign treatment and the fact that individuals just below and above it have similar underlying characteristics.

We employ a sharp RDD, as noncompliance is impossible due to Lifelines automatically sending warnings for blood sugar values exceeding the defined thresholds. As discussed in Table 1, Lifelines sent out different types of warnings depending on the thresholds. However, we focus on the effect of receiving any warning regardless of type, as the wording differences are minimal. We will also investigate this issue further in a sensitivity analysis. Thus, in our main analysis, we define the treatment for individual i in wave 1A as

$$t_i \equiv \mathbb{1}(HbA1c_{i1A} \geq 6.1\% \mid glucose_{i1A} \geq 7.1 mmol/L) \tag{1}$$

where t_i is the treatment indicator for individual i who received a warning in wave 1A. Due to the presence of two running variables in the warning assignment rule, glucose and HbA1c, a multidimensional setting is necessary. To use a multidimensional RDD, we first need to account for the difference in information contained in the running variables. Indeed, glucose measures the short-run blood sugar average and HbA1c the long-run one, necessitating a standardization of both running variables. We employ the standard deviation computed by gender (indexed as g),[10] as we will conduct the analysis

[10]The standard deviation of glucose is 0.65 for men and 0.56 for women. The standard deviation of

separately by gender. Furthermore, we also centre the variables at their respective cutoff values.

$$glucose\_std_i = \frac{glucose_i - 7.1}{sd_g(glucose_i)}$$
$$HbA1c\_std_i = \frac{HbA1c_i - 6.1}{sd_g(HbA1c_i)}$$

To have a single running variable, as required by the RDD, we create a new running variable equal to the minimum distance from the cutoff of both glucose and HbA1c [32]. Furthermore, since glucose and HbA1c are centred and standardised, the new running variable is centred and standardised as well. The new running variable z_i is computed as follows

$$z_i = \begin{cases} -min(\mid glucose\_std_i \mid, \mid HbA1c\_std_i \mid) & \text{if } glucose\_std_i < 0 \ \& \ HbA1c\_std_i < 0 \\ \mid glucose\_std_i \mid & \text{if } glucose\_std_i \geq 0 \ \& \ HbA1c\_std_i < 0 \\ \mid HbA1c\_std_i \mid & \text{if } glucose\_std_i < 0 \ \& \ HbA1c\_std_i \geq 0 \\ \sqrt{glucose\_std_i^2 + HbA1c\_std_i^2} & \text{if } glucose\_std_i \geq 0 \ \& \ HbA1c\_std_i \geq 0 \end{cases} \tag{2}$$

The observed outcome is

$$y_i = (1 - t_i) \cdot y_i(0) + t_i \cdot y_i(1) = \begin{cases} y_i(0) & if \quad z_i < 0 \\ y_i(1) & if \quad z_i \geq 0 \end{cases} \tag{3}$$

where y_i is the observed outcome. Following the usual notation in the causal inference literature, $y_i(0)$ is the potential outcome in the absence of treatment and $y_i(1)$ is the potential outcome in the presence of treatment.

The sharp RDD design estimates the following treatment effect:

$$\tau \equiv \mathbb{E}[(y_i(1) - y_i(0)|z_i = 0] = \lim_{z \downarrow 0} \mathbb{E}[(y_i|z_i = 0)] - \lim_{z \uparrow 0} \mathbb{E}[(y_i|z_i = 0)] \tag{4}$$

the treatment effect τ is local, it reflects the causal impact of marginally crossing either of the cutoffs. To estimate τ, we estimate Equation 5. We stratify results by gender and estimate the effect of marginally receiving a warning at baseline on the outcomes (y_{iw}) in waves $1B$ (on average 1.5 years later) and $2A$ (on average 4 years later).

$$y_{iw} = \beta_0 + \tau t_{i1A} + \beta_1 z_{i1A} + \beta_2 z_{i1A} \cdot t_{i1A} + \beta_3 y_{i1A} + \mathbf{X}'_{i1A}\gamma + \epsilon_{iw} \qquad w = 1B, 2A \tag{5}$$

HbA1c is 0.36 for men and 0.33 for women.

we report this coefficient in the next sections. The treatment dummy is t_{i1A}, and z_{i1A} the centred running variable, $\mathbf{X_{i1A}}$ a vector of demographic controls measured at baseline (namely, a second-order polynomial of age, educational level, and whether the individual was born outside of the Netherlands), year of inclusion at baseline and current year fixed effects. By including y_{i1A} as a control, we correct for possible baseline differences in the outcome variable, as suggested in [33].

To estimate Equation 5, we employ a weighted OLS, where a triangular kernel assigns larger weights to observations close to the cutoff, and we apply a first-order polynomial on both sides of the cutoff.[11] Following [35], we choose a linear polynomial to avoid overfitting, as the bandwidth is narrow. We estimate Equation 5 only for those in a bandwidth of 1.5 standard deviations (sd) from the cutoff. We enforce this bandwidth to guarantee the compatibility of our results across the various outcomes, as data-driven bandwidths suggested by [35] change for each outcome. Notably, with the 1.5 sd bandwidth, as reported in Table 2, the treated group is largely made of those who received a prediabetes warning. Given the discrete nature of the running variable, we do not cluster the standard errors at every unique point of the running variable because such practice leads to poor coverage properties [36].[12] Instead, following [36], we estimate heteroskedasticity-robust standard errors, which are provided by an option of the rdrobust package [34].

To guarantee a causal interpretation of the estimates two conditions need to be fulfilled: 1) sorting in or out of treatment is impossible, thus, the running variable cannot be manipulated; 2) individuals around the threshold are similar. Regarding the first assumption, our running variables are unlikely to be manipulated. First, participants did not know the exact cutoffs for the diabetes risk warning before receiving the letter. Second, participants receive a warning if either threshold is crossed. While glucose levels reflect short-run blood sugar levels and may be easier to manipulate, this is not the case for HbA1c. HbA1c reflects blood sugar levels over the past two to three months [37]. HbA1c manipulation would only be possible if participants changed their lifestyle for at least two months before the lab test. The absence of visible discontinuities in HbA1c and glucose levels at the cutoff values, as shown in Figure 3, provides evidence for our claims. Furthermore, in Tables B7-B9 in Appendix B, as a sensitivity check, we replicate the main analysis using HbA1c as the sole running variable to identify

[11] We use the rdrobust package developed by Calonico et al. [34]

[12] They find that, for small bandwidths, standard errors clustered at every unique value of the running variable produce confidence intervals that have an actual coverage rate much lower than the nominal 95 percent. The coverage rate here is the proportion of times that a confidence interval successfully contains the true value of the parameter being estimated.

the effect of a warning for individuals crossing only the HbA1c 6.1% threshold. As discussed above, this group contains most of the treated observations. We find similar findings to those discussed in our main result section (Section 5.1).

The second assumption implies that individuals around the threshold only differ because of the warning they have received. To validate this assumption, we follow a common approach in the literature by assessing whether predetermined covariates exhibit discontinuities at the cutoff at baseline. Any such discontinuities would indicate potential preexisting differences between individuals above and below the threshold. We estimate a model similar to Equation 5, using the baseline value of the control variable as the outcome and excluding controls from the right-hand side. Following [35], for this test, we use bandwidths that minimise the asymptotic coverage error rate (CER) of the bias-corrected confidence interval. We show the results of this falsification test in Table 3. We find that only women's smoking behaviour varies discontinuously at the cutoff at the 90% level. We replicate this exercise for the other subgroups present in our analysis in Tables C1 and C2 in Appendix C.

4 Descriptives

In this section, we present baseline descriptive statistics for individuals in the range of 1.5 sd below the cutoff (control group) and 1.5 sd above the cutoff (treated group), based on the standard deviation of the running variable computed by gender.

Looking at the descriptive statistics for the whole sample in Table 4, we find similar patterns for both genders in differences across the threshold. Regarding the control variables, those in the range of 1.5 sd above the threshold are significantly older, less likely to have a university or an upper secondary education, more likely to have lower secondary education and to be born outside the Netherlands. Considering the outcome variables, the differences are significant for all variables except retirement. Those under the threshold are more likely to be employed (this is even more so for women), work more hours, are less likely to be on disability insurance and (insignificantly) less likely to be retired. Furthermore, individuals 1.5 sd below the cutoff are less likely to visit their GP, smoke, are more likely to drink alcohol and report at least very good health. While for both genders those below the cutoff have lower BMI, the differences around the cutoff are more pronounced for women than for men.

When comparing age groups across the threshold in Tables D1 and D2 in Appendix D, descriptive statistics for individuals older than 40 are very similar to those of the

complete sample. This is unsurprising as most of those who received a warning are aged 40 or above. Among individuals under 40, while many differences across the threshold lose significance, their signs generally align with those observed in the overall sample. The few instances where the direction differs are when the differences for individuals under 40 are insignificant (i.e. smoking probability for men).

5 Results

In Section 5.1, we present the main results, where we stratify by baseline age groups (28–40 and 41–60) and gender. The age stratification has two major advantages. Firstly, we are able to investigate whether the age restriction for screening programmes that other countries impose is suitable for the Netherlands. Additionally, it takes into account that individuals diagnosed with diabetes before the age of 40 have higher risks of diabetes complications [38, 39] and all-cause mortality [25] than those diagnosed after 40. Section 5.2 extends the analysis by further stratifying the sample according to educational attainment. The graphical representation of the results by age and gender can be found in Figures E1 to E12 in Appendix E.

5.1 Main results

As shown in Table 5, younger women experience a significant positive increase in their employment rate of around 8 percentage points (pp) in both, the short-term and long-term, after marginally crossing either of the warning thresholds. The short-term effect is partly driven by a significant reduction in the take-up of disability insurance of 2 pp. There are signs that the weekly hours of younger women also increase after receiving a warning; however, this effect is not significant. In contrast, older women do not engage more actively in labour market activities after receiving a warning. Neither employment rate, weekly hours, nor disability insurance take-up show economically or statistically significant changes. Moreover, in the long-term older women are more likely to retire by nearly 2 pp.

The labour market results of men somewhat mirror the results of women. In the short term, young men below the age of 40 experience a significant increase in the employment rate of 4 pp accompanied by a significant decrease in disability insurance take-up of 2 pp due to a warning. Additionally, weekly hours increase significantly by more than 2 hours per week. However, in contrast to younger women, these effects seem to evaporate in the long term. With regard to employment rate and weekly hours, older men, similarly to older women, do not show a change in work attachment after

receiving a warning, as neither labour market measure changes significantly. However, notably the uptake of disability insurance of older men significantly decreases in the long term by 2 pp and no change in retirement take-up is detectable.

Thus, for younger individuals, we find positive short-term employment effects of a diabetes warning which persist for women in the long term. In contrast, the employment outcomes of older individuals seem to be only affected in the long term showing opposing directions by gender. The retirement rate of women increases whereas the disability insurance take-up of men decreases, both significantly. However, neither change translates into a significant change of the employment rate.

Additionally, we are interested in whether diabetes warnings lead to changes in health-related outcomes (see Table 6 and 7). Not only is this interesting from a medical perspective, but could also provide information on potential pathways for the labour market effects. For this analysis, we use health behaviours and outcomes common in the literature. Furthermore, our data allow us to present results for self-reported health, which have not been explored in previous studies that rely primarily on administrative data lacking this type of information.

The most striking effect is the reduction in being in very good health for younger women when passing the warning threshold. In the short term, we observe a 10 pp statistically significant reduction; the long-term coefficient still remains at 7 pp, although not statistically significant. With regards to health behaviour and health outcomes, neither statistically nor economically relevant changes in health behaviour and health outcomes can be observed for younger women in the short term. In the long term, some coefficients turn signs and/or become larger in absolute terms. However, they all stay statistically insignificant.

Women 40 and older react differently when crossing the warning threshold than younger women not only with regards to the labour market outcomes but also with regards to the health outcomes. Surprisingly, receiving a warning at the margin increases the percentage of older women reporting being in very good health by around 3 pp. in the short term (at the 10 percent significance level). An effect that does, however, not persist in the long term, as it reduces to close to zero. Interestingly, we find a sizable although insignificant increase in the probability of exercising of 2 pp. However, all other short-term measures of health behaviour, like GP visits, smoking, as well as the objective health measures (diabetes diagnosis, BMI and HbA1c) are close to zero.

Also in the long-term, no changes with regards to GP visits or smoking are observable for older women after a marginal warning, a similar result as for the short run. However, as a long-term measure, we have information available on the alcohol consumption, which is not collected in the earlier wave. This shows a sizable reduction of drinking alcohol by 3 pp (significant on the 10 percent level). Another notable finding is that the percentage of diagnosed older women increases in the long term by around 3 pp when warned at the margin. No changes are observable with regards to BMI and HbA1c for older women. In particular, the result with regards to HbA1c is consistent with previous studies ([8], [9]). Furthermore, we detect a sizable effect on the 4-year mortality rate of app. -0.6 pp, which is, however, not significant.

When examining health-related responses to a marginal warning of men and comparing these to women's responses, distinct gender patterns emerge. Among younger men, smoking prevalence declines by approximately 7 pp in the short term and 8 pp in the longer term, with both estimates reaching statistical significance at the 10 percent level. These behavioural adjustments are not mirrored among younger women, for whom no discernible changes in smoking are observed. Also in contrast to women, we find indications that older men increase their use of general practitioner (GP) services, with the probability of a yearly consultation rising by approximately 4 pp following the warning. Conversely, while women appear to revise their self-assessed health status in response to the threshold, we observe no statistically significant changes in self-reported health among any of the male subgroups considered. Moreover, older men, in contrast to older women, are not significantly more likely to receive a diabetes diagnosis.

The finding that young women are significantly less likely to report at least very good health after a crossing of the warning thresholds can be explained in different ways: 1) warned individuals became more self-conscious about their health because of the warning and reported lower health without actually experiencing any new issues; 2) the decrease in self-reported health reflects a worsening of their underlying health. Given that younger women after crossing the warning threshold actually do not show changes in other health-related indicators and in addition even reduce the take-up of disability insurance, we think that the first explanation is more likely. We derive supporting evidence by [40] and [41], who document that younger individuals tend to underestimate their health and base their health assessment on diagnosis rather than functioning, whereas older individuals tend to overestimate their health. These assessments might be particularly true after an information shock, as we not only observe that younger women are less likely to be at least in very good health, but also that older women are more likely to report being at least in very good health after the marginal

warning.

Interpreting the effect of the warning on the probability of diabetes diagnosis also calls for further clarification. We consider diagnosed with DM those who self-reported a diagnosis between waves as medication data is unavailable after baseline. As discussed in Appendix A, if participants' reporting behaviour remained consistent throughout the waves, we expect most individuals diagnosed between waves to report their new status. At first glance, the coefficient's size might seem surprisingly small on the diagnosis variable for either gender. To explain this, we note that the RDD estimates the effect of crossing either the HbA1c 6.1% or glucose 7.1 mmol/L thresholds or both. As reported in Table 2, more than 80% of those warned only crossed the HbA1c threshold and can thus be considered prediabetic. Prediabetes is a phase before diabetes when individuals can prevent DM from developing through lifestyle changes. Consequently, prediabetic individuals should not be diagnosed with DM unless their condition deteriorates or their GPs deem it appropriate.

Moreover, our findings indicate a statistically insignificant increase in diabetes diagnosis rates for older men, whereas, for older women, the increase (2.8 pp) is significant in the long run. To explain the gender differences in diagnosed diabetes, particularly in the long run, we refer to the literature on gender differences in diagnosis. [42] and [43] find that, on average, men are diagnosed 3-4 years earlier than women. This trend of earlier diagnosis for men seems to be confirmed in our data, as the average diagnosis rate 0.5 sd below the cutoff is higher for men than for women in the long run.

Following the broader literature on disease diagnoses, differences in diagnosis rates between genders could stem from the dynamics of the patient-GP relationship [44]. Indeed, in the Netherlands, men are more likely to receive a diagnosis than women when they visit their GP for common somatic symptoms [45]. This difference is partly driven by women being prescribed fewer diagnostic interventions (i.e. laboratory analyses and physical examinations) than men. However, even when diagnostic interventions are performed, women are still less likely to be diagnosed with a disease than men [46]. In our context, warned women not only receive a diagnostic intervention (the blood sample test) like in the case of [46] but also a written warning from Lifelines. The significant increase in DM diagnosis for older women might be driven by women feeling more confident in asking their GP for further checks and eventually obtaining a diagnosis, or by GPs, who also receive the warning from Lifelines, taking such warnings more seriously.

5.2 By education level

In this section, we stratify the analysis by educational level (below university education vs. university education), serving as a proxy for SES. This stratification is motivated by evidence that individuals with lower SES are more likely to experience diabetes-related complications over time [47] and present a higher incidence of (diagnosed) diabetes [48].

When examining the labour market outcomes for women, as reported in Table 8, it is evident that the positive employment effects among younger women following a marginal diabetes warning are concentrated among those with lower educational attainment. Within this group, we observe sizable increases in employment, including a rise of approximately 10 pp in employment probability and an average increase of 4 hours in weekly working time. Also among older women, heterogeneity by education is salient. The observed increase in retirement appears to be concentrated among the low-educated, whereas the decline in disability uptake is more pronounced among the higher-educated.

A broadly similar pattern emerges for younger men, as shown in Table 9. Similar to women, the effects found in the main sample are largely driven by lower levels of education. For younger men with lower educational attainment, we find a short-term increase in employment, while for older men with lower educational attainment, the long-term decline in employment rate of nearly 5 pp is the more salient feature, part of which is accounted for by an increase in retirement. In addition, we find a statistically significant short-term reduction in disability claims of about 2 pp among older men. At first glance, the concurrent decline in both employment and disability claims may seem difficult to reconcile. However, it is important to note that the data include part-time disability, such that disability and employment are not mutually exclusive. In addition, some individuals may transition from disability to unemployment or exit the labour force entirely, without necessarily entering retirement. Finally, we find a short-run increase in the employment probability of older men with high education, of roughly 4 pp.

Thus, among younger individuals, positive employment responses are concentrated among the low-educated, with especially strong effects for women. Older individuals exhibit more mixed responses: low-educated women and men show increased retirement, while higher-educated groups reduce disability uptake. Moreover, there is a pronounced heterogeneity among older men: the warning leads to both lower employment for the lower educated in the long-term and higher employment in the short-term for the highly

educated.

Turning to the health outcomes for women, as reported in Tables 10 and 11, younger women of both educational groups appear to report lower levels of self-assessed health following the warning, though this pattern is more pronounced among younger, highly educated women. It should be noted, however, that these effects are no longer statistically significant. Among the same group, we observe a short-term reduction in BMI of approximately 0.7 points, suggesting some behavioural response in terms of weight management. Among older women, the patterns differ by education. Those with lower educational attainment show a decline in general practitioner visits over the longer term - a somewhat unexpected finding - while their higher-educated counterparts exhibit increases in health service use, including more frequent doctor visits and a higher incidence of diabetes diagnoses. Most remarkably, this group also shows a significant decline in the 4-year mortality rate by 1.5 pp.

As in the main analysis, we observe only limited changes in health outcomes for men upon crossing the warning thresholds, even when stratifying by educational attainment (Tables 12 and 13). The patterns remain distinct from those observed for women. The reduction in smoking is concentrated among low-educated men, suggesting that health behaviour responses are not uniform across groups. A further notable finding is that older, higher-educated men are less likely to rate their health as very good following the warning (around -8 pp) - a result that was not apparent in the main sample.

6 Sensitivity analysis

To show the robustness of our results, we replicate the analysis by gender and age, taking into account multiple hypothesis testing and varying the choice of kernel, bandwidth, and polynomial. As reported in Tables B1-B12 in Appendix B, our analysis is robust to the use of a uniform kernel. Employing a smaller bandwidth (1 sd), most of our results still hold. However, while the sign and size stay similar, the short-term results for employment and hours worked for men under 40 become insignificant, probably due to a loss of statistical power due to the smaller sample size. Lastly, when considering a second-order polynomial, the sign of the estimated coefficients remains similar but the estimates for labour market outcomes become less precise, likely due to overfitting, as discussed in Section 3.

In the analysis, we estimate the same equation for many outcomes. Due to random chance, we run the risk of finding statistically significant coefficients. To address this

issue, we compute p-values that account for multiple hypothesis testing within each domain. We partition our outcomes into three domains, namely, labour market outcomes, health behaviours and health outcomes. Staying on the conservative side, we include self-reported health in the health outcomes domain. For each domain, we compute the false discovery rate (FDR) based on the method proposed in [49]. FDR is the expected proportion of rejections that are type I errors (false rejections) and is suitable for randomised control trials. We apply this method because, in the RDD setup, the treatment can be considered randomly assigned near the threshold when individuals cannot manipulate the running variable [33].

As shown in Tables F1-F9 in Appendix F, most of the results from Section 5 are also significant with the FDR correction. For example, in the analysis by age and gender, the short-term labour market results remain statistically significant. On the other hand, some health behaviours and outcomes lose statistical significance, such as the lower probability of drinking alcohol for women over 40 and the higher probability of visiting the GP for men over 40. Notably, in very few cases, the FDR p-values are lower than the conventional p-values. This can happen when many null hypotheses are rejected because, when many true rejections happen, this method tolerates some false rejections too, but still maintains the false discovery rate low.

7 Discussion and conclusions

We conduct a comprehensive evaluation of the effects of a health intervention, namely a diabetes warning. Building on the literature on diabetes screening's effectiveness, we are the first to analyse the employment effects of a diabetes risk warning. We further contribute to the literature by presenting the first results on health outcomes and behaviours by gender as well as age, and by providing evidence on effect heterogeneity by educational level. We used data from Lifelines, a cohort study based in the North of the Netherlands, in which participants received diabetes risk warnings after performing blood tests. With a multidimensional RDD, we estimated the effect of receiving a warning on labour market and health outcomes.

Our findings show that low-educated, younger individuals increase their labour market activities after receiving a marginal warning. The effect is particularly strong and long-lasting for women. However, this is not accompanied by a similar remarkable change in our health indicators. Additionally, the health reactions by gender are quite heterogeneous. Concerning young individuals of lower educational attainment, for example, men reduce smoking, whereas women report a lower probability of being in

very good health. Turning to other demographic groups, older individuals exhibit more mixed responses concerning labour market results: low-educated women and men show increased retirement rates, while higher-educated groups reduce disability uptake. Moreover, there is a pronounced heterogeneity among older men: the warning leads to both lower employment for the lower educated in the long term and higher employment in the short term for the highly educated.

In addition, it should be noted that highly educated older women seem to gain the most in terms of health from receiving a warning. Their mortality rate declines significantly in the long term. This could be related to a higher probability of consulting a general practitioner and a higher rate of diabetes diagnoses. Remarkably, at the same time, we do not observe changes in labour market activities.

Thus, we do not find a strong one-directional connection between labour market activities and change in health indicators after receiving a warning, which might not be surprising, given the multitude of theoretical predictions. A potential explanation for the finding that particularly young, lower educated individuals increase their labour market activities could be related to working more frequently in manual jobs than other demographic groups. For these activities, diabetes-related side effects might be particularly detrimental. Although we do not find strong health-related changes, it might be that some actually take place, given the medical trajectory that individuals usually take after a warning, but are not captured by our indicators. Thus, this suggests the employment effects might be related to changes in productivity. An alternative explanation is that the diabetes warning may convey to individuals that a pre-existing but previously undiagnosed health condition is both identifiable and manageable. This reassurance could plausibly increase individuals' perceived health capital and thereby strengthen their engagement with the labour market.

In light of the positive labour market responses observed among younger individuals, there may be merit in considering the introduction of a screening programme also to younger age groups, as the results indicate the potential of screening programmes to at least partly counteract the commonly found negative employment effect of diagnosed diabetes. Furthermore, our findings highlight the importance of incorporating productivity measures when assessing the effectiveness of screening programmes, a practice that remains uncommon. As noted by [50] in their review of cost-effectiveness study guidelines across various countries, only 35% of the guidelines recommended including the impacts of indirect costs, such as changes in labour market outcomes.

Furthermore, the health benefits of receiving a warning for older, higher educated women should receive particular attention, as they point towards room for improvement in the interaction between female patients and GPs. These warnings lead simultaneously to a higher rate of GP visits, diabetes diagnosis and lower mortality for this particular group. From a public health perspective, a variety of relevant questions arise in this context, among which are: Why do only highly educated women benefit so strongly in health terms from the warning? Why does it appear that women suffer from a delayed diabetes diagnosis? Answering these questions is left for future research.

8 CRediT authorship contribution statement

Claudio Annibali: Conceptualization, Data Curation, Formal analysis, Methodology, Writing – original draft, Funding acquisition. **Annette Bergemann**: Conceptualization, Methodology, Supervision, Writing – original draft, Funding acquisition. **Rob Alessie**: Supervision, Writing – original draft, Funding acquisition.

9 Funding statement

This work was supported by the University of Groningen and the Centre for Public Health in Economics and Business.

10 Acknowledgements

The authors wish to acknowledge the services of the Lifelines Cohort Study, the contributing research centres delivering data to Lifelines, and all the study participants. The authors acknowledge the Centre for Public Health in Economics and Business for providing the necessary funding. The authors thank Dalton Conley, Guglielmo Cori, Zoë McLaren, Jochen Mierau, David Richter, Matthias Rieger, and Jannis Stöckel for their feedback. Lastly, the authors are thankful for the comments and suggestions received during the European Society for Population Economics conference, the Lowlands Health Economists' Study Group conference, the American-European Health Economics Study Group 9th Edition, the Joint Workshop CPHEB x ZEW Mannheim, the VIU Summer Institute on Ageing, the Second Dutch Health Econometrics conference, Workshop in Causal Health Economics in Uppsala, FEBRI PhD conference and World Congress of Epidemiology as well as seminars at Stellenbosch University, SALDRU (Cape Town University), and Ca' Foscari University.

11 Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT to improve the readability of the text. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

References

[1] International Diabetes Federation. (2025). *Idf diabetes atlas, 11th edn* (tech. rep.). International Diabetes Federation. Brussels, Belgium. https://www.diabetesatlas.org

[2] American Diabetes Association. (2013). Diagnosis and classification of diabetes mellitus. *Diabetes Care*, *37*(Supplement_1), S81–S90. https://doi.org/10.2337/dc14-s081

[3] Peters, M., Huisman, E., Schoonen, M., & Wolffenbuttel, B. (2017). The current total economic burden of diabetes mellitus in the Netherlands. *Neth. J. Med.*, *75*(7), 281–297.

[4] Bjarkø, V. V., Haug, E. B., Sørgjerd, E. P., Stene, L. C., Ruiz, P. L.-D., Birkeland, K. I., Berg, T. J., Gulseth, H. L., Iversen, M. M., Langhammer, A., & Åsvold, B. O. (2022). Undiagnosed diabetes: Prevalence and cardiovascular risk profile in a population-based study of 52,856 individuals. The hunt study, Norway. *Diabet. Med.*, *39*(6). https://doi.org/10.1111/dme.14829

[5] Paprott, R., Rosario, A. S., Busch, M. A., Du, Y., Thiele, S., Scheidt-Nave, C., & Heidemann, C. (2014). Association between hemoglobin a1c and all-cause mortality: Results of the mortality follow-up of the German national health interview and examination survey 1998. *Diabetes Care*, *38*(2), 249–256. https://doi.org/10.2337/dc14-1787

[6] Pedron, S., Emmert-Fees, K., Laxy, M., & Schwettmann, L. (2019). The impact of diabetes on labour market participation: A systematic review of results and methods. *BMC Public Health*, *19*(1). https://doi.org/10.1186/s12889-018-6324-6

[7] Bergemann, A., Grönqvist, E., & Guðbjörnsdóttir, S. (2019). Health and labor markets, research in labor Economics. In *Diabetes morbidity after displacement* (pp. 99–154). Emerald Publishing Limited. https://doi.org/10.1108/S0147-912120190000047005

[8] Kim, H. B., Lee, S. A., & Lim, W. (2019). Knowing is not half the battle: Impacts of information from the national health screening program in Korea. *J. Health Econ.*, *65*, 1–14. https://doi.org/10.1016/j.jhealeco.2019.01.003

[9] Iizuka, T., Nishiyama, K., Chen, B., & Eggleston, K. (2021). False alarm? Estimating the marginal value of health signals. *J. Public Econ.*, *195*, 104368. https://doi.org/10.1016/j.jpubeco.2021.104368

[10] Alalouf, M., Miller, S., & Wherry, L. R. (2024). What difference does a diagnosis make?: Evidence from marginal patients. *Am. J. Health Econ.*, *10*(1), 97–131. https://doi.org/10.1086/724415

[11] Gaggero, A., Gil, J., Jiménez-Rubio, D., & Zucchelli, E. (2022). Does health information affect lifestyle behaviours? The impact of a diabetes diagnosis. *Soc. Sci. Med.*, *314*, 115420. https://doi.org/10.1016/j.socscimed.2022.115420

[12] Blundell, R., Britton, J., Dias, M. C., & French, E. (2021). The impact of health on labor supply near retirement. *J. Hum. Resour.* https://doi.org/10.3368/jhr.58.3.1217-9240R4

[13] Krol, M., Papenburg, J., Tan, S. S., Brouwer, W., & Hakkaart, L. (2015). A noticeable difference? Productivity costs related to paid and unpaid work in economic evaluations on expensive drugs. *Eur. J. Health Econ.*, *17*(4), 391–402. https://doi.org/10.1007/s10198-015-0685-x

[14] Pike, J., & Grosse, S. D. (2018). Friction cost estimates of productivity costs in cost-of-illness studies in comparison with human capital estimates: A review. *Appl. Health Econ. Health Policy.*, *16*(6), 765–778. https://doi.org/10.1007/s40258-018-0416-4

[15] Sanders, G. D., Neumann, P. J., Basu, A., Brock, D. W., Feeny, D., Krahn, M., Kuntz, K. M., Meltzer, D. O., Owens, D. K., Prosser, L. A., Salomon, J. A., Sculpher, M. J., Trikalinos, T. A., Russell, L. B., Siegel, J. E., & Ganiats, T. G. (2016). Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses: Second panel on cost-effectiveness in health and medicine. *JAMA*, *316*(10), 1093. https://doi.org/10.1001/jama.2016.12195

[16] Currie, J., & Madrian, B. C. (1999). Health, health insurance and the labor market. In C. Dustmann & T. Lemieux (Eds.), *Handbook of labor Economics* (pp. 3309–3416, Vol. 3). Elsevier. https://doi.org/https://doi.org/10.1016/S1573-4463(99)30041-9

[17] Blundell, R., Britton, J., Dias, M. C., French, E., & Zou, W. (2022). *The dynamic effect of health on the employment of older workers: Impacts by gender, country, and race* (tech. rep.). Michigan Retirement and Disability Research Center.

[18] Stephens, J., Melvin, & Toohey, D. (2022). The impact of health on labor market outcomes: Evidence from a large-scale health experiment. *AEJ: Applied*, *14*(3), 367–99. https://doi.org/10.1257/app.20180686

[19] Sarafidis, P. A., McFarlane, S. I., & Bakris, G. L. (2006). Gender disparity in outcomes of care and management for diabetes and the metabolic syndrome. *Curr. Diab. Rep.*, *6*(3), 219–224. https://doi.org/10.1007/s11892-006-0038-3

[20] Gnatiuc, L., Herrington, W., Halsey, J., Tuomilehto, J., Fang, X., Kim, H., & De Bacquer, D. e. a. (2018). Sex-specific relevance of diabetes to occlusive vascular and other mortality: A collaborative meta-analysis of individual data from 980,793 adults from 68 prospective studies. *Lancet Diabetes Endocrinol.*, *6*(7), 538–546. https://doi.org/10.1016/s2213-8587(18)30079-2

[21] Rossi, M. C., Cristofaro, M. R., Gentile, S., Lucisano, G., Manicardi, V., Mulas, M. F., Napoli, A., Nicolucci, A., Pellegrini, F., Suraci, C., & Giorda, C. (2013). Sex disparities in the quality of diabetes care: Biological and cultural factors may play a different role for different outcomes. *Diabetes Care*, *36*(10), 3162–3168. https://doi.org/10.2337/dc13-0184

[22] OECD. (2019, September). *Part-time and partly equal: Gender and work in the Netherlands* (tech. rep.). OECD. https://doi.org/10.1787/204235cf-en

[23] International Diabetes Federation. (2021). *Idf diabetes atlas, 10th edn* (tech. rep.). International Diabetes Federation. Brussels, Belgium. https://www.diabetesatlas.org

[24] Barker, M. M., Zaccardi, F., Brady, E. M., Gulsin, G. S., Hall, A. P., Henson, J., Htike, Z. Z., Khunti, K., McCann, G. P., Redman, E. L., Webb, D. R., Wilmot, E. G., Yates, T., Yeo, J., Davies, M. J., & Sargeant, J. A. (2022). Age at diagnosis of type 2 diabetes and cardiovascular risk factor profile: A pooled analysis. *World J. Diabetes*, *13*(3), 260–271. https://doi.org/10.4239/wjd.v13.i3.260

[25] Sattar, N., Rawshani, A., Franzén, S., Rawshani, A., Svensson, A.-M., Rosengren, A., McGuire, D. K., Eliasson, B., & Gudbjörnsdottir, S. (2019). Age at diagnosis of type 2 diabetes mellitus and associations with cardiovascular and mortality risks: Findings from the Swedish national diabetes registry. *Circulation*, *139*(19), 2228–2237. https://doi.org/10.1161/circulationaha.118.037885

[26] Hardman, R., Begg, S., & Spelten, E. (2020). What impact do chronic disease self-management support interventions have on health inequity gaps related to socioeconomic status: A systematic review. *BMC Health Serv. Res.*, *20*(1). https://doi.org/10.1186/s12913-020-5010-4

[27] Scholtens, S., Smidt, N., Swertz, M. A., Bakker, S. J., Dotinga, A., Vonk, J. M., van Dijk, F., van Zon, S. K., Wijmenga, C., Wolffenbuttel, B. H., & Stolk, R. P. (2014). Cohort profile: Lifelines, a three-generation cohort study and biobank. *Int. J. Epidemiol.*, *44*(4), 1172–1180. https://doi.org/10.1093/ije/dyu229

[28] American Diabetes Association. (2017). 2. classification and diagnosis of diabetes: Standards of medical care in diabetes—2018. *Diabetes Care*, *41*(Supplement_1), S13–S27. https://doi.org/10.2337/dc18-s002

[29] Rutten, G., De Grauw, W., Nijpels, G., Goudswaard, A., Uitewaal, P., Van der Does, F., Heine, R., Van Ballegooie, E., Verduijn, M., & Bouma, M. (2006). Nhg-standaard diabetes mellitus type 2. *Huisarts Wet*, *49*(3), 137–52.

[30] Holleman, F., de Vries, J., & Hoekstra, J. (2006). De standaard 'diabetes mellitus type 2'(tweede herziening) van het nederlandshuisartsen genootschap; reactie vanuit deinterne geneeskunde. *Ned Tijdschr Geneeskd.*, (150), 2235–7.

[31] Beran, D., Yudkin, J. S., & de Courten, M. (2005). Access to care for patients with insulin-requiring diabetes in developing countries. *Diabetes Care*, *28*(9), 2136–2140. https://doi.org/10.2337/diacare.28.9.2136

[32] Cattaneo, M. D., Idrobo, N., & Titiunik, R. (2024, March). *A practical introduction to regression discontinuity designs: Extensions*. Cambridge University Press. https://doi.org/10.1017/9781009441896

[33] Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in Economics. *J. Econ. Lit.*, *48*(2), 281–355. https://doi.org/10.1257/jel.48.2.281

[34] Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust data-driven inference in the regression-discontinuity design. *Stata J.*, *14*(4), 909–946. https://doi.org/10.1177/1536867x1401400413

[35] Cattaneo, M. D., Idrobo, N., & Titiunik, R. (2019, November). *A practical introduction to regression discontinuity designs: Foundations* (R. M. Alverez & N. Beck, Eds.). Cambridge University Press. https://doi.org/10.1017/9781108684606

[36] Kolesár, M., & Rothe, C. (2018). Inference in regression discontinuity designs with a discrete running variable. *Am. Econ. Rev.*, *108*(8), 2277–2304. https://doi.org/10.1257/aer.20160945

[37] Sherwani, S. I., Khan, H. A., Ekhzaimy, A., Masood, A., & Sakharkar, M. K. (2016). Significance of HbA1c test in diagnosis and prognosis of diabetic patients. *Biomark. Insights*, *11*, BMI.S38440. https://doi.org/10.4137/bmi.s38440

[38] Al-Saeed, A. H., Constantino, M. I., Molyneaux, L., D'Souza, M., Limacher-Gisler, F., Luo, C., Wu, T., Twigg, S. M., Yue, D. K., & Wong, J. (2016). An inverse relationship between age of type 2 diabetes onset and complication risk and mortality: The impact of youth-onset type 2 diabetes. *Diabetes Care*, *39*(5), 823–829. https://doi.org/10.2337/dc15-0991

[39] Hillier, T. A., & Pedula, K. L. (2003). Complications in young adults with early-onset type 2 diabetes. *Diabetes Care*, *26*(11), 2999–3005. https://doi.org/10.2337/diacare.26.11.2999

[40] Oksuzyan, A., Dańko, M., Caputo, J., & Jasilionis, V., D.and Shkolnikov. (2019). Is the story about sensitive women and stoical men true? Gender differences in health after adjustment for reporting behavior. *Soc Sci Med*, 41–50. https://doi.org/10.1016/j.socscimed.2019.03.002

[41] Lazarevic, P., & Brandt, M. (2020). Diverging ideas of health? Comparing the basis of health ratings across gender, age, and country. *Soc Sci Med*, (267), 1129–13. https://doi.org/10.1016/j.socscimed.2020.112913

[42] Wright, A. K., Welsh, P., Gill, J. M. R., Kontopantelis, E., Emsley, R., Buchan, I., Ashcroft, D. M., Rutter, M. K., & Sattar, N. (2020). Age-, sex- and ethnicity-related differences in body weight, blood pressure, HbA1c and lipid levels at the

diagnosis of type 2 diabetes relative to people without diabetes. *Diabetologia*, *63*(8), 1542–1553. https://doi.org/10.1007/s00125-020-05169-6

[43] Wändell, P. E., & Carlsson, A. C. (2014). Gender differences and time trends in incidence and prevalence of type 2 diabetes in Sweden—a model explaining the diabetes epidemic worldwide today? *Diabetes Res. Clin. Pract.*, *106*(3), e90–e92. https://doi.org/10.1016/j.diabres.2014.09.013

[44] Groeneveld, J. M., Ballering, A. V., van Boven, K., Akkermans, R. P., Olde Hartman, T. C., & Uijen, A. A. (2020). Sex differences in incidence of respiratory symptoms and management by general practitioners. *Fam. Pract.*, *37*(5), 631–636. https://doi.org/10.1093/fampra/cmaa040

[45] Ballering, A. V., Muijres, D., Uijen, A. A., Rosmalen, J. G., & olde Hartman, T. C. (2021). Sex differences in the trajectories to diagnosis of patients presenting with common somatic symptoms in primary care: An observational cohort study. *J. Psychosom. Res.*, *149*, 110589. https://doi.org/10.1016/j.jpsychores.2021.110589

[46] Ballering, A. V., Rosmalen, J. G. M., & Hartman, T. C. o. (2022). Differences between women and men are present in the rate of diagnosed diseases after a diagnostic intervention is conducted in primary care. *J. Am. Board Fam. Med.*, *35*(1), 73–84. https://doi.org/10.3122/jabfm.2022.01.210289

[47] Tatulashvili, S., Fagherazzi, G., Dow, C., Cohen, R., Fosse, S., & Bihan, H. (2020). Socioeconomic inequalities and type 2 diabetes complications: A systematic review. *Diabetes Metab.*, *46*(2), 89–99. https://doi.org/10.1016/j.diabet.2019.11.001

[48] Robbins, J. M., Vaccarino, V., Zhang, H., & Kasl, S. V. (2005). Socioeconomic status and diagnosed diabetes incidence. *Diabetes Res. Clin. Pract.*, *68*(3), 230–236. https://doi.org/10.1016/j.diabres.2004.09.007

[49] Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *J. Am. Stat. Assoc.*, *103*(484), 1481–1495. https://doi.org/10.1198/016214508000000841

[50] Zhao, Y., Feng, H.-m., Qu, J., Luo, X., Ma, W.-j., & Tian, J.-h. (2018). A systematic review of pharmacoeconomic guidelines. *J. Med. Econ.*, *21*(1), 85–96. https://doi.org/10.1080/13696998.2017.1387118

[51] Lifelines Wiki. (2023, October). General assessments [October 2023].

12 Figures

Figure 1: Timeline of Lifelines' waves

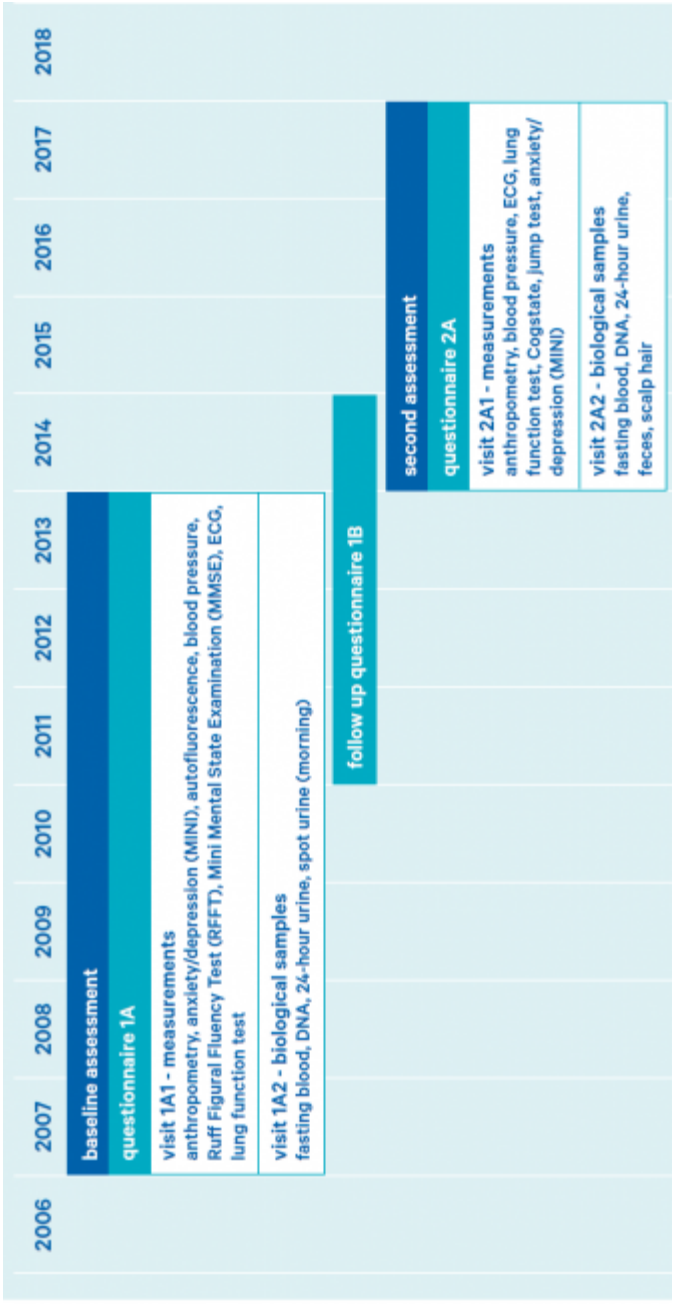
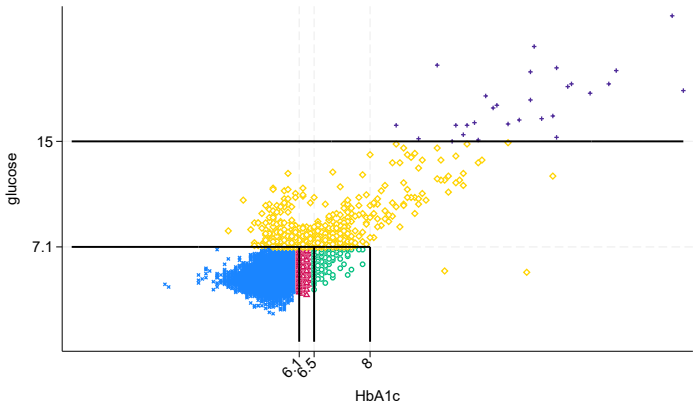
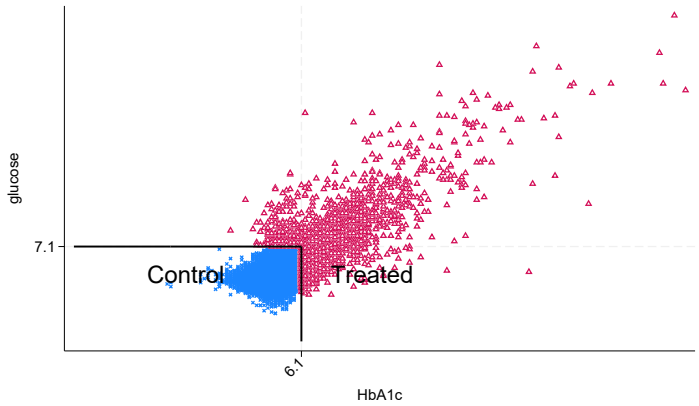


Image taken from the publicly available [51].

Figure 2: Explicative graphs of the various setups



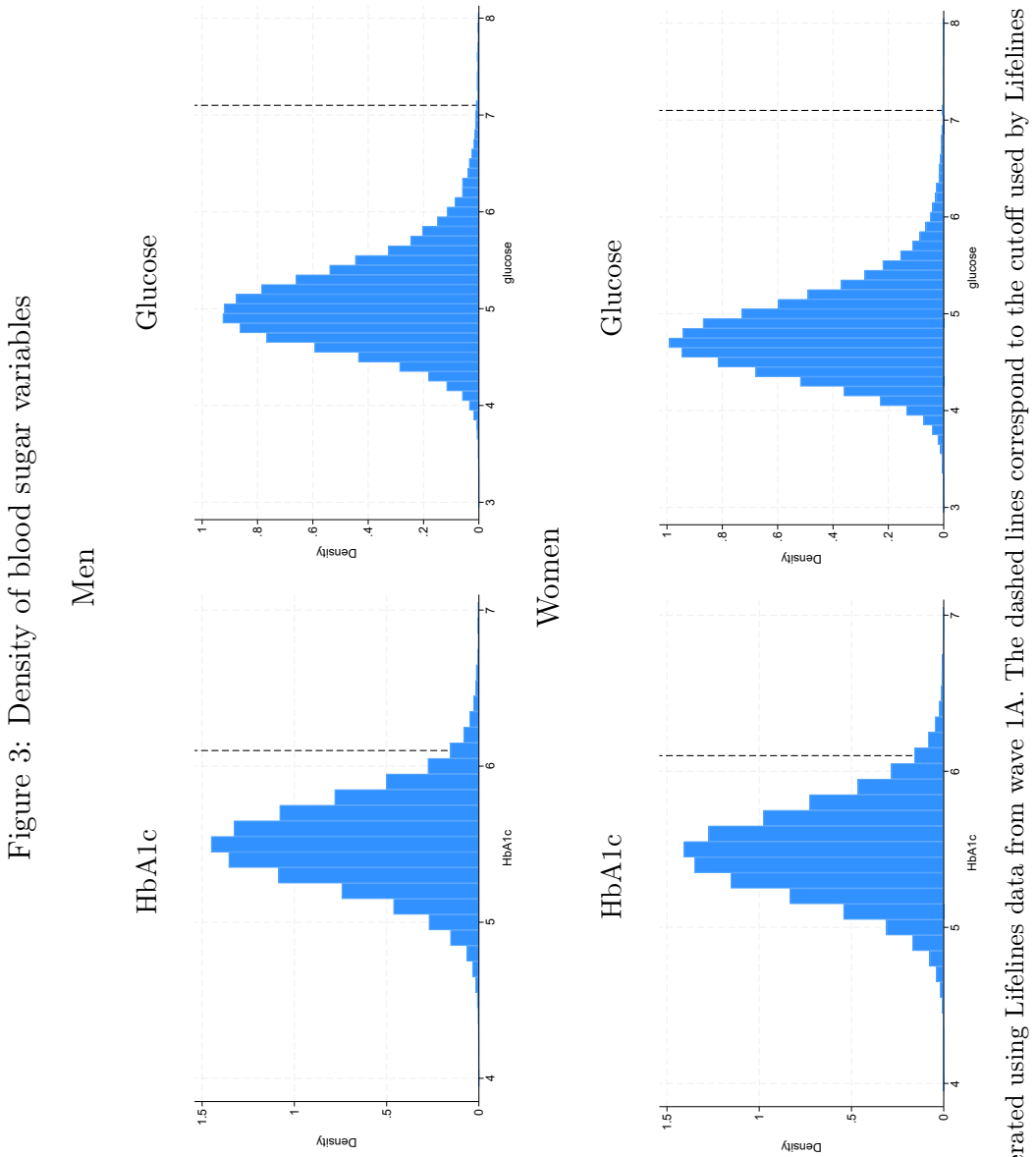
A: Lifelines' setup



B: RDD setup

The graphs were produced using Lifelines data from the first wave. Figure A shows the setup used by Lifelines to issue the different warnings, as explained in Table 1, where the different colours and symbols correspond to different warnings. Figure B shows the setup we employ in the RDD, where we only consider receiving a warning, irrespective of its type, as treatment.

Figure 3: Density of blood sugar variables



The graphs were generated using Lifelines data from wave 1A. The dashed lines correspond to the cutoff used by Lifelines to issue the warnings.

13 Tables

Table 1: Summary of the warnings issued by Lifelines

Warning	1	2	3	4
Value	HbA1c$\in$ $[6.1\%, 6.5\%)$ and glucose$<$ 7.1 mmol/L	HbA1c$\in$ $[6.5\%, 7.9\%]$ and glucose$<$ 7.1 mmol/L	HbA1c$\geq$ 8% or glucose$\geq$ 7.1 mmol/L	glucose$\geq$ 15 mmol/L
Diagnosis according to the ADA	prediabetes*	diabetes	diabetes	diabetes
Message	"the average glucose level in the blood is slightly increased. If Diabetes Mellitus is not known, it is recommended to check this."	"the average glucose level in the blood is slightly increased. Would you like to have this checked again by your doctor?"	"the average glucose level in the blood is increased. Will you have this checked again by your doctor?"	the GP was called by Lifelines' endocrinologist

Warnings issued at baseline by Lifelines depending on the various thresholds. ADA stands for American Diabetes Association.
*The official threshold for the ADA is glucose$\geq$ 7 mmol/L. However, the majority of the individuals in this group would still be considered prediabetic by the ADA.

Table 2: Distribution of the different warnings at baseline by values of the running variable.

	men		women	
	total mean	1.5 sd above mean	total mean	1.5 sd above mean
prediabetes warning 1	0.697	0.822	0.807	0.896
diabetes warning 2	0.085	0.059	0.080	0.058
diabetes warning 3	0.218	0.119	0.112	0.046
Observations	1770	1500	2212	1994

The left column for both men and women contains information on the frequency of those warned by Lifelines computed on the total sample. The right column displays the relative frequency only of those who received a warning and are situated in a range of 1.5 sd of the running variable above the cutoff. We drop those who received warning 4 as they are too fat from the threshold and would not enter our estimation.

Table 3: Falsification test for discontinuities in predetermined covariates around the cutoff

Variable	τ	CER-Optimal Bandwidth	Robust Inference p-value	Effective N Left	Right
Women					
age	-0.054	0.730	0.999	4250	1626
university education	-0.008	1.042	0.702	8339	1767
lower secondary education	0.016	0.894	0.593	8339	1761
born outside NL	-0.001	0.818	0.866	4250	1630
employed	-0.0119	1.039	0.574	8339	1767
weekly hours worked	-0.043	1.129	0.972	8287	1749
disability insurance	-0.0023	1.072	0.771	8367	1770
retired	0.002	1.041	0.440	8339	1767
GP visit	-0.018	0.777	0.350	4250	1629
currently smoking	-0.044	0.927	0.053	8272	1738
alcohol drinker	0.006	0.850	0.784	4228	1618
at least very good health	-0.034	1.010	0.135	8339	1767
BMI	0.289	0.782	0.470	4250	1629
Men					
age	-0.706	0.874	0.116	6230	1297
university education	0.008	0.931	0.854	6301	1306
lower secondary education	-0.011	1.101	0.662	6365	1312
born outside NL	-0.007	0.989	0.421	6301	1307
employed	0.024	0.837	0.174	6230	1295
weekly hours worked	0.689	0.893	0.401	6194	1289
disability insurance	-0.006	0.713	0.579	3136	1186
retired	0.008	0.649	0.136	3136	1184
GP visit	0.026	0.700	0.420	3136	1186
currently smoking	-0.017	0.665	0.729	3092	1163
alcohol drinker	-0.009	0.846	0.649	6183	1283
at least very good health	-0.014	0.827	0.692	3185	1199
BMI	0.055	0.688	0.968	3136	1186

This falsification test looks for discontinuities in predetermined covariates at baseline. We estimate Equation 5 where the outcome is the baseline value of the control variable and we exclude the controls on the right-hand side of the equation. The coefficients are obtained via a multidimensional RDD performed on Lifelines data. We employed CER-optimal bandwidths, as suggested in [35], and a first-order polynomial on both sides of the cutoff.

Table 4: Descriptive statistics at baseline by gender and values of the running variable

	1.5 sd below		1.5 sd above		
	$\bar{x}_b$	sd	$\bar{x}_a$	sd	$\bar{x}_b$-$\bar{x}_a$
Women					
controls					
age	46.311	(7.828)	49.680	(6.942)	-3.370***
highly educated	0.271	(0.445)	0.188	(0.391)	0.083***
lowly educated	0.729	(0.445)	0.812	(0.391)	-0.083***
born outside NL	0.029	(0.167)	0.035	(0.184)	-0.006
outcomes					
employment	0.841	(0.366)	0.776	(0.417)	0.064***
workweek hours	20.965	(13.368)	18.636	(13.992)	2.330***
disabled	0.029	(0.167)	0.044	(0.204)	-0.015***
retired	0.003	(0.055)	0.005	(0.074)	-0.002
gp visit	0.823	(0.382)	0.873	(0.333)	-0.050***
smoker	0.214	(0.410)	0.238	(0.426)	-0.025*
alcohol drinking	0.706	(0.456)	0.671	(0.470)	0.035**
very good health	0.327	(0.469)	0.237	(0.425)	0.090***
BMI	26.325	(4.795)	28.909	(5.923)	-2.584***
Observations	20446		1834		22280
Men					
controls					
age	45.609	(7.716)	48.393	(7.037)	-2.783***
highly educated	0.295	(0.456)	0.249	(0.433)	0.046***
lowly educated	0.705	(0.456)	0.751	(0.433)	-0.046***
born outside NL	0.022	(0.146)	0.037	(0.190)	-0.016***
outcomes					
employment	0.939	(0.240)	0.914	(0.280)	0.024***
workweek hours	38.507	(14.327)	37.731	(15.658)	0.776
disabled	0.022	(0.146)	0.036	(0.186)	-0.014***
retired	0.006	(0.080)	0.009	(0.096)	-0.003
gp visit	0.727	(0.446)	0.772	(0.419)	-0.046***
smoker	0.271	(0.444)	0.295	(0.456)	-0.025*
alcohol drinking	0.895	(0.306)	0.860	(0.348)	0.036***
very good health	0.368	(0.482)	0.273	(0.446)	0.095***
BMI	26.846	(3.646)	28.738	(4.380)	-1.892***
Observations	15457		1388		16845

*** p<0.01, ** p<0.05, * p<0.1

Statistics obtained using the first wave of Lifelines data. We compare individuals in the range of 1.5 sd of the running variable below the threshold to those 1.5 sd above the threshold. $\bar{x}$ corresponds to the average in the considered range below ($\bar{x}_b$) or above ($\bar{x}_a$) the cutoff. The last column reports significance stars from a t-value test performed on the difference between $\bar{x}_b$ and $\bar{x}_a$.

Table 5: Effect of a warning on labour market outcomes by gender and age

	1B (after 1.5 y)			2A (after 4 y)		
	τ	average 0.5 sd before cutoff	Effective N	τ	average 0.5 sd before cutoff	Effective N
Women						
Under 40						
employed	.0758**	.8553	3846	.0862*	.8475	2791
	(2.433)			(1.824)		
weekly hours worked	1.1727	20.8	3706	2.3977	21.3	2765
	(1.124)			(1.609)		
disability insurance	-.0212*	.0395	3846	-.0094	.0339	2791
	(-1.875)			(-.638)		
Over 40						
employed	.0191	.7799	16870	-.0095	.7355	13603
	(1.581)			(-.517)		
weekly hours worked	.1793	18.4	15973	-.1992	17.6	13429
	(.505)			(-.442)		
disability insurance	-.0077	.0564	16870	-.0024	.0533	13603
	(-1.3)			(-.411)		
early retirement	-.0008	.0085	16870	.0163*	.0266	13603
	(-.154)			(1.845)		
Men						
Under 40						
employed	.0431**	.9328	3068	.0304	.9398	2286
	(2.009)			(.896)		
weekly hours worked	2.4274**	38.5	2892	-.007	38.9	2267
	(2.05)			(.001)		
disability insurance	-.0236*	.042	3068	-.0192	.0241	2286
	(-1.685)			(-1.295)		
Over 40						
employed	.0068	.8913	12215	-.026	.8559	9748
	(.831)			(-1.485)		
weekly hours worked	.1861	36.1	11444	-.5713	34.0	9649
	(.49)			(-.711)		
disability insurance	-.0129	.0673	12215	-.024**	.0785	9748
	(-1.505)			(-2.26)		
early retirement	.0059	.009	12215	.0115	.0328	9748
	(1.517)			(1.223)		

*** p<0.01, ** p<0.05, * p<0,1

Results are based on the estimation of Equation 5 using an RD design with a triangular kernel and local linear regressions, applied to data from Lifelines waves 1B and 2A. The bandwidth is set to 1.5 standard deviations of the running variable, calculated by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. We control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects. For reference, we report the mean of the dependent variable for individuals located within 0.5 standard deviations below the cutoff. The number of observations used in each regression corresponds to those falling within the estimation bandwidth, reflecting the effective sample size for the RD estimation.

Table 6: Effect of a warning on women's health behaviours and outcomes by age

	1B (after 1.5 y)			2A (after 4 y)		
	τ	average 0.5 sd before cutoff	Effective N	τ	average 0.5 sd before cutoff	Effective N
Under 40						
health behaviours						
GP visits	-.0152	.8026	3846	.0377	.7712	2791
	(-.288)			(.788)		
smoking	-.0079	.1908	3846	.0401	.1373	2243
	(-.43)			(.785)		
drinking				-.0627	.7379	2261
				(-.961)		
exercising	-.0151	.4013	3846			
	(-.431)					
subjective measure						
very good health	-.1015**	.3224	3846	-.0711	.3814	2791
	(-2.036)			(-1.104)		
health outcomes						
diagnosed	.0044	.0263	3846	-.0074	.0339	2791
	(.263)			(-.168)		
BMI	.1493	27.2	3846	-.3574	27.3	2791
	(.754)			(-.883)		
HbA1c				-.0092	5.7	2791
				(-.167)		
death				-.0011	.	4433
				(-.571)		
Over 40						
health behaviours						
GP visits	.0047	.8124	16870	-.01	.823	13603
	(.209)			(-.728)		
smoking	-.0037	.2332	16870	-.0059	.2	11905
	(-.383)			(-.658)		
drinking				-.0324*	.7426	11976
				(-1.797)		
exercising	.0256	.2386	16870			
	(1.015)					
subjective measure						
very good health	.0267*	.2324	16870	-.0033	.275	13603
	(1.692)			(-.002)		
health outcomes						
diagnosed	.0057	.0062	16870	.0275***	.019	13603
	(1.458)			(2.764)		
BMI	-.0174	26.8	16870	-.0106	27.1	13603
	(-.339)			(-.125)		
HbA1c				.0109	5.9	13603
				(.426)		
death				-.0055	.0115	19039
				(-1.393)		

*** p<0.01, ** p<0.05, * p<0.1

Results are based on the estimation of Equation 5 using an RD design with a triangular kernel and local linear regressions, applied to data from Lifelines waves 1B and 2A. The bandwidth is set to 1.5 standard deviations of the running variable, calculated by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. We control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects. Mortality is defined as dying by the end of 2018. For reference, we report the mean of the dependent variable for individuals located within 0.5 standard deviations below the cutoff. The number of observations used in each regression corresponds to those falling within the estimation bandwidth, reflecting the effective sample size for the RD estimation.

Table 7: Effect of a warning on men's health behaviours and outcomes by age

	1B (after 1.5 y)			2A (after 4 y)		
	τ	average 0.5 sd before cutoff	Effective N	τ	average 0.5 sd before cutoff	Effective N
Under 40						
health behaviours						
GP visits	-.0093	.6218	3068	-.0126	.7229	2286
	(-.185)			(-.279)		
smoking	-.0734*	.395	3068	-.0833*	.3478	1785
	(-1.914)			(-1.8)		
drinking				-.0493	.9286	1811
				(-.976)		
exercising	-.0058	.2521	3068			
	(-.047)					
subjective measure						
very good health	-.0347	.3445	3068	.0028	.3735	2286
	(-.823)			(.006)		
health outcomes						
diagnosed	.0083	.	3068	-.001	.0241	2286
	(.611)			(-.164)		
BMI	-.0957	27.1	3068	.0527	27.2	2286
	(-.424)			(.216)		
HbA1c				-.0183	5.8	2286
				(-.389)		
death				-.005	.0069	3669
				(-.965)		
Over 40						
health behaviours						
GP visits	.041*	.685	12215	.0137	.7404	9748
	(1.727)			(.553)		
smoking	-.0022	.2511	12215	.0003	.2044	8205
	(-.117)			(.234)		
drinking				.0085	.8941	8270
				(.613)		
exercising	.026	.1962	12215			
	(1.093)					
subjective measure						
very good health	-.0234	.2915	12215	-.0255	.3153	9748
	(-.993)			(-1.066)		
health outcomes						
diagnosed	.0055	.0056	12215	.0147	.03	9748
	(.712)			(.752)		
BMI	-.0339	27.3	12215	.0267	27.5	9748
	(-.195)			(.418)		
HbA1c				.0189	5.9	9748
				(.401)		
death				-.0006	.0146	13924
				(-.104)		

*** p<0.01, ** p<0.05, * p<0.1

Results are based on the estimation of Equation 5 using an RD design with a triangular kernel and local linear regressions, applied to data from Lifelines waves 1B and 2A. The bandwidth is set to 1.5 standard deviations of the running variable, calculated by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. We control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects. Mortality is defined as dying by the end of 2018. For reference, we report the mean of the dependent variable for individuals located within 0.5 standard deviations below the cutoff. The number of observations used in each regression corresponds to those falling within the estimation bandwidth, reflecting the effective sample size for the RD estimation.

Table 8: Effect of a warning on women's labour market outcomes by age and educational level.

	under 40						over 40					
	1B (after 1.5 y)			2A (after 4 y)			1B (after 1.5 y)			2A (after 4 y)		
	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N
low education												
employed	.0861**	.8265	2321	.101*	.8219	1637	.0146	.7606	12768	-.007	.7091	10239
	(2.015)			(1.675)			(1.004)			(-.312)		
weekly hours worked	1.6423	18.3	2213	3.7509**	19.0	1618	-.0135	16.8	12042	-.135	15.9	10096
	(1.229)			(2.058)			(.014)			(-.246)		
disability insurance	-.0226	.051	2321	-.0125	.0548	1637	-.0035	.0584	12768	.0012	.0545	10239
	(-1.47)			(-.583)			(-.612)			(.02)		
early retirement							.0026	.0049	12768	.02**	.0242	10239
							(.825)			(2.003)		
high education												
employed	.0483	.9074	1497	.0525	.8889	1135	.0217	.8608	3969	-.01	.8341	3263
	(1.453)			(.627)			(1.068)			(-.363)		
weekly hours worked	.0931	25.0	1465	-.4462	25.0	1128	.5438	24.5	3803	-.3415	23.8	3235
	(.12)			(-.32)			(.552)			(-.371)		
disability insurance	-.0188	.0185	1497	-.0072	.	1135	-.0214*	.0476	3969	-.0185	.0461	3263
	(-1.277)			(-.361)			(-1.649)			(-1.08)		
early retirement							-.0129	.022	3969	.0004	.0369	3263
							(-1.472)			(.198)		

Results are based on the estimation of Equation 5 using an RD design with a triangular kernel and local linear regressions, applied to data from Lifelines waves 1B and 2A. The bandwidth is set to 1.5 standard deviations of the running variable, calculated by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. We control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects. For reference, we report the mean of the dependent variable for individuals located within 0.5 standard deviations below the cutoff. The number of observations used in each regression corresponds to those falling within the estimation bandwidth, reflecting the effective sample size for the RD estimation.

Table 9: Effect of a warning on men's labour market outcomes by age and educational level.

	under 40						over 40					
	1B (after 1.5 y)			2A (after 4 y)			1B (after 1.5 y)			2A (after 4 y)		
	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N
low education												
employed	.0535*	.9091	2064	.0252	.931	1505	-.0016	.8904	8632	-.0455**	.8483	6809
	(1.829)			(.517)			(.019)			(-2.133)		
weekly hours worked	3.009**	37.4	1950	-.6616	39.5	1490	.3802	36.3	8065	-1.0802	33.9	6726
	(1.975)			(-.353)			(.641)			(-1.041)		
disability insurance	-.0312	.0568	2064	-.0283	.0345	1505	-.0198*	.0776	8632	-.0297**	.0918	6809
	(-1.597)			(-1.291)			(-1.81)			(-2.148)		
early retirement							.0052	.0076	8632	.0204**	.0259	6809
							(1.46)			(2.01)		
high education												
employed	.0067	1.	986	.0298	.9583	765	.0407**	.8913	3470	.0116	.8724	2848
	(.563)			(.879)			(2.318)			(.388)		
weekly hours worked	1.1853	41.7	925	.0594	37.6	762	.4917	35.3	3275	.3052	34.1	2832
	(.665)			(.032)			(.678)			(.186)		
disability insurance	.0042	.	986	.0012	.	765	-.0015	.0391	3470	-.0152	.0459	2848
	(1.114)			(.242)			(-.153)			(-1.186)		
early retirement							.01	.013	3470	-.0076	.051	2848
							(.882)			(-.308)		

Results are based on the estimation of Equation 5 using an RD design with a triangular kernel and local linear regressions, applied to data from Lifelines waves 1B and 2A. The bandwidth is set to 1.5 standard deviations of the running variable, calculated by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. We control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects. For reference, we report the mean of the dependent variable for individuals located within 0.5 standard deviations below the cutoff. The number of observations used in each regression corresponds to those falling within the estimation bandwidth, reflecting the effective sample size for the RD estimation.

Table 10: Effect of a warning on women's health behaviours outcomes by age and with low educational.

	under 40						over 40					
	1B (after 1.5 y)			2A (after 4 y)			1B (after 1.5 y)			2A (after 4 y)		
	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N
health behaviours												
GP visit	-.0321	.8367	2321	.086	.8082	1637	-.0032	.818	12768	-.0386**	.8436	10239
	(-.607)			(1.352)			(-.225)			(-2.02)		
smoking	-.0363	.2449	2321	.0178	.1774	1303	-.0076	.2582	12768	-.0081	.2181	9010
	(-1.374)			(.239)			(-.646)			(-.729)		
alcohol				-.1156	.7143	1314				-.0385*	.7244	9064
				(-1.265)						(-1.814)		
exercising	-.0632	.4592	2321				.0299	.2255	12768			
	(-1.057)						(1.059)					
subjective measure												
very good health	-.0803	.2857	2321	-.0588	.3562	1637	.0237	.2226	12768	.0026	.2558	10239
	(-1.383)			(-.779)			(1.313)			(.232)		
health outcomes												
diagnosed	.0027	.0204	2321	-.0139	.0274	1637	.0039	.0069	12768	.0164	.0218	10239
	(.211)			(-.428)			(.953)			(1.528)		
BMI	.3987	27.6	2321	-.5307	27.9	1637	-.0314	27.2	12768	-.0214	27.4	10239
	(1.589)			(-.937)			(-.408)			(-.187)		
HbA1c				.0218	5.6616	1637				.0105	5.8657	10239
				(.321)						(.357)		
death				-.0025	.	2703				-.0032	.0104	14506
				(-.938)						(-.752)		

Results are based on the estimation of Equation 5 using an RD design with a triangular kernel and local linear regressions, applied to data from Lifelines waves 1B and 2A. The bandwidth is set to 1.5 standard deviations of the running variable, calculated by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. We control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects. Mortality is defined as dying by the end of 2018. For reference, we report the mean of the dependent variable for individuals located within 0.5 standard deviations below the cutoff. The number of observations used in each regression corresponds to those falling within the estimation bandwidth, reflecting the effective sample size for the RD estimation.

Table 11: Effect of a warning on women's health behaviours outcomes by age and with high educational level.

	under 40						over 40					
	1B (after 1.5 y)			2A (after 4 y)			1B (after 1.5 y)			2A (after 4 y)		
	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N
health behaviours												
GP visit	.0044	.7407	1497	-.107	.7111	1135	.0236	.8022	3969	.0915**	.7512	3263
	(.209)			(-.659)			(.527)			(2.116)		
smoking	.0458	.0926	1497	.0795	.075	922	.0055	.1282	3969	-.0077	.1176	2809
	(.76)			(1.243)			(.28)			(-.402)		
alcohol				.0676	.775	929				-.0258	.8182	2825
				(.689)						(-.797)		
exercising	.0722	.2963	1497				.0061	.2894	3969			
	(.664)						(.159)					
subjective measure												
very good health	-.1558	.3889	1497	-.1178	.4222	1135	.0503	.2674	3969	-.0172	.3456	3263
	(-1.621)			(-1.08)			(1.418)			(-.162)		
health outcomes												
diagnosed	.0094	.037	1497	.0156	.0444	1135	.0115	.0037	3969	.0732***	.0092	3263
	(.07)			(.233)			(1.268)			(2.816)		
BMI	-.7005**	26.5	1497	-.0482	26.3	1135	.0488	25.6	3969	.0738	25.9	3263
	(-2.08)			(.079)			(.23)			(.339)		
HbA1c				-.074	5.6533	1135				.0125	5.8074	3263
				(-1.155)						(.294)		
death				.0017	.	1691				-.0148*	.0165	4376
				(1.517)						(-1.83)		

Results are based on the estimation of Equation 5 using an RD design with a triangular kernel and local linear regressions, applied to data from Lifelines waves 1B and 2A. The bandwidth is set to 1.5 standard deviations of the running variable, calculated by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. We control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects. Mortality is defined as dying by the end of 2018. For reference, we report the mean of the dependent variable for individuals located within 0.5 standard deviations below the cutoff. The number of observations used in each regression corresponds to those falling within the estimation bandwidth, reflecting the effective sample size for the RD estimation.

Table 12: Effect of a warning on men's health behaviours outcomes by age and with low educational.

	under 40						over 40					
	1B (after 1.5 y)			2A (after 4 y)			1B (after 1.5 y)			2A (after 4 y)		
	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N
health behaviours												
GP visit	-.0116	.6591	2064	.0349	.6724	1505	.0326	.6834	8632	.0259	.7385	6809
	(-.27)			(.482)			(1.27)			(.959)		
smoking	-.0676	.4773	2064	-.1422**	.4167	1181	-.0031	.2846	8632	-.0005	.2512	5749
	(-1.469)			(-2.362)			(-.16)			(.191)		
alcohol				-.0463	.898	1200				.0151	.8741	5801
				(-.676)						(.873)		
exercising	.0164	.2159	2064				.0413	.1781	8632			
	(.337)						(1.599)					
subjective measure												
very good health	.017	.2614	2064	.0077	.2759	1505	-.0044	.2694	8632	-.0128	.2814	6809
	(.241)			(.169)			(-.094)			(-.543)		
health outcomes												
diagnosed	-.0002	.	2064	-.0204	.0345	1505	.0013	.0076	8632	.015	.0279	6809
	(.224)			(-.743)			(-.081)			(.715)		
BMI	-.2348	27.7	2064	-.1222	27.8	1505	-.0231	27.5	8632	-.0307	27.8	6809
	(-1.271)			(-.309)			(.052)			(-.142)		
HbA1c				-.0145	5.8	1505				.0227	5.9	6809
				(-.37)						(.42)		
death				-.0005	.	2475				-.0036	.0158	9930
				(-.08)						(-.531)		

Results are based on the estimation of Equation 5 using an RD design with a triangular kernel and local linear regressions, applied to data from Lifelines waves 1B and 2A. The bandwidth is set to 1.5 standard deviations of the running variable, calculated by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. We control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects. Mortality is defined as dying by the end of 2018. For reference, we report the mean of the dependent variable for individuals located within 0.5 standard deviations below the cutoff. The number of observations used in each regression corresponds to those falling within the estimation bandwidth, reflecting the effective sample size for the RD estimation.

Table 13: Effect of a warning on men's health behaviours outcomes by age and with high educational level.

	under 40						over 40					
	1B (after 1.5 y)			2A (after 4 y)			1B (after 1.5 y)			2A (after 4 y)		
	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N
health behaviours												
GP visit	.0097	.5	986	-.1567	.875	765	.0642	.687	3470	.0023	.7449	2848
	(.185)			(-1.622)			(1.263)			(-.092)		
smoking	-.0747	.1667	986	.0519	.2	592	.0053	.1609	3470	.0044	.0864	2375
	(-1.062)			(.865)			(.229)			(.194)		
alcohol				-.0657	1.	598				-.0086	.9444	2388
				(-.884)						(-.391)		
exercising	-.031	.3333	986				-.0095	.2478	3470			
	(-.296)						(-.346)					
subjective measure												
very good health	-.1012	.5667	986	.0059	.625	765	-.0764*	.3522	3470	-.0591	.4031	2848
	(-1.241)			(-.111)			(-1.67)			(-1.079)		
health outcomes												
diagnosed	.027	.	986	.0121	.	765	.0144	.	3470	.0142	.0357	2848
	(.616)			(.105)			(1.327)			(.329)		
BMI	.3975	25.6	986	.041	25.9	765	-.0451	26.6	3470	.1792	26.6	2848
	(.79)			(.273)			(-.417)			(1.252)		
HbA1c				-.0336	5.8	765				.0138	5.8	2848
				(-.251)						(.226)		
death				-.0176	.0278	1167				.0046	.0078	3863
				(-1.151)						(.465)		

Results are based on the estimation of Equation 5 using an RD design with a triangular kernel and local linear regressions, applied to data from Lifelines waves 1B and 2A. The bandwidth is set to 1.5 standard deviations of the running variable, calculated by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. We control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects. Mortality is defined as dying by the end of 2018. For reference, we report the mean of the dependent variable for individuals located within 0.5 standard deviations below the cutoff. The number of observations used in each regression corresponds to those falling within the estimation bandwidth, reflecting the effective sample size for the RD estimation.

A Diagnosed diabetes and medication data

To identify individuals diagnosed with diabetes at baseline, we used self-reported diagnosis data, 2,292 individuals, and blood-sugar-lowering medication use data (39 additional observations). From how the questionnaire was structured, diagnosed diabetes is considered an absorbing state.

Data on medication use is only available for wave 1A. Until July 2012, individuals self-reported medication by questionnaire (74,958 participants). After July 2012, medication information was obtained only via medication wrappers that participants could bring to one of the two in-person visits to the chosen location for the physical examinations (42,642 participants). In Table A1, the share of individuals reporting medication use but not a diagnosis increased but remained small after July 2012. If the reporting behaviour remains consistent in wave 2A, we expect most individuals diagnosed between waves to report their new status. Nonetheless, we are aware that we might miss a fraction of those newly diagnosed because of the lack of medication data.

Table A1: Percentages of diagnosed individuals by self-report and/or medication usage

time	Reported diagnosis and medication	Reported diagnosis but no medication	Reported medication but no diagnosis	N
before July 2012	65.98	32.72	1.30	1464
after July 2012	65.51	32.18	2.31	867

Table obtained using Lifelines data for wave 1A. We only employed data on those who either self-reported a previous diabetes diagnosis or reported use of diabetes medication.

B Sensitivity analysis

In this Section, we conduct sensitivity analyses of the main results presented in Tables 5, 6, and 7. In Tables B1-B3, we employ a uniform kernel instead of a triangular one. In Tables B4-B6, we use observations in a bandwidth of 1 sd of the running variable. In Tables B7-B9, as discussed in Section 3, we analyse the effect of warning 1 (for prediabetes) on labour and health outcomes. The running variable is standardised HbA1c and we only include individuals with glucose$<$ 7.1 mmol/L and HbA1c$<$ 6.5% in the estimation. These constraints ensure that individuals in the sample either received no warning or received only warning 1. We find similar results to those found in the general sample. Lastly, in Tables B10-B12, we use a second-order polynomial in the estimation.

Table B1: Effect of a warning on labour market outcomes - uniform kernel.

	1B (after 1.5 y)			2A (after 4 y)		
	τ	average 0.5 sd before cutoff in levels	Effective N	τ	average 0.5 sd before cutoff in levels	Effective N
Women						
Under 40						
employed	.0636**	.8553	3846	.0609	.8475	2791
	(2.408)			(1.534)		
weekly hours worked	1.0132	20.7635	3706	1.9165	21.3136	2765
	(1.069)			(1.327)		
disability insurance	-.0255**	.0395	3846	-.0181	.0339	2791
	(-2.401)			(-1.232)		
Over 40						
employed	.0146	.7799	16870	-.0004	.7355	13603
	(1.243)			(.092)		
weekly hours worked	.2366	18.4122	15973	.0764	17.5522	13429
	(.802)			(.255)		
disability insurance	-.0075	.0564	16870	-.004	.0533	13603
	(-1.261)			(-.751)		
early retirement	-.0008	.0085	16870	.0134*	.0266	13603
	(-.08)			(1.674)		
Men						
Under 40						
employed	.04*	.9328	3068	.0313	.9398	2286
	(1.932)			(.85)		
weekly hours worked	2.3156**	38.531	2892	.4341	38.9157	2267
	(2.048)			(.175)		
disability insurance	-.0072	.042	3068	-.0201	.0241	2286
	(-.611)			(-1.481)		
Over 40						
employed	.0024	.8913	12215	-.0191	.8559	9748
	(.362)			(-1.136)		
weekly hours worked	.0112	36.0869	11444	-.5126	33.9799	9649
	(.072)			(-.732)		
disability insurance	-.0113	.0673	12215	-.0194**	.0785	9748
	(-1.431)			(-2.045)		
early retirement	.0051	.009	12215	.0117	.0328	9748
	(1.404)			(1.251)		

*** p<0.01, ** p<0.05, * p<0.1

Results of the estimation of Equation 5. Estimates are generated via a regression discontinuity design, with a uniform kernel and a polynomial of order 1 on both sides of the cutoff using data from waves 1B and 2A from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects.

Table B2: Effect of a warning on women's health behaviours and outcomes - uniform kernel.

	1B (after 1.5 y)			2A (after 4 y)		
	τ	average 0.5 sd before cutoff in levels	Effective N	τ	average 0.5 sd before cutoff in levels	Effective N
Under 40						
GP visits	-.0085	.8026	3846	.0101	.7712	2791
	(-.199)			(.661)		
smoking	-.0157	.1908	3846	.0423	.1373	2243
	(-.482)			(.712)		
drinking				-.0373	.7379	2261
				(-.878)		
exercising	.0045	.4013	3846			
	(-.033)					
very good health	-.0906*	.3224	3846	-.0357	.3814	2791
	(-1.906)			(-.86)		
diagnosed	.0066	.0263	3846	-.0103	.0339	2791
	(.205)			(.019)		
BMI	.0313	27.2304	3846	-.3988	27.2774	2791
	(.238)			(-1.395)		
HbA1c				-.0543	5.6585	2791
				(-.932)		
death				-.0004	.	4433
				(-.098)		
Over 40						
GP visits	.0049	.8124	16870	-.0092	.823	13603
	(.013)			(-.699)		
smoking	-.0065	.2332	16870	-.0104	.2	11905
	(-1.03)			(-1.261)		
drinking				-.0278*	.7426	11976
				(-1.765)		
exercising	.0194	.2386	16870			
	(.574)					
very good health	.0222*	.2324	16870	.0085	.275	13603
	(1.747)			(.942)		
diagnosed	.0016	.0062	16870	.0176**	.019	13603
	(1.008)			(2.139)		
BMI	.0016	26.8297	16870	-.0307	27.0939	13603
	(-.002)			(-.449)		
HbA1c				.0158	5.8537	13603
				(.75)		
death				-.004	.0115	19039
				(-1.187)		

*** p<0.01, ** p<0.05, * p<0.1

Results of the estimation of Equation 5. Estimates are generated via a regression discontinuity design, with a uniform kernel and a polynomial of order 1 on both sides of the cutoff using data from waves 1B and 2A from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects. Mortality is defined as dying by the end of 2018.

Table B3: Effect of a warning on men's health behaviours and outcomes - uniform kernel.

	1B (after 1.5 y)			2A (after 4 y)		
	τ	average 0.5 sd before cutoff in levels	Effective N	τ	average 0.5 sd before cutoff in levels	Effective N
Under 40						
GP visits	-.0314	.6218	3068	.0064	.7229	2286
	(-.637)			(.042)		
smoking	-.066**	.395	3068	-.0783*	.3478	1785
	(-1.98)			(-1.839)		
drinking				-.0746	.9286	1811
				(-1.604)		
exercising	.0113	.2521	3068			
	(.155)					
very good health	-.047	.3445	3068	-.0372	.3735	2286
	(-1.131)			(-.803)		
diagnosed	.0143	.	3068	.0123	.0241	2286
	(1.253)			(.374)		
BMI	-.1448	27.1204	3068	-.0274	27.2477	2286
	(-.634)			(-.09)		
HbA1c				-.0249	5.8	2286
				(-.592)		
death				.0021	.0069	3669
				(.37)		
Over 40						
GP visits	.0338*	.685	12215	.0125	.7404	9748
	(1.709)		(.59)	(.539)		
smoking	-.0063	.2511	12215	-.002	.2044	8205
	(-.560)		(.994)	(-.008)		
drinking				.0066	.8941	8270
				(.398)		
exercising	.0209	.1962	12215			
	(.996)					
very good health	-.0226	.2915	12215	-.0261	.3153	9748
	(-.741)			(-1.203)		
diagnosed	.0044	.0056	12215	.0242	.03	9748
	(.683)			(1.41)		
BMI	-.009	27.2811	12215	.0212	27.49	9748
	(.025)			(.584)		
HbA1c				.0413	5.8749	9748
				(1.081)		
death				-.005	.0146	13924
				(-.988)		

*** p<0.01, ** p<0.05, * p<0.1

Results of the estimation of Equation 5. Estimates are generated via a regression discontinuity design, with a uniform kernel and a polynomial of order 1 on both sides of the cutoff using data from waves 1B and 2A from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects. Mortality is defined as dying by the end of 2018.

Table B4: Effect of a warning on labour market outcomes - smaller bandwidth.

	1B (after 1.5 y)			2A (after 4 y)		
	τ	average 0.5 sd before cutoff in levels	Effective N	τ	average 0.5 sd before cutoff in levels	Effective N
Women						
Under 40						
employed	.069**	.8553	1276	.1082**	.8475	938
	(1.999)			(2.006)		
weekly hours worked	.6796	20.7635	1230	2.0926	21.3136	930
	(.609)			(1.274)		
disability insurance	-.0269*	.0395	1276	-.0087	.0339	938
	(-1.677)			(-.581)		
Over 40						
employed	.0257*	.7799	7880	-.0147	.7355	6397
	(1.798)			(-.719)		
weekly hours worked	.4198	18.4122	7438	-.3194	17.5522	6309
	(1.017)			(-.594)		
disability insurance	-.0084	.0564	7880	-.0057	.0533	6397
	(-1.115)			(-.697)		
early retirement	-.0031	.0085	7880	.0169	.0266	6397
	(-.672)			(1.64)		
Men						
Under 40						
employed	.032	.9328	1028	.0187	.9398	758
	(1.076)			(.445)		
weekly hours worked	2.6612	38.531	959	-.15	38.9157	751
	(1.55)			(-.071)		
disability insurance	-.042*	.042	1028	-.0027	.0241	758
	(-1.789)			(-.143)		
Over 40						
employed	.0103	.8913	5691	-.0287	.8559	4511
	(.849)			(-1.318)		
weekly hours worked	.375	36.0869	5347	-.3458	33.9799	4462
	(.597)			(-.354)		
disability insurance	-.0155	.0673	5691	-.0349**	.0785	4511
	(-1.376)			(-2.425)		
early retirement	.0073*	.009	5691	.0097	.0328	4511
	(1.648)			(.837)		

*** p<0.01, ** p<0.05, * p<0.1

Results of the estimation of Equation 5. Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from waves 1B and 2A from Lifelines. The estimation bandwidth is 1 standard deviation computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects.

Table B5: Effect of a warning on women's health behaviours and outcomes - smaller bandwidth.

	1B (after 1.5 y)			2A (after 4 y)		
	τ	average 0.5 sd before cutoff in levels	Effective N	τ	average 0.5 sd before cutoff in levels	Effective N
Under 40						
GP visits	-.0272	.8026	1276	.0119	.7712	938
	(-.443)			(.24)		
smoking	.0006	.1908	1276	.053	.1373	775
	(-.048)			(1.015)		
drinking				-.0702	.7379	781
				(-.906)		
exercising	-.0366	.4013	1276			
	(-.559)					
very good health	-.1116*	.3224	1276	-.0796	.3814	938
	(-1.868)			(-1.035)		
diagnosed	-.0097	.0263	1276	-.0243	.0339	938
	(-.402)			(-.721)		
BMI	.2631	27.2304	1276	-.2661	27.2774	938
	(1.053)			(-.549)		
HbA1c				-.0015	5.6585	938
				(-.023)		
death				.	.	1463
				(-.081)		
Over 40						
GP visits	.0016	.8124	7880	-.019	.823	6397
	(.045)			(-.942)		
smoking	-.0005	.2332	7880	-.0112	.2	5654
	(-.03)			(-.844)		
drinking				-.0356	.7426	5691
				(-1.61)		
exercising	.0284	.2386	7880			
	(1.065)					
very good health	.0398**	.2324	7880	-.0088	.275	6397
	(1.994)			(-.288)		
diagnosed	.0052	.0062	7880	.0273**	.019	6397
	(1.033)			(2.407)		
BMI	-.0661	26.8297	7880	.0298	27.0939	6397
	(-.867)			(.206)		
HbA1c				.0108	5.8537	6397
				(.429)		
death				-.0026	.0115	8902
				(-.578)		

*** p<0.01, ** p<0.05, * p<0.1

Results of the estimation of Equation 5. Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from waves 1B and 2A from Lifelines. The estimation bandwidth is 1 standard deviation computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects. Mortality is defined as dying by the end of 2018.

Table B6: Effect of a warning on men's health behaviours and outcomes - smaller bandwidth.

	1B (after 1.5 y)			2A (after 4 y)		
	τ	average 0.5 sd before cutoff in levels	Effective N	τ	average 0.5 sd before cutoff in levels	Effective N
Under 40						
GP visits	.0644	.6218	1028	-.045	.7229	758
	(.818)			(-.553)		
smoking	-.0898*	.395	1028	-.1143*	.3478	604
	(-1.891)			(-1.737)		
drinking				-.0351	.9286	618
				(-.645)		
exercising	-.0313	.2521	1028			
	(-.394)					
very good health	-.0148	.3445	1028	.0072	.3735	758
	(-.285)			(.07)		
diagnosed	.0146	.	1028	-.0053	.0241	758
	(1.019)			(-.22)		
BMI	-.0498	27.1204	1028	.3331	27.2477	758
	(-.163)			(.862)		
HbA1c				-.0012	5.8	758
				(-.054)		
death				-.0138	.0069	1238
				(-1.436)		
Over 40						
GP visits	.0481*	.685	5691	.0066	.7404	4511
	(1.692)			(.215)		
smoking	.0029	.2511	5691	.008	.2044	3830
	(.209)			(.456)		
drinking				.0082	.8941	3858
				(.448)		
exercising	.0241	.1962	5691			
	(.871)					
very good health	-.0278	.2915	5691	-.0357	.3153	4511
	(-1.052)			(-1.22)		
diagnosed	.0039	.0056	5691	.0001	.03	4511
	(.51)			(-.171)		
BMI	-.0399	27.2811	5691	.0587	27.49	4511
	(-.304)			(.551)		
HbA1c				-.0024	5.8749	4511
				(-.173)		
death				.0014	.0146	6498
				(.195)		

*** p<0.01, ** p<0.05, * p<0.1

Results of the estimation of Equation 5. Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from waves 1B and 2A from Lifelines. The estimation bandwidth is 1 standard deviation computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects. Mortality is defined as dying by the end of 2018.

Table B7: Effect of a warning on labour market outcomes - only HbA1c.

	1B (after 1.5 y)			2A (after 4 y)		
	τ	average 0.5 sd before cutoff in levels	Effective N	τ	average 0.5 sd before cutoff in levels	Effective N
Women						
Under 40						
employed	.0758*	.8553	2276	.0798*	.8547	1659
	(1.714)			(1.662)		
weekly hours worked	1.1644	20.8	2199	2.1426	21.5	1641
	(.51)			(.687)		
disability insurance	-.0211	.0395	2276	-.0083	.0342	1659
	(-1.163)			(.017)		
Over 40						
employed	.0216	.7797	11839	-.0101	.7354	9578
	(1.509)			(-.799)		
weekly hours worked	.2552	18.4	11219	-.2083	17.5	9456
	(1.326)			(-.447)		
disability insurance	-.0074	.0574	11839	-.0032	.0537	9578
	(-.861)			(-.985)		
early retirement	-.0013	.0085	11839	.0153	.0268	9578
	(-.816)			(1.272)		
Men						
Under 40						
employed	.0444	.9364	1810	.0281	.9367	1366
	(.272)			(.01)		
weekly hours worked	2.3259	38.7	1705	-.5226	39.1	1355
	(.84)			(.153)		
disability insurance	-.0323	.0455	1810	-.0198	.0253	1366
	(-1.533)			(.546)		
Over 40						
employed	.0072	.8932	8439	-.0275	.8612	6722
	(.608)			(-1.102)		
weekly hours worked	.3106	36.0	7915	-.4814	34.1	6647
	(.579)			(-.112)		
disability insurance	-.0171	.0688	8439	-.029**	.0814	6722
	(-1.014)			(-2.284)		
early retirement	.0058	.0095	8439	.0135	.0287	6722
	(1.31)			(1.174)		

*** p<0.01, ** p<0.05, * p<0.1

Results of the estimation of Equation 5. Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from waves 1B and 2A from Lifelines. The estimation bandwidth is 1.5 standard deviations of HbA1c computed by gender at baseline. We excluded individuals with baseline glucose levels larger or equal to 7.1 mmol/L. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects.

Table B8: Effect of a warning on women's health behaviours and outcomes - only HbA1c.

	1B (after 1.5 y)			2A (after 4 y)		
	τ	average 0.5 sd before cutoff in levels	Effective N	τ	average 0.5 sd before cutoff in levels	Effective N
Under 40						
GP visits	-.0125	.8026	2276	.0317	.7692	1659
	(-.245)			(.051)		
smoking	-.0102	.1908	2276	.0418	.1386	1346
	(.054)			(.608)		
drinking				-.0468	.7353	1361
				(-.827)		
exercising	-.0108	.4013	2276			
	(-1.08)					
very good health	-.0995*	.3224	2276	-.0718	.3846	1659
	(-1.95)			(-.926)		
diagnosed	.0015	.0263	2276	-.0088	.0342	1659
	(-.552)			(-.628)		
BMI	.1729	27.2304	2276	-.3848	27.223	1659
	(1.34)			(-.293)		
HbA1c				-.0161	5.6564	1659
				(.119)		
death				-.0009	.	2614
				(.112)		
Over 40						
GP visits	.0084	.8115	11839	-.0112	.8236	9578
	(.372)			(-1.105)		
smoking	-.0045	.2335	11839	-.0076	.2007	8440
	(.16)			(-1.646)		
drinking				-.0372	.7426	8496
				(-1.503)		
exercising	.0298	.2397	11839			
	(1.008)					
very good health	.025*	.2335	11839	-.012	.2771	9578
	(1.818)			(-.622)		
diagnosed	.0042	.0062	11839	.0245	.0182	9578
	(.271)			(1.45)		
BMI	-.0287	26.8103	11839	-.0087	27.0846	9578
	(-1.158)			(.751)		
HbA1c				.0059	5.8561	9578
				(.256)		
death				-.0062	.0116	13322
				(.28)		

*** p<0.01, ** p<0.05, * p<0.1

Results of the estimation of Equation 5. Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from waves 1B and 2A from Lifelines. The estimation bandwidth is 1.5 standard deviations of HbA1c computed by gender at baseline. We excluded individuals with baseline glucose levels larger or equal to 7.1 mmol/L. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects. Mortality is defined as dying by the end of 2018.

Table B9: Effect of a warning on men's health behaviours and outcomes - only HbA1c.

	1B (after 1.5 y)			2A (after 4 y)		
	τ	average 0.5 sd before cutoff in levels	Effective N	τ	average 0.5 sd before cutoff in levels	Effective N
Under 40						
GP visits	.0234	.6091	1810	.0222	.6962	1366
	(1.438)			(.013)		
smoking	-.0889*	.4091	1810	-.1051	.3846	1071
	(-1.86)			(-1.553)		
drinking				-.0445	.9242	1089
				(-.112)		
exercising	-.0171	.2545	1810			
	(-.396)					
very good health	-.0262	.3273	1810	.016	.3797	1366
	(.319)			(-.341)		
diagnosed	.0106	.	1810	.0059	.0253	1366
	(1.543)			(.164)		
BMI	-.0796	27.184	1810	.1875	27.0625	1366
	(-.26)			(1.219)		
HbA1c				.0194	5.8152	1366
				(.212)		
death				-.0088	.0074	2172
				(-1.487)		
Over 40						
GP visits	.0435	.6856	8439	.0159	.73	6722
	(1.128)			(.518)		
smoking	-.0006	.2515	8439	.0043	.2103	5686
	(.277)			(.64)		
drinking				.011	.8952	5722
				(.01)		
exercising	.0307	.1981	8439			
	(.664)					
very good health	-.028	.2954	8439	-.0304	.3213	6722
	(-.98)			(-1.506)		
diagnosed	.0045	.0047	8439	.0114	.0226	6722
	(.504)			(.113)		
BMI	-.065	27.2147	8439	.0129	27.4348	6722
	(-.39)			(.428)		
HbA1c				.016	5.8854	6722
				(-.14)		
death				.0009	.0155	9615
				(-.308)		

*** p<0.01, ** p<0.05, * p<0.1

Results of the estimation of Equation 5. Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from waves 1B and 2A from Lifelines. The estimation bandwidth is 1.5 standard deviations of HbA1c computed by gender at baseline. We excluded individuals with baseline glucose levels larger or equal to 7.1 mmol/L. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects. Mortality is defined as dying by the end of 2018.

Table B10: Effect of a warning on labour market outcomes - second-order polynomial.

	1B (after 1.5 y)			2A (after 4 y)		
	τ	average 0.5 sd before cutoff in levels	Effective N	τ	average 0.5 sd before cutoff in levels	Effective N
Women						
Under 40						
employed	.0723	.8553	3846	.1274*	.8475	2791
	(1.603)			(1.856)		
weekly hours worked	.6407	20.7635	3706	2.4953	21.3136	2765
	(.465)			(1.258)		
disability insurance	-.0298	.0395	3846	-.0052	.0339	2791
	(-1.226)			(-.107)		
Over 40						
employed	.032*	.7799	16870	-.0216	.7355	13603
	(1.73)			(-.878)		
weekly hours worked	.703	18.4122	15973	-.4477	17.5522	13429
	(1.299)			(-.663)		
disability insurance	-.0089	.0564	16870	-.0084	.0533	13603
	(-.881)			(-.711)		
early retirement	-.0054	.0085	16870	.018	.0266	13603
	(-1.002)			(1.323)		
Men						
Under 40						
employed	.0464	.9328	3068	.021	.9398	2286
	(1.208)			(.527)		
weekly hours worked	3.4399	38.531	2892	-.1457	38.9157	2267
	(1.56)			(.05)		
disability insurance	-.0501	.042	3068	.0061	.0241	2286
	(-1.627)			(.203)		
Over 40						
employed	.0118	.8913	12215	-.0323	.8559	9748
	(.693)			(-1.262)		
weekly hours worked	.3625	36.0869	11444	-.4426	33.9799	9649
	(.433)			(-.401)		
disability insurance	-.0188	.0673	12215	-.0448**	.0785	9748
	(-1.254)			(-2.402)		
early retirement	.0078	.009	12215	.0089	.0328	9748
	(1.351)			(.599)		

*** p<0.01, ** p<0.05, * p<0.1

Results of the estimation of Equation 5. Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 2 on both sides of the cutoff using data from waves 1B and 2A from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects.

Table B11: Effect of a warning on women's health behaviours and outcomes - second-order polynomial.

	1B (after 1.5 y)			2A (after 4 y)		
	τ	average 0.5 sd before cutoff in levels	Effective N	τ	average 0.5 sd before cutoff in levels	Effective N
Under 40						
GP visits	-.0158	.8026	3846	.0168	.7712	2791
	(-.221)			(.146)		
smoking	.0045	.1908	3846	.0536	.1373	2243
	(.036)			(.821)		
drinking				-.0886	.7379	2261
				(-.935)		
exercising	-.0847	.4013	3846			
	(-.902)					
very good health	-.1392*	.3224	3846	-.0979	.3814	2791
	(-1.8)			(-.933)		
diagnosed	-.02	.0263	3846	-.0275	.0339	2791
	(-.695)			(-.661)		
BMI	.4499	27.2304	3846	-.1315	27.2774	2791
	(1.384)			(-.132)		
HbA1c				.0196	5.6585	2791
				(.305)		
death				.0006	.	4433
				(.44)		
Over 40						
GP visits	-.0017	.8124	16870	-.0311	.823	13603
	(-.073)			(-1.121)		
smoking	.0035	.2332	16870	-.0195	.2	11905
	(.343)			(-1)		
drinking				-.0412	.7426	11976
				(-1.344)		
exercising	.0329	.2386	16870			
	(1.059)					
very good health	.0564**	.2324	16870	-.0145	.275	13603
	(1.995)			(-.592)		
diagnosed	.0053	.0062	16870	.0272**	.019	13603
	(.798)			(1.995)		
BMI	-.1184	26.8297	16870	.0649	27.0939	13603
	(-1.149)			(.449)		
HbA1c				.0089	5.8537	13603
				(.236)		
death				.0004	.0115	19039
				(.07)		

*** p<0.01, ** p<0.05, * p<0.1

Results of the estimation of Equation 5. Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 2 on both sides of the cutoff using data from waves 1B and 2A from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects. Mortality is defined as dying by the end of 2018.

Table B12: Effect of a warning on men's health behaviours and outcomes - second-order polynomial.

	1B (after 1.5 y)			2A (after 4 y)		
	τ	average 0.5 sd before cutoff in levels	Effective N	τ	average 0.5 sd before cutoff in levels	Effective N
Under 40						
GP visits	.0862	.6218	3068	-.0835	.7229	2286
	(.998)			(-.843)		
smoking	-.0936	.395	3068	-.1063	.3478	1785
	(-1.532)			(-1.189)		
drinking				-.0148	.9286	1811
				(-.216)		
exercising	-.0254	.2521	3068			
	(-.258)					
very good health	-.0239	.3445	3068	.0464	.3735	2286
	(-.329)			(.558)		
diagnosed	.0175	.	3068	-.0043	.0241	2286
	(1.164)			(-.145)		
BMI	.0014	27.1204	3068	.4808	27.2477	2286
	(.015)			(1.167)		
HbA1c				.0033	5.8	2286
				(.012)		
death				-.0156	.0069	3669
				(-1.203)		
Over 40						
GP visits	.0542	.685	12215	.0019	.7404	9748
	(1.415)			(.116)		
smoking	.0058	.2511	12215	.0151	.2044	8205
	(.371)			(.629)		
drinking				.0099	.8941	8270
				(.423)		
exercising	.0226	.1962	12215			
	(.587)					
very good health	-.0329	.2915	12215	-.0357	.3153	9748
	(-1.162)			(-.987)		
diagnosed	.0054	.0056	12215	-.008	.03	9748
	(.711)			(-.565)		
BMI	-.0496	27.2811	12215	.0594	27.49	9748
	(-.258)			(.438)		
HbA1c				-.0238	5.8749	9748
				(-.682)		
death				.004	.0146	13924
				(.464)		

*** p<0.01, ** p<0.05, * p<0.1

Results of the estimation of Equation 5. Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 2 on both sides of the cutoff using data from waves 1B and 2A from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects. Mortality is defined as dying by the end of 2018.

C Falsification tests for subgroups

In this Section, we replicate the falsification test for the subgroups present in our analysis. Tables C1 and C2 show the results by age.

Table C1: Falsification test by age for women

Variable	τ	CER-Optimal Bandwidth	Robust Inference p-value	Effective N Left	Effective N Right
Women					
Under 40					
age	0.204	0.912	0.761	1276	154
university education	-0.010	0.907	0.926	1276	154
born outside NL	-0.031	1.102	0.258	1281	155
employed	-0.020	1.109	0.690	1281	155
weekly hours worked	0.897	0.949	0.678	1267	153
disability insurance	0.009	0.888	0.775	553	142
GP visit	-0.035	0.946	0.429	1276	154
currently smoking	0.035	0.876	0.626	544	138
alcohol drinker	0.083	0.883	0.363	546	139
at least very good health	-0.033	0.900	0.648	1276	154
BMI	0.962	1.000	0.403	1276	154
Over 40					
age	0.362	0.642	0.276	3688	1482
university education	-0.011	1.085	0.658	7086	1615
born outside NL	0.000	0.793	0.938	3701	1489
employed	-0.014	0.996	0.537	7063	1613
weekly hours worked	-0.213	1.137	0.828	7015	1595
disability insurance	-0.002	0.961	0.869	7063	1613
retired	0.003	0.972	0.353	7063	1613
GP visit	-0.010	0.771	0.601	3701	1489
currently smoking	-0.058	0.830	0.042	3670	1469
alcohol drinker	0.000	0.855	0.936	3682	1479
at least very good health	-0.036	1.035	0.140	7063	1613
BMI	0.279	0.793	0.491	3701	1489

This falsification test looks for discontinuities in predetermined covariates at baseline. We estimate Equation 5 where the outcome is the baseline value of the control variable and we exclude the controls on the right-hand side of the equation. The coefficients are obtained via a multidimensional RDD performed on Lifelines data. We employed CER-optimal bandwidths, as suggested in [35], and a first-order polynomial on both sides of the cutoff.

Table C2: Falsification test by age for men

Variable	τ	CER-Optimal Bandwidth	Robust Inference p-value	Effective N Left	Effective N Right
Under 40					
age	0.598	0.925	0.301	1063	150
university education	0.015	1.204	0.831	2040	156
born outside NL	-0.003	1.369	0.855	2055	159
employed	0.005	0.998	0.876	1063	150
weekly hours worked	-0.004	1.455	0.990	3417	167
disability insurance	0.000	1.276	0.988	2055	158
GP visit	0.086	1.075	0.248	1079	150
currently smoking	-0.059	1.362	0.426	2014	156
alcohol drinker	-0.100	0.789	0.066	472	139
at least very good health	-0.087	1.067	0.262	1063	150
BMI	-0.209	0.937	0.700	1063	150
Over 40					
age	-0.205	0.906	0.514	5182	1149
university education	0.003	0.932	0.970	5238	1156
born outside NL	-0.011	0.899	0.302	5182	1149
employed	0.027	0.804	0.184	2709	1057
weekly hours worked	0.963	0.855	0.344	5155	1138
disability insurance	-0.009	0.715	0.453	2670	1047
retired	0.010	0.605	0.100	2645	1038
GP visit	0.014	0.710	0.678	2670	1047
currently smoking	-0.004	0.657	0.962	2637	1025
alcohol drinker	-0.004	0.892	0.870	5143	1136
at least very good health	0.005	0.804	0.854	2709	1057
BMI	0.088	0.702	0.885	2670	1047

This falsification test looks for discontinuities in predetermined covariates at baseline. We estimate Equation 5 where the outcome is the baseline value of the control variable and we exclude the controls on the right-hand side of the equation. The coefficients are obtained via a multidimensional RDD performed on Lifelines data. We employed CER-optimal bandwidths, as suggested in [35], and a first-order polynomial on both sides of the cutoff.

D Further descriptives

In this Section, we present descriptive statistics for women (Table D1) and men (Table D2) by age group.

Table D1: Descriptive statistics at baseline for women by age and values of the running variable

	under 40					over 40				
	1.5 sd below		1.5 sd above			1.5 sd below		1.5 sd above		
	$\bar{x}_b$	sd	$\bar{x}_a$	sd	$\bar{x}_b$-$\bar{x}_a$	$\bar{x}_b$	sd	$\bar{x}_a$	sd	$\bar{x}_b$-$\bar{x}_a$
women										
controls										
age	34.510	(3.431)	35.494	(3.099)	-0.984***	49.222	(5.536)	51.055	(5.526)	-1.833***
highly educated	0.394	(0.489)	0.290	(0.455)	0.104**	0.241	(0.428)	0.178	(0.383)	0.063***
lowly educated	0.606	(0.489)	0.710	(0.455)	-0.104**	0.759	(0.428)	0.822	(0.383)	-0.063***
born outside NL	0.043	(0.204)	0.037	(0.189)	0.006	0.025	(0.156)	0.035	(0.183)	-0.010*
outcomes										
employment	0.888	(0.316)	0.815	(0.390)	0.073**	0.829	(0.376)	0.773	(0.419)	0.057***
workweek hours	22.760	(12.554)	20.809	(13.901)	1.951	20.523	(13.525)	18.425	(13.987)	2.098***
disabled	0.017	(0.129)	0.043	(0.204)	-0.026*	0.032	(0.175)	0.044	(0.204)	-0.012**
retired	0.000	(0.000)	0.006	(0.079)	-0.006***	0.004	(0.061)	0.005	(0.073)	-0.002
gp visit	0.844	(0.363)	0.901	(0.299)	-0.058*	0.818	(0.386)	0.870	(0.336)	-0.052***
smoker	0.230	(0.421)	0.278	(0.449)	-0.047	0.210	(0.407)	0.234	(0.424)	-0.025*
alcohol drinking	0.658	(0.475)	0.660	(0.475)	-0.003	0.718	(0.450)	0.672	(0.470)	0.046***
very good health	0.354	(0.478)	0.278	(0.449)	0.077*	0.320	(0.466)	0.233	(0.423)	0.087***
BMI	26.108	(5.287)	29.702	(7.528)	-3.595***	26.379	(4.664)	28.832	(5.740)	-2.453***
Observations	4046		162		4208	16400		1672		18072

*** p<0.01, ** p<0.05, * p<0.1

Using data from the first wave of Lifelines, we compare women within 1.5 standard deviations below the threshold of the running variable to those within 1.5 standard deviations above it, stratified by age.

Table D2: Descriptive statistics at baseline for men by age and values of the running variable

	under 40					over 40				
	1.5 sd below		1.5 sd above			1.5 sd below		1.5 sd above		
	$\bar{x}_b$	sd	$\bar{x}_a$	sd	$\bar{x}_b$-$\bar{x}_a$	$\bar{x}_b$	sd	$\bar{x}_a$	sd	$\bar{x}_b$-$\bar{x}_a$
men										
controls										
age	34.560	(3.422)	35.802	(2.817)	-1.243***	48.642	(5.473)	50.056	(5.594)	-1.414***
highly educated	0.327	(0.469)	0.228	(0.421)	0.098**	0.286	(0.452)	0.252	(0.434)	0.034*
lowly educated	0.673	(0.469)	0.772	(0.421)	-0.098**	0.714	(0.452)	0.748	(0.434)	-0.034*
born outside NL	0.025	(0.156)	0.043	(0.204)	-0.018	0.021	(0.143)	0.037	(0.188)	-0.016***
outcomes										
employment	0.955	(0.207)	0.938	(0.241)	0.017	0.934	(0.248)	0.911	(0.285)	0.023**
workweek hours	39.753	(13.538)	39.735	(14.790)	0.018	38.164	(14.518)	37.466	(15.756)	0.699
disabled	0.015	(0.120)	0.037	(0.189)	-0.022*	0.024	(0.152)	0.036	(0.186)	-0.012**
retired	0.000	(0.000)	0.000	(0.000)	0.000	0.008	(0.090)	0.011	(0.102)	-0.002
gp visit	0.713	(0.453)	0.809	(0.395)	-0.096**	0.730	(0.444)	0.768	(0.423)	-0.037**
smoker	0.342	(0.475)	0.333	(0.473)	0.009	0.251	(0.434)	0.290	(0.454)	-0.040**
alcohol drinking	0.898	(0.303)	0.864	(0.344)	0.034	0.894	(0.307)	0.859	(0.348)	0.035***
very good health	0.397	(0.489)	0.228	(0.421)	0.168***	0.360	(0.480)	0.279	(0.449)	0.081***
BMI	26.334	(3.844)	29.117	(5.260)	-2.783***	26.987	(3.577)	28.688	(4.251)	-1.701***
Observations	3329		162		3491	12128		1226		13354

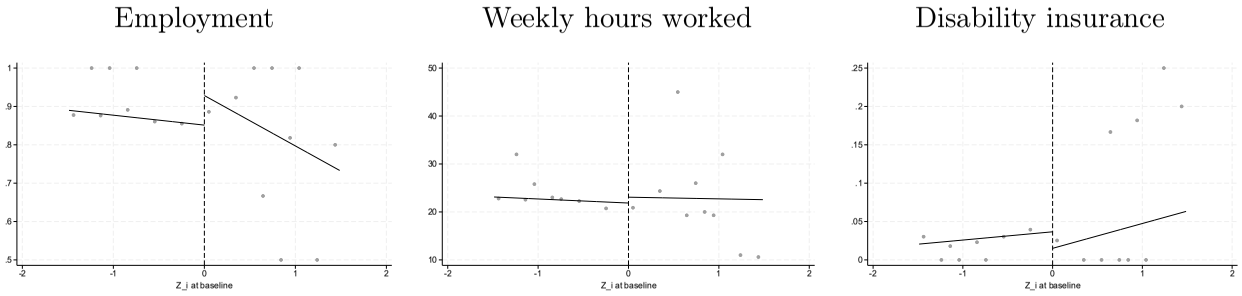
*** p<0.01, ** p<0.05, * p<0.1

Using data from the first wave of Lifelines, we compare men within 1.5 standard deviations below the threshold of the running variable to those within 1.5 standard deviations above it, stratified by age.

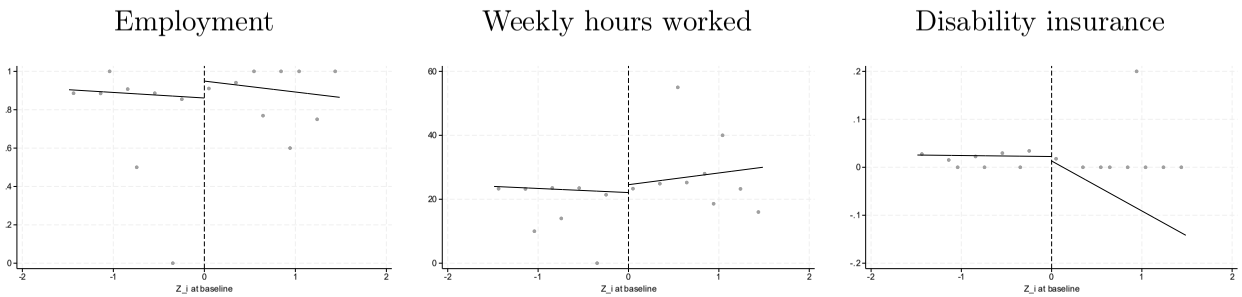
E Graphical RDD results

Figure E1: Results for women's labour market outcomes under 40

wave 1B

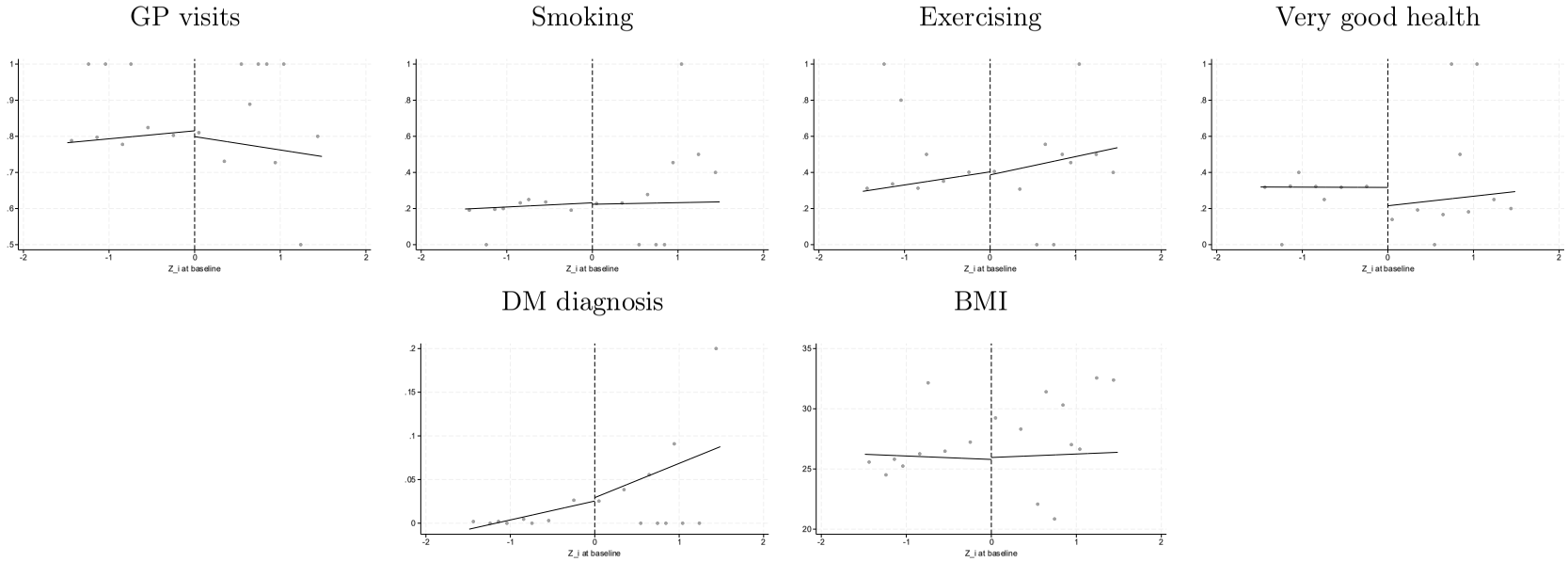


wave 2A



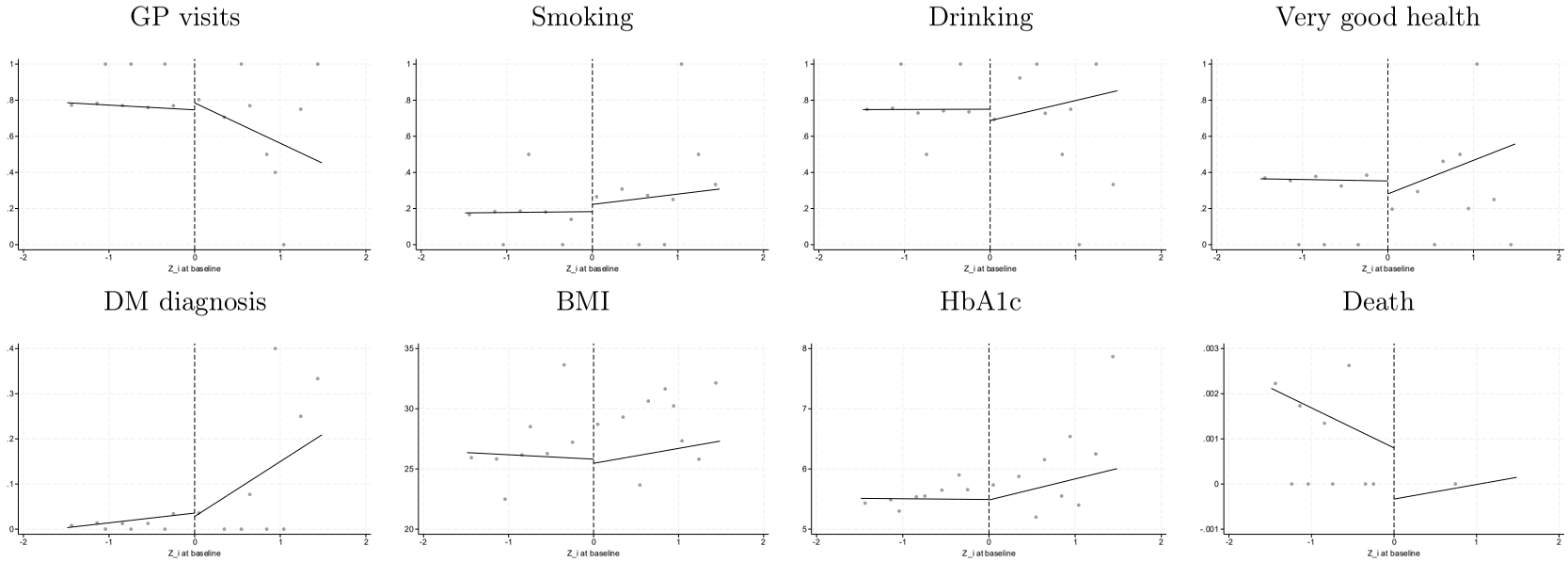
Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from waves 1B and 2A from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects.

Figure E2: Results for women's health behaviours and outcomes in wave 1B under 40



Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from waves 1B from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects.

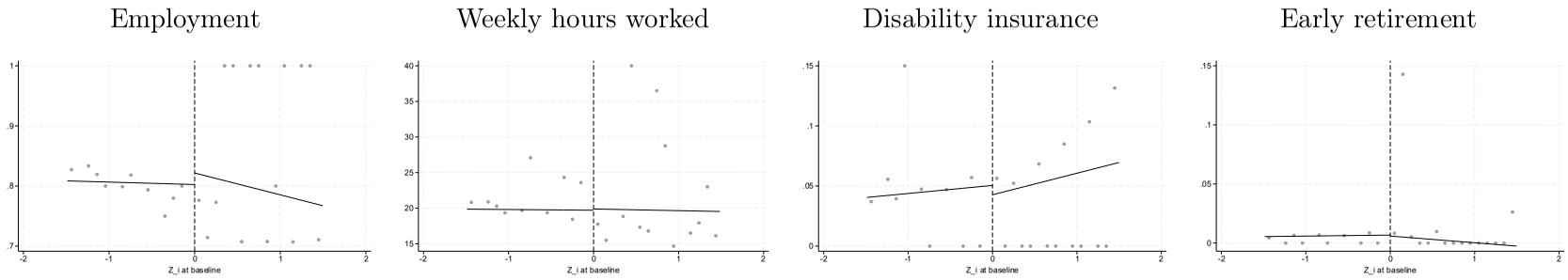
Figure E3: Results for women's health behaviours and outcomes in wave 2A under 40



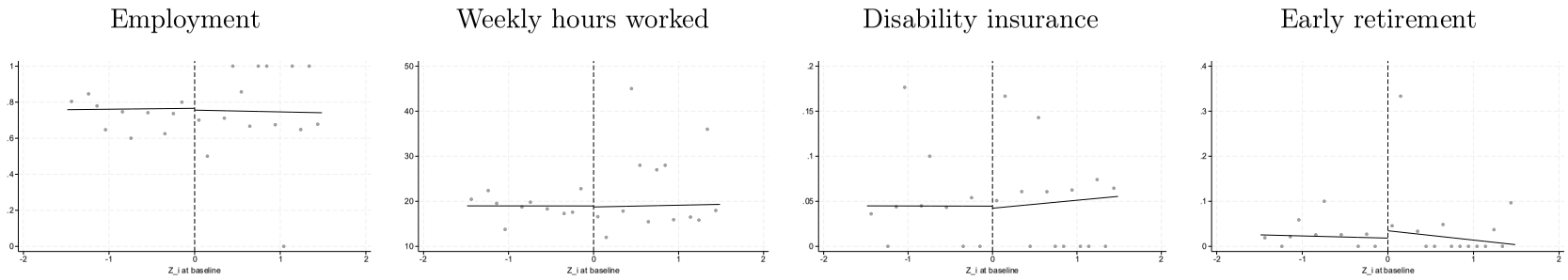
Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from wave 2A from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects.

Figure E4: Results for women's labour market outcomes over 40

wave 1B

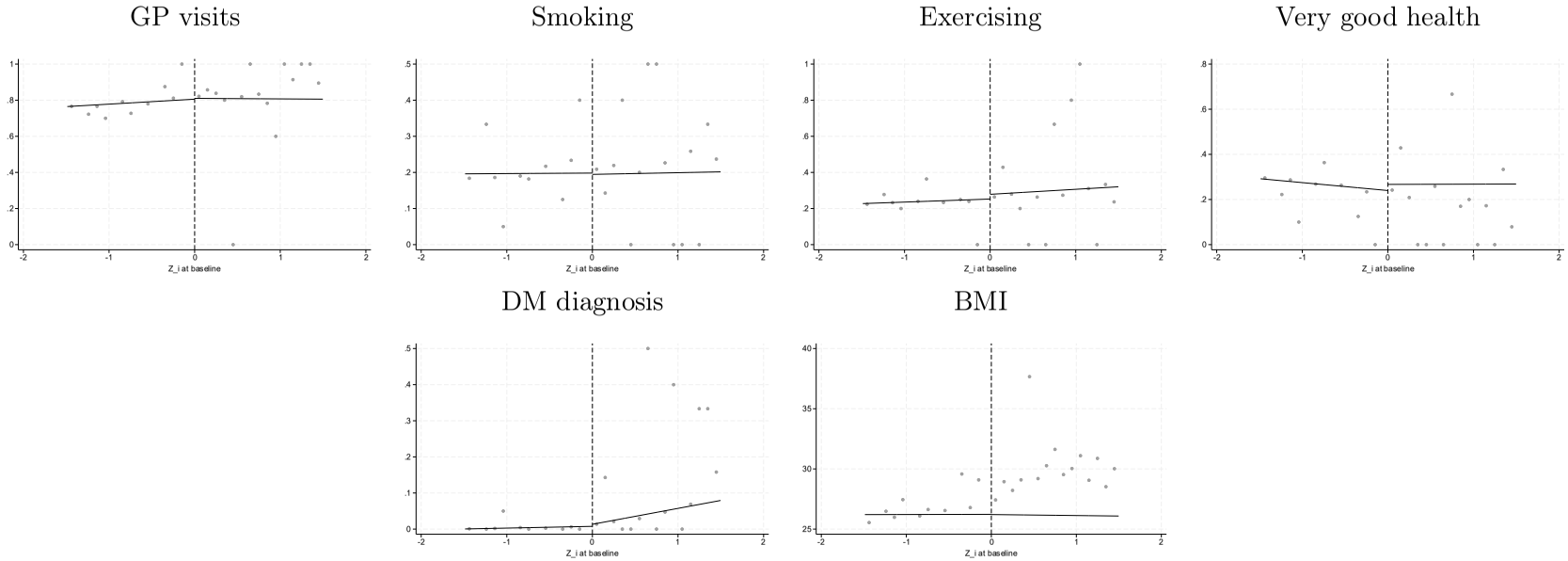


wave 2A



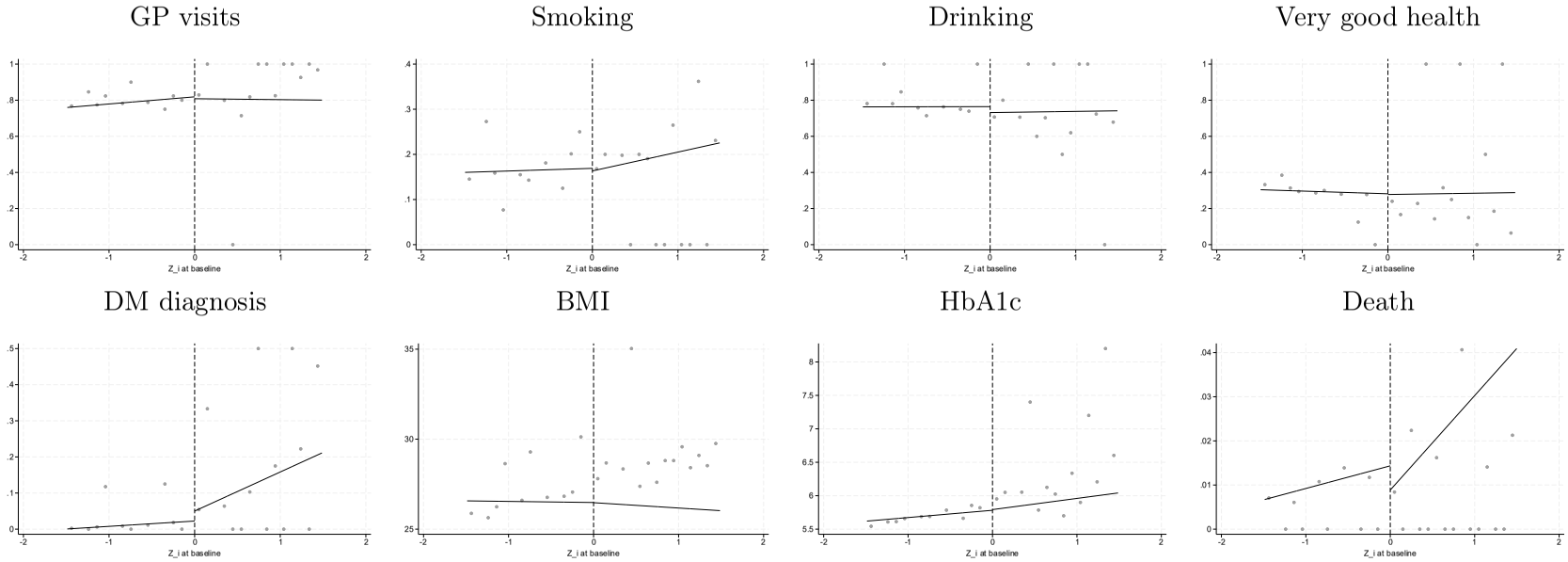
Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from waves 1B and 2A from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects.

Figure E5: Results for women's health behaviours and outcomes in wave 1B over 40



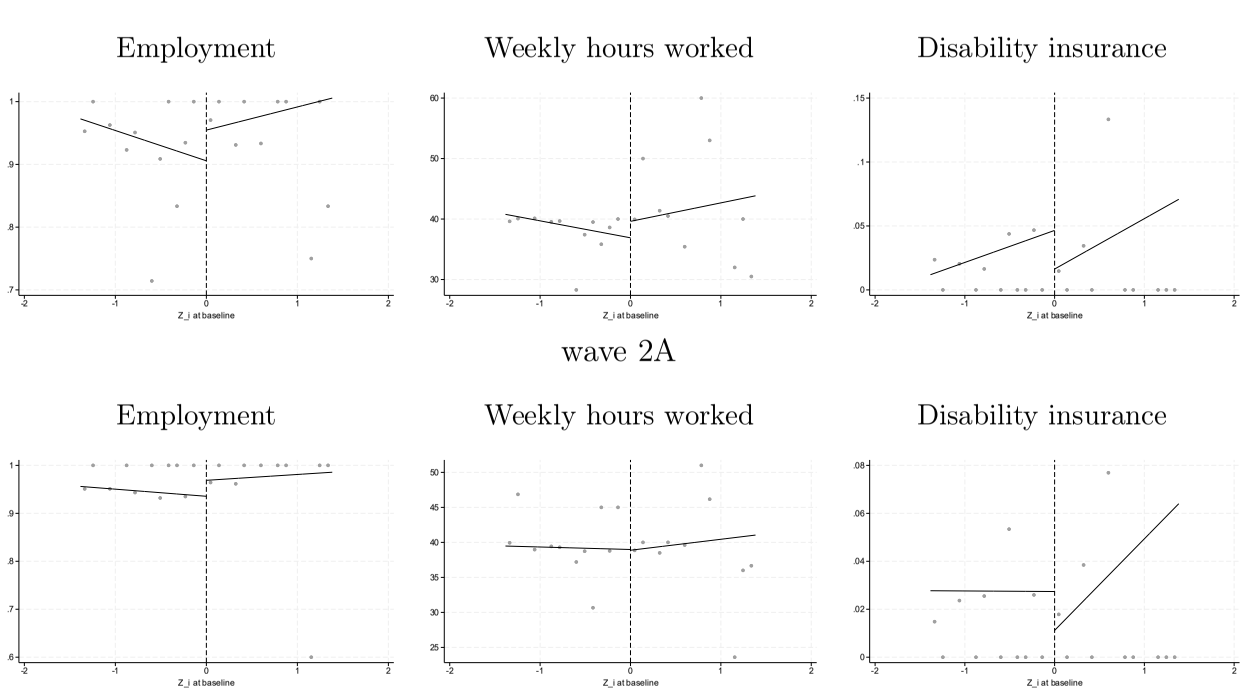
Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from waves 1B from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects.

Figure E6: Results for women's health behaviours and outcomes in wave 2A over 40



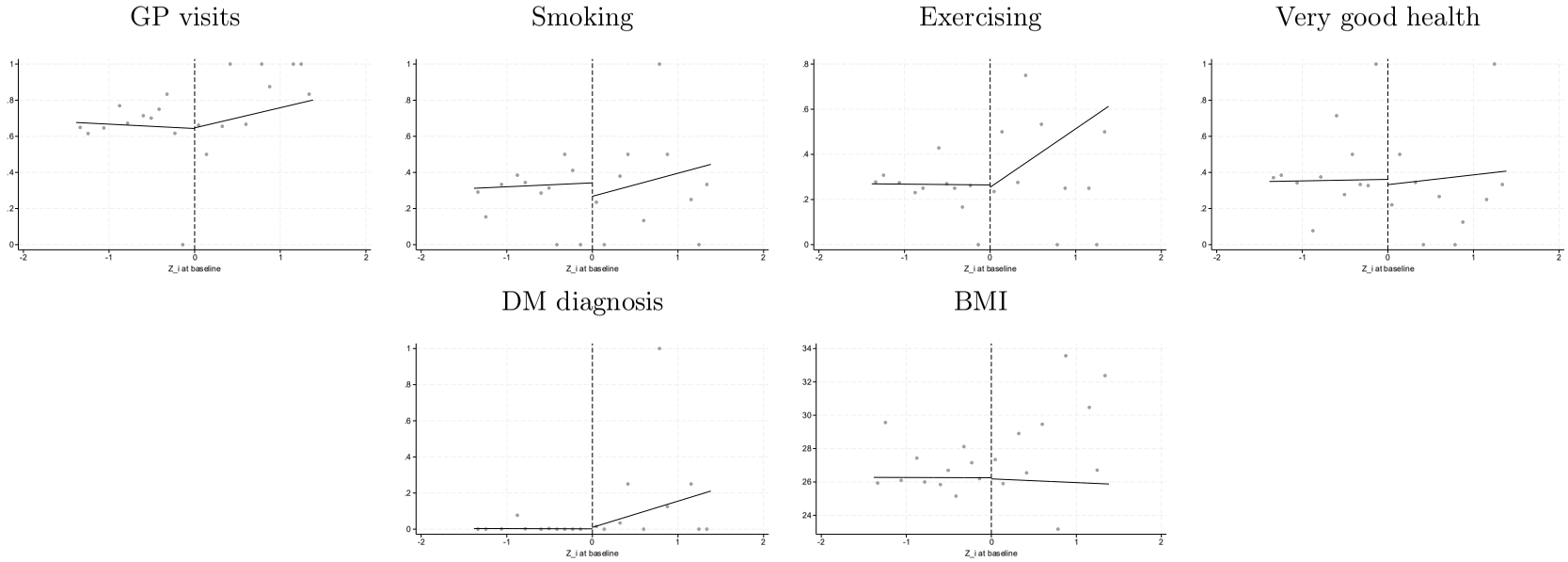
Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from wave 2A from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects.

Figure E7: Results for men's labour market outcomes under 40



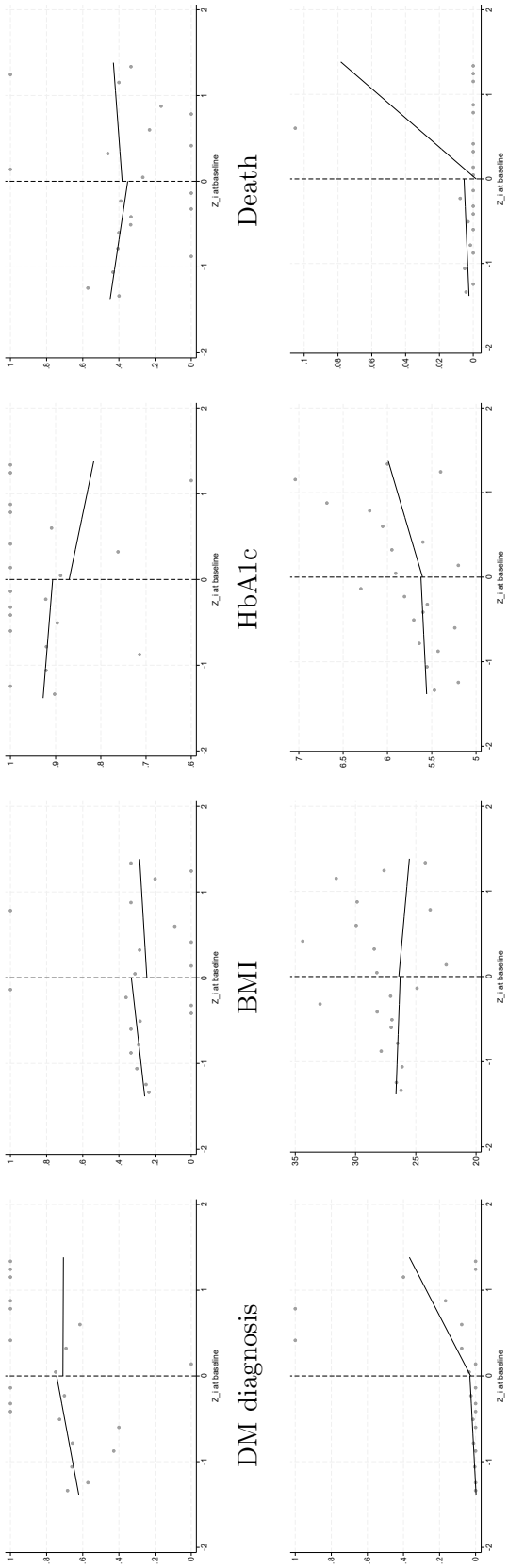
Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from waves 1B and 2A from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects.

Figure E8: Results for men's health behaviours and outcomes in wave 1B under 40



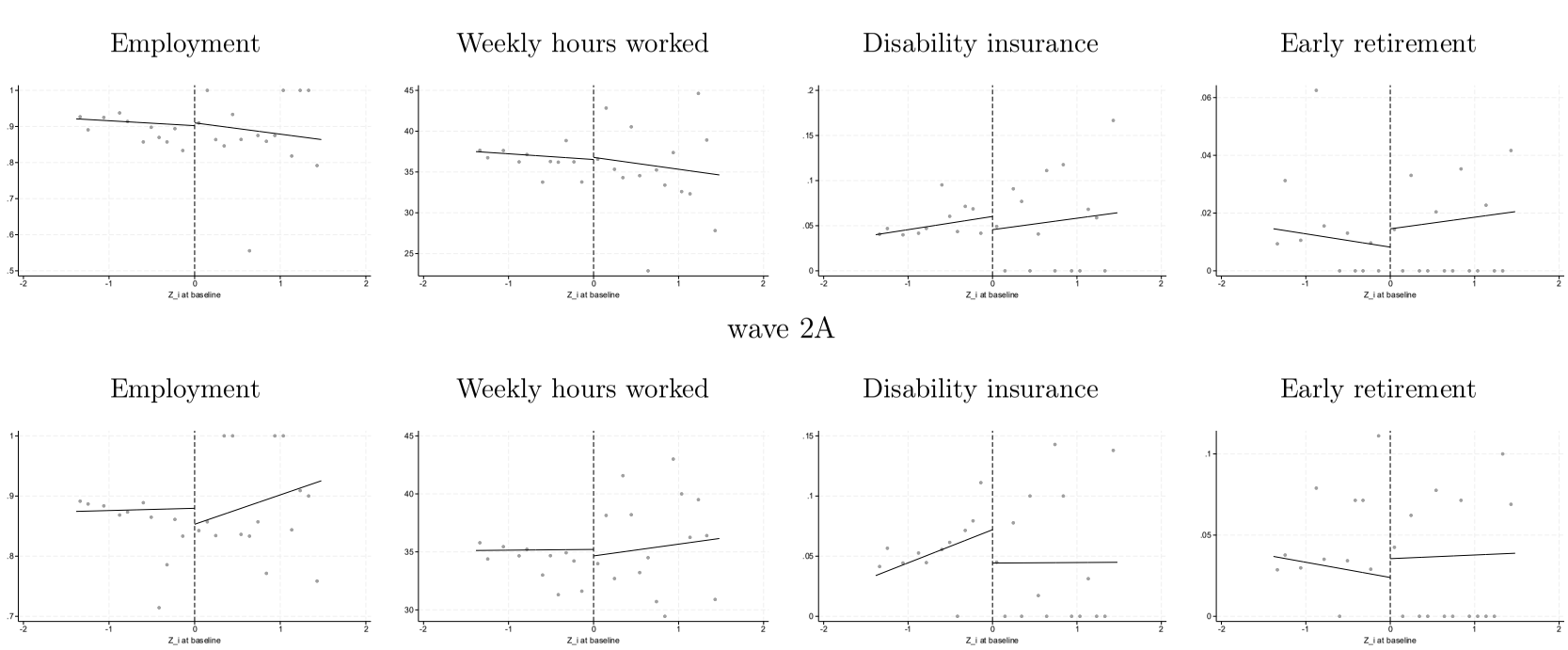
Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from waves 1B from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects.

Figure E9: Results for men's health behaviours and outcomes in wave 2A under 40



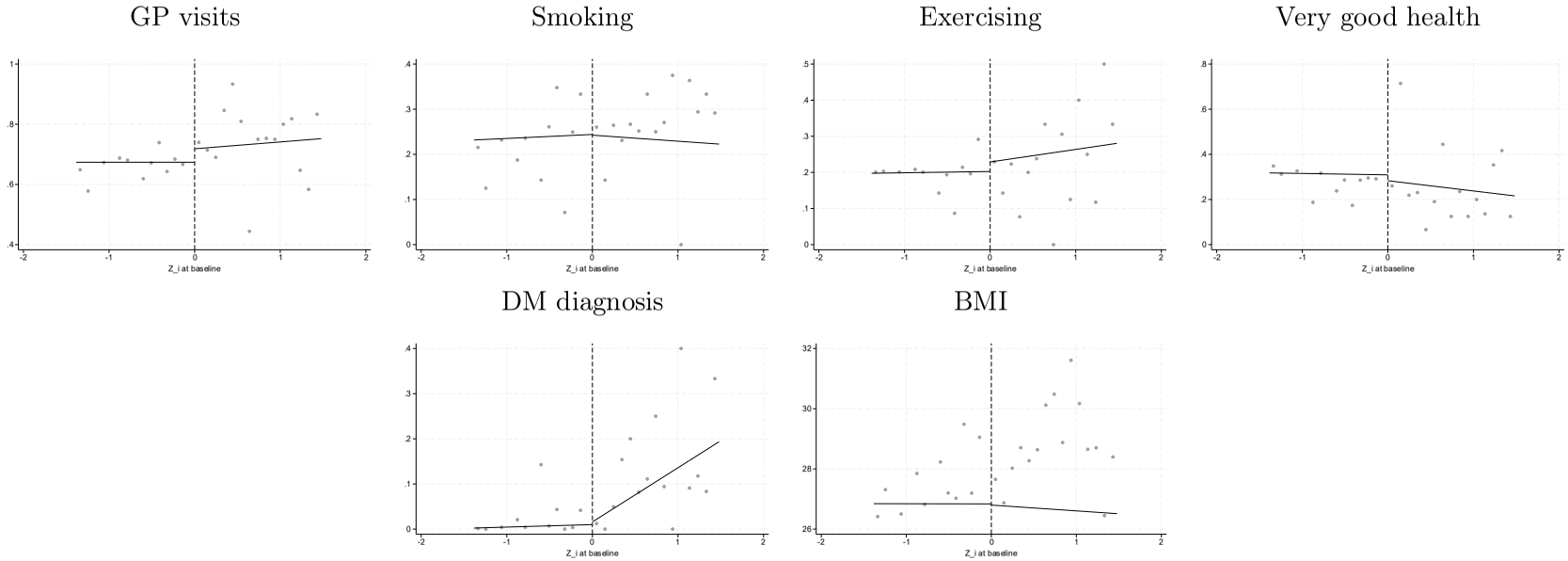
Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from wave 2A from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects.

Figure E10: Results for men's labour market outcomes over 40



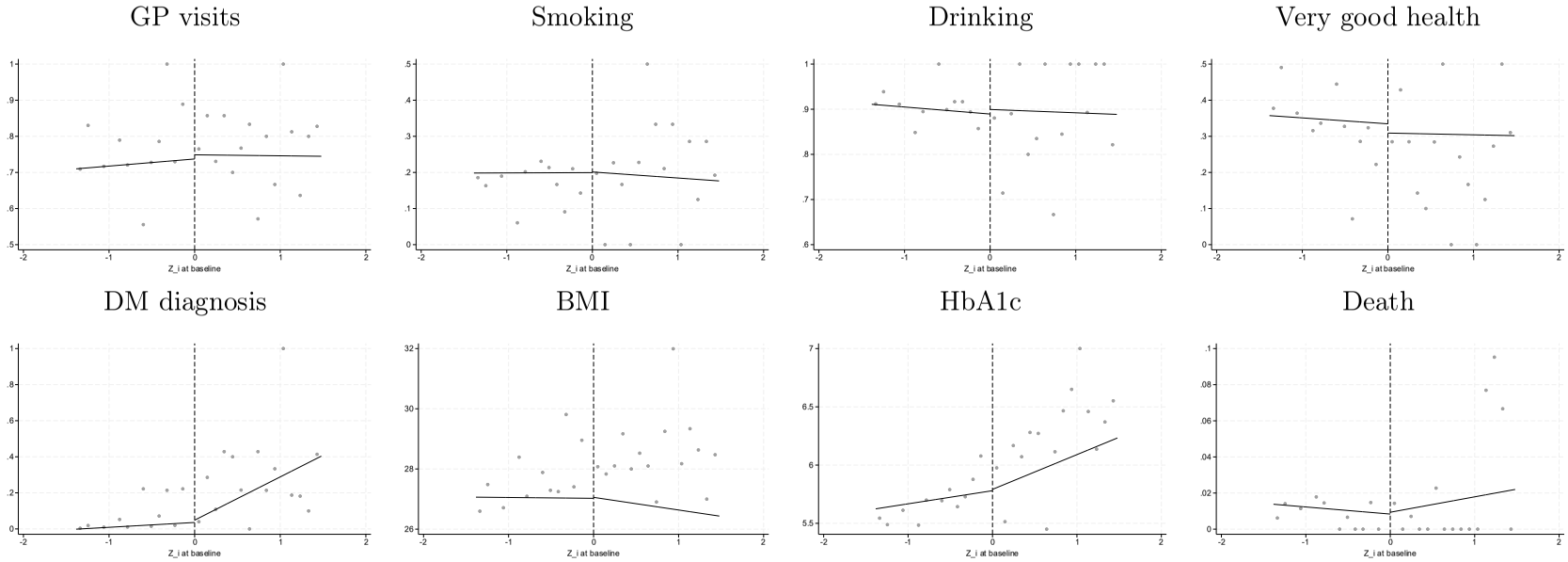
Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from waves 1B and 2A from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects.

Figure E11: Results for men's health behaviours and outcomes in wave 1B over 40



Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from waves 1B from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects.

Figure E12: Results for men's health behaviours and outcomes in wave 2A over 40



Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from wave 2A from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects.

F Multiple-testing p-values

In this Section, we present results with the FDR p-values. Tables F1, F2, and F3 replicate the main analysis by gender and age. Tables F4-F9 replicate the analysis by gender, age, and educational level.

Table F1: Effect of a warning on labour market outcomes by age and gender.

	Under 40						Over 40					
	1B (after 1.5 y)			2A (after 4 y)			1B (after 1.5 y)			2A (after 4 y)		
	τ	p-value	FDR p-value	τ	p-value	FDR p-value	τ	p-value	FDR p-value	τ	p-value	FDR p-value
Women												
employment	.0758	0.015	0.047	.0862	0.068	0.193	.0191	0.114	0.633	-.0095	0.605	1.000
	(2.433)			(1.824)			(1.581)			(-.517)		
weekly hours worked	1.1727	0.261	0.101	2.3977	0.108	0.193	.1793	0.613	0.692	-.1992	0.659	1.000
	(1.124)			(1.609)			(.505)			(-.442)		
disability insurance	-.0212	0.061	0.065	-.0094	0.523	0.212	-.0077	0.194	0.633	-.0024	0.681	1.000
	(-1.875)			(-.638)			(-1.3)			(-.411)		
early retirement							-.0008	0.877	0.782	.0163	0.065	0.352
							(-.154)			(1.845)		
Men												
employment	.0431	0.045	0.072	.0304	0.370	1.000	.0068	0.406	0.371	-.026	0.138	0.261
	(2.009)			(.896)			(.831)			(-1.485)		
weekly hours worked	2.4274	0.040	0.072	-.007	1.000	1.000	.1861	0.624	0.454	-.5713	0.477	0.380
	(2.05)			(.001)			(.49)			(-.711)		
disability insurance	-.0236	0.092	0.072	-.0192	0.195	1.000	-.0129	0.132	0.360	-.024	0.024	0.106
	(-1.685)			(-1.295)			(-1.505)			(-2.26)		
early retirement							.0059	0.129	0.360	.0115	0.221	0.284
							(1.517)			(1.223)		

Results of the estimation of Equation 5. Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from waves 1B and 2A from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects. The FDR p-values are calculated using the method presented in [49].

Table F2: Effect of a warning on women's health behaviours and outcomes by age.

	Under 40						Over 40					
	1B (after 1.5 y)			2A (after 4 y)			1B (after 1.5 y)			2A (after 4 y)		
	τ	p-value	FDR p-value	τ	p-value	FDR p-value	τ	p-value	FDR p-value	τ	p-value	FDR p-value
Health behaviours												
GP visits	-.0152	0.773	1.000	.0377	0.475	0.904	.0047	0.834	1.000	-.01	0.470	0.524
	(-.288)			(.788)			(.209)			(-.728)		
smoking	-.0079	0.667	1.000	.0401	0.425	0.904	-.0037	0.701	1.000	-.0059	0.515	0.524
	(-.43)			(.785)			(-.383)			(-.658)		
drinking				-.0627	0.403	0.904				-.0324	0.066	0.246
				(-.961)						(-1.797)		
exercising	-.0151	0.666	1.000				.0256	0.310	1.000			
	(-.431)						(1.015)					
Health outcomes												
very good health	-.1015	0.057	0.209	-.0711	0.289	1.000	.0267	0.094	0.275	-.0033	0.973	1.000
	(-2.036)			(-1.104)			(1.692)			(-.002)		
diagnosed	.0044	0.884	1.000	-.0074	0.845	1.000	.0057	0.144	0.275	.0275	0.006	0.029
	(.263)			(-.168)			(1.458)			(2.764)		
BMI	.1493	0.482	0.932	-.3574	0.355	1.000	-.0174	0.752	0.335	-.0106	0.923	1.000
	(.754)			(-.883)			(-.339)			(-.125)		
HbA1c				-.0092	0.870	1.000				.0109	0.666	1.000
				(-.167)						(.426)		
death				-.0011	0.660	1.000				-.0055	0.162	0.480
				(-.571)						(-1.393)		

Results of the estimation of Equation 5. Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from waves 1B and 2A from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects. Mortality is defined as dying by the end of 2018. The FDR p-values are calculated using the method presented in [49].

Table F3: Effect of a warning on men's health behaviours and outcomes by age.

	Under 40						Over 40					
	1B (after 1.5 y)			2A (after 4 y)			1B (after 1.5 y)			2A (after 4 y)		
	τ	p-value	FDR p-value	τ	p-value	FDR p-value	τ	p-value	FDR p-value	τ	p-value	FDR p-value
Health behaviours												
GP visits	-.0093	0.853	1.000	-.0126	0.821	0.664	.041	0.084	0.338	.0137	0.586	1.000
	(-.185)			(-.279)			(1.727)			(.553)		
smoking	-.0734	0.056	0.201	-.0833	0.071	0.269	-.0022	0.907	0.700	.0003	0.834	1.000
	(-1.914)			(-1.8)			(-.117)			(.234)		
drinking				-.0493	0.266	0.363				.0085	0.515	1.000
				(-.976)						(.613)		
exercising	-.0058	0.963	1.000				.026	0.274	0.379			
	(-.047)						(1.093)					
Health outcomes												
very good health	-.0347	0.399	1.000	.0028	0.973	1.000	-.0234	0.329	1.000	-.0255	0.295	1.000
	(-.823)			(.006)			(-.993)			(-1.066)		
diagnosed	.0083	0.537	1.000	-.001	0.899	1.000	.0055	0.481	1.000	.0147	0.453	1.000
	(.611)			(-.164)			(.712)			(.752)		
BMI	-.0957	0.682	1.000	.0527	0.813	1.000	-.0339	0.835	1.000	.0267	0.680	1.000
	(-.424)			(.216)			(-.195)			(.418)		
HbA1c				-.0183	0.728	1.000				.0189	0.693	1.000
				(-.389)						(.401)		
death				-.005	0.339	1.000				-.0006	0.918	1.000
				(-.965)						(-.104)		

Results of the estimation of Equation 5. Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from waves 1B and 2A from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects. Mortality is defined as dying by the end of 2018. The FDR p-values are calculated using the method presented in [49].

Table F4: Effect of a warning on women's labour market outcomes by educational level.

	under 40						over 40					
	1B (after 1.5 y)			2A (after 4 y)			1B (after 1.5 y)			2A (after 4 y)		
		average			average			average			average	
	τ	0.5 sd below	Effective N	τ	0.5 sd below	Effective N	τ	0.5 sd below	Effective N	τ	0.5 sd below	Effective N
low education												
employed	.0861	0.033	0.112	.101	0.085	0.111	.0146	0.335	1.000	-.007	0.676	1.000
	(2.015)			(1.675)			(1.004)			(-.312)		
weekly hours worked	1.6423	0.194	0.149	3.7509	0.033	0.111	-.0135	0.978	1.000	-.135	0.732	1.000
	(1.229)			(2.058)			(.014)			(-.246)		
disability	-.0226	0.109	0.123	-.0125	0.405	0.156	-.0035	0.530	1.000	.0012	0.991	1.000
	(-1.47)			(-.583)			(-.612)			(.02)		
retired							.0026	0.422	1.000	.02	0.049	0.244
							(.825)			(2.003)		
high education												
employed	.0483	0.146	0.434	.0525	0.531	1.000	.0217	0.286	0.393	-.01	0.716	1.000
	(1.453)			(.627)			(1.068)			(-.363)		
weekly hours worked	.0931	0.904	0.434	-.4462	0.749	1.000	.5438	0.581	0.410	-.3415	0.711	1.000
	(.12)			(-.32)			(.552)			(-.371)		
disability	-.0188	0.202	0.434	-.0072	0.718	1.000	-.0214	0.099	0.393	-.0185	0.280	1.000
	(-1.277)			(-.361)			(-1.649)			(-1.08)		
retired							-.0129	0.141	0.393	.0004	0.843	1.000
							(-1.472)			(.198)		

Results of the estimation of Equation 5. Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from waves 1B and 2A from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects. The FDR p-values are calculated using the method presented in [49].

Table F5: Effect of a warning on men's labour market outcomes by educational level.

	under 40						over 40					
	1B (after 1.5 y)			2A (after 4 y)			1B (after 1.5 y)			2A (after 4 y)		
	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N
low education												
employment	.0535	0.069	0.116	.0252	0.609	1.000	-.0016	0.946	0.899	-.0455	0.036	0.065
	(1.829)			(.517)			(.019)			(-2.133)		
work_hours	3.009	0.050	0.116	-.6616	0.690	1.000	.3802	0.477	0.466	-1.0802	0.318	0.087
	(1.975)			(-.353)			(.641)			(-1.041)		
disability	-.0312	0.117	0.116	-.0283	0.221	1.000	-.0198	0.064	0.346	-.0297	0.029	0.065
	(-1.597)			(-1.291)			(-1.81)			(-2.148)		
retired							.0052	0.146	0.346	.0204	0.045	0.065
							(1.46)			(2.01)		
high education												
employment	.0067	0.573	1.000	.0298	0.380	1.000	.0407	0.020	0.090	.0116	0.698	1.000
	(.563)			(.879)			(2.318)			(.388)		
work_hours	1.1853	0.506	1.000	.0594	0.975	1.000	.4917	0.498	0.992	.3052	0.852	1.000
	(.665)			(.032)			(.678)			(.186)		
disability	.0042	0.265	1.000	.0012	0.809	1.000	-.0015	0.878	1.000	-.0152	0.235	1.000
	(1.114)			(.242)			(-.153)			(-1.186)		
retired							.01	0.378	0.992	-.0076	0.758	1.000
							(.882)			(-.308)		

Results of the estimation of Equation 5. Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from waves 1B and 2A from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects. The FDR p-values are calculated using the method presented in [49].

Table F6: Effect of a warning on women's health behaviours outcomes by age and with low educational.

	under 40						over 40					
	1B (after 1.5 y)			2A (after 4 y)			1B (after 1.5 y)			2A (after 4 y)		
	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N
health behaviours												
GP visit	-.0321	0.457	0.806	.086	0.203	0.611	-.0032	0.827	1.000	-.0386	0.044	0.104
	(-.607)			(1.352)			(-.225)			(-2.02)		
smoking	-.0363	0.149	0.806	.0178	0.797	0.611	-.0076	0.527	1.000	-.0081	0.477	0.189
	(-1.374)			(.239)			(-.646)			(-.729)		
alcohol				-.1156	0.253	0.611				-.0385	0.062	0.104
				(-1.265)						(-1.814)		
exercising	-.0632	0.320	0.806				.0299	0.295	1.000			
	(-1.057)						(1.059)					
health outcomes												
very good health	-.0803	0.217	0.482	-.0588	0.466	1.000	.0237	0.197	1.000	.0026	0.844	1.000
	(-1.383)			(-.779)			(1.313)			(.232)		
diagnosed	.0027	0.948	0.482	-.0139	0.641	1.000	.0039	0.336	1.000	.0164	0.124	1.000
	(.211)			(-.428)			(.953)			(1.528)		
BMI	.3987	0.123	0.482	-.5307	0.323	1.000	-.0314	0.698	1.000	-.0214	0.868	1.000
	(1.589)			(-.937)			(-.408)			(-.187)		
HbA1c				.0218	0.744	1.000				.0105	0.714	1.000
				(.321)						(.357)		
death				-.0025	0.416	1.000				-.0032	0.448	1.000
				(-.938)						(-.752)		

Results of the estimation of Equation 5. Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from waves 1B and 2A from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects. The FDR p-values are calculated using the method presented in [49].

Table F7: Effect of a warning on women's health behaviours outcomes by age and with high educational level.

	under 40						over 40					
	1B (after 1.5 y)			2A (after 4 y)			1B (after 1.5 y)			2A (after 4 y)		
	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N
health behaviours												
GP visit	.0044	0.834	1.000	-.107	0.510	1.000	.0236	0.598	1.000	.0915	0.034	0.115
	(.209)			(-.659)			(.527)			(2.116)		
smoking	.0458	0.447	1.000	.0795	0.214	1.000	.0055	0.780	1.000	-.0077	0.688	0.847
	(.76)			(1.243)			(.28)			(-.402)		
alcohol				.0676	0.491	1.000				-.0258	0.425	0.740
				(.689)						(-.797)		
exercising	.0722	0.506	1.000				.0061	0.874	1.000			
	(.664)						(.159)					
health outcomes												
very good health	-.1558	0.105	0.127	-.1178	0.280	0.875	.0503	0.156	0.444	-.0172	0.871	1.000
	(-1.621)			(-1.08)			(1.418)			(-.162)		
diagnosed	.0094	0.944	0.460	.0156	0.816	0.875	.0115	0.205	0.444	.0732	0.005	0.025
	(.07)			(.233)			(1.268)			(2.816)		
BMI	-.7005	0.038	0.127	-.0482	0.937	0.875	.0488	0.818	0.444	.0738	0.735	1.000
	(-2.08)			(.079)			(.23)			(.339)		
HbA1c				-.074	0.248	0.875				.0125	0.769	1.000
				(-1.155)						(.294)		
death				.0017	0.129	0.875				-.0148	0.067	0.156
				(1.517)						(-1.83)		

Results of the estimation of Equation 5. Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from waves 1B and 2A from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects. The FDR p-values are calculated using the method presented in [49].

Table F8: Effect of a warning on men's health behaviours outcomes by age and with low educational.

	under 40						over 40					
	1B (after 1.5 y)			2A (after 4 y)			1B (after 1.5 y)			2A (after 4 y)		
	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N
health behaviours												
GP visit	-.0116	0.786	1.000	.0349	0.577	0.625	.0326	0.210	0.459	.0259	0.344	1.000
	(-.27)			(.482)			(1.27)			(.959)		
smoking	-.0676	0.142	0.739	-.1422	0.017	0.056	-.0031	0.849	0.459	-.0005	0.871	1.000
	(-1.469)			(-2.362)			(-.16)			(.191)		
alcohol				-.0463	0.400	0.625				.0151	0.359	1.000
				(-.676)						(.873)		
exercising	.0164	0.737	1.000				.0413	0.107	0.459			
	(.337)						(1.599)					
health outcomes												
very good health	.017	0.821	1.000	.0077	0.903	1.000	-.0044	0.948	1.000	-.0128	0.608	1.000
	(.241)			(.169)			(-.094)			(-.543)		
diagnosed	-.0002	0.815	1.000	-.0204	0.490	1.000	.0013	0.927	1.000	.015	0.477	1.000
	(.224)			(-.743)			(-.081)			(.715)		
BMI	-.2348	0.207	1.000	-.1222	0.772	1.000	-.0231	0.972	1.000	-.0307	0.881	1.000
	(-1.271)			(-.309)			(.052)			(-.142)		
HbA1c				-.0145	0.752	1.000	(-.408)			.0227	0.682	1.000
				(-.37)						(.42)		
death				-.0005	0.946	1.000				-.0036	0.597	1.000
				(-.08)						(-.531)		

Results of the estimation of Equation 5. Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from waves 1B and 2A from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects. The FDR p-values are calculated using the method presented in [49].

Table F9: Effect of a warning on men's health behaviours outcomes by age and with high educational level.

	under 40						over 40					
	1B (after 1.5 y)			2A (after 4 y)			1B (after 1.5 y)			2A (after 4 y)		
	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N	τ	average 0.5 sd below	Effective N
health behaviours												
GP visit	.0097	0.853	1.000	-.1567	0.105	0.460	.0642	0.207	1.000	.0023	0.927	1.000
	(.185)			(-1.622)			(1.263)			(-.092)		
smoking	-.0747	0.288	1.000	.0519	0.387	0.460	.0053	0.819	1.000	.0044	0.847	1.000
	(-1.062)			(.865)			(.229)			(.194)		
alcohol				-.0657	0.377	0.460				-.0086	0.696	1.000
				(-.884)						(-.391)		
exercising	-.031	0.767	1.000				-.0095	0.729	1.000			
	(-.296)						(-.346)					
health outcomes												
very good health	-.1012	0.215	1.000	.0059	0.912	1.000	-.0764	0.095	0.383	-.0591	0.281	1.000
	(-1.241)			(-.111)			(-1.67)			(-1.079)		
diagnosed	.027	0.538	1.000	.0121	0.916	1.000	.0144	0.184	0.383	.0142	0.742	1.000
	(.616)			(.105)			(1.327)			(.329)		
BMI	.3975	0.429	1.000	.041	0.785	1.000	-.0451	0.677	0.383	.1792	0.210	1.000
	(.79)			(.273)			(-.417)			(1.252)		
HbA1c				-.0336	0.802	1.000				.0138	0.821	1.000
				(-.251)						(.226)		
death				-.0176	0.250	1.000				.0046	0.642	1.000
				(-1.151)						(.465)		

Results of the estimation of Equation 5. Estimates are generated via a regression discontinuity design, with a triangular kernel and a polynomial of order 1 on both sides of the cutoff using data from waves 1B and 2A from Lifelines. The estimation bandwidth is 1.5 standard deviations computed by gender at baseline. We employ robust standard errors and display robust z-statistics in parentheses. In the regressions, we control for the baseline value of the outcome variable, a second-order polynomial of age, educational level, country of birth, year of inclusion and current year fixed effects. The FDR p-values are calculated using the method presented in [49].