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

DISCUSSION PAPER SERIES

IZA DP No. 15403

Sick and Depressed? The Causal Impact of a Diabetes Diagnosis on Depression

Alessio Gaggero
Joan Gil
Dolores Jiménez-Rubio
Eugenio Zucchelli

JUNE 2022

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15403

Sick and Depressed? The Causal Impact of a Diabetes Diagnosis on Depression

Alessio Gaggero
University of Granada

Dolores Jiménez-Rubio
University of Granada

Joan Gil
Universitat de Barcelona

Eugenio Zucchelli
MIAS, Lancaster University, IZA and Universidad Autónoma de Madrid

JUNE 2022

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

Sick and Depressed? The Causal Impact of a Diabetes Diagnosis on Depression*

There is sparse evidence on the impact of health information on mental health as well as on the mechanisms governing this relationship. We estimate the causal impact of health information on mental health via the effect of a diabetes diagnosis on depression. We employ a fuzzy regression discontinuity design (RDD) exploiting the exogenous cut-off value in the diagnosis of type-2 diabetes provided by a biomarker (glycated haemoglobin) and information on diagnosed clinical depression drawn from rich administrative longitudinal data from Spain. We find that overall a type-2 diabetes diagnosis increases the probability of becoming depressed, however this effect appears to be driven mostly by women. Results also appear to differ by changes in lifestyle induced by the diabetes diagnosis: while women who did not lose weight are more likely to develop depression, men who did lose weight present a reduced probability of being depressed. Results are robust to alternative parametric and non-parametric specifications and placebo tests.

JEL Classification: C21, I10, I12

Keywords: diabetes, depression, fuzzy regression discontinuity design, administrative longitudinal data, lifestyle changes

Corresponding author:
Eugenio Zucchelli
Department of Economic Analysis
Economic Theory and Economic History
Universidad Autónoma de Madrid (UAM)
C/Francisco Tomás y Valiente 5
28049, Madrid
Spain

E-mail: eugenio.zucchelli@uam.es

* We would like to thank participants to the EvaluAES session of the Conference of the Spanish Health Economics Association (Zaragoza, June 2022) for their useful suggestions. We gratefully acknowledge financial support from the Spanish Ministry of Science, Innovation and Universities (grant number PID2019-105688RB-I00). Eugenio Zucchelli also acknowledges support from the Tomás y Valiente Fellowship funded by the Madrid Institute for Advanced Study (MIAS), Universidad Autónoma de Madrid (UAM), and grants H2019/HUM-5793 and PID2019-111765GB-I00 funded by the Regional Government of Madrid and the Spanish Ministry of Science, Innovation and Universities, respectively.

1. Introduction

An increasing body of evidence suggests the relevance of health information in influencing key health-behaviours. For instance, the medical literature finds that the information provided by portable devices (e.g. Patel et al., 2015; Jo et al., 2019) or the diagnosis of specific types of cancer (e.g. Jazieh et al., 2006; Burris et al., 2015) might trigger behavioural changes and ultimately affect health outcomes. More recently, the economics literature has started exploring the role of health information by focusing on the impact of the diagnosis of chronic conditions, including hypertension and diabetes (e.g. Zhao et al., 2013; Kim et al., 2019; Gaggero, 2020; Gaggero et al. 2021; Iizuka et al., 2021). While these recent economics studies often employ causal inference methods and are thus capable of identifying causal effects, they mostly focus on changes in health-behaviours while ignoring potential spillover effects, including those on relevant health outcomes such as mental health.

Mental health and depression have been consistently found to be linked with health-behaviours. A series of studies (e.g. Conway et al., 2016; Conry et al., 2011; Scott and Happell, 2011; Jacka et al., 2010; Luppino et al., 2010; Stathopoulou et al., 2006; Lasser et al., 2000) show that individuals with healthier patterns of lifestyle behaviours present lower levels of psychological distress. In addition, the relationship between chronic conditions and mental health is well-documented with several papers suggesting strong correlations between major chronic conditions such as diabetes and mental health (e.g. Pan et al., 2010; Mezuk et al., 2013; Deschênes et al., 2015; Feng and Astell-Burt, 2017; Robinson et al., 2018). However, the existing literature does not appear to have comprehensively investigated the role of health information in influencing mental health.

The main objective of this paper is to identify the causal impact of health information on mental health by identifying the impact of the diagnosis of type-2 diabetes (T2DM) on clinical depression via a regression discontinuity design (RDD). More specifically, we exploit the discontinuity offered by the exogenous cut-off of a biomarker commonly used for the diagnosis of T2DM (i.e. glycated haemoglobin, HbA1c) to estimate the impact of a T2DM diagnosis on diagnosed clinical depression using rich longitudinal administrative data from Spain. Our results show that overall, a T2DM diagnosis increases the probability of developing clinical depression and that this result seems to be driven by women. Importantly, the effect of health information following a diagnosis appears to differ considerably by changes in lifestyle induced by the diabetes diagnosis: while men who lose weight following a T2DM diagnosis show a reduced probability of developing clinical depression, a diabetes diagnosis increases

the probability of developing clinical depression irrespective of lifestyle changes among women. These results are robust to a series of sensitivity analysis, including alternative econometric specifications and placebo tests.

This paper offers several contributions to the literature. First, we provide novel causal evidence of the impact of health information on depression, one of the most widespread mental disorders affecting around 280 million individuals globally (WHO, 2020). While the growing literature on the role of health information has mainly focused on its effect on health-behaviours especially among individuals with chronic conditions, to the best of our knowledge no previous studies have attempted to identify the causal impact of health information on diagnosed clinical depression. Second, our analysis suggests that the effect of a type-2 diabetes diagnosis on depression might vary by gender as well as by the lifestyle changes induced by it. More specifically, our analysis suggests the presence of a potentially protective effect of health information among male patients: a type-2 diabetes diagnosis appears to induce male individuals to reduce their weight and this might in turn decrease their risk of developing clinical depression. Hence, this contributes directly to the literature on the relevance of health information and the mechanisms through which it might affect health outcomes. From a policy perspective this might be also potentially useful as it highlights that the provision of health information (in the form of a diabetes diagnosis) could positively affect two highly relevant health outcomes (weight losses and mental health). Fourth, our empirical analysis also contributes to the large strand of the literature concerned with the determinants of mental health among individuals with major chronic conditions, including obesity and type-2 diabetes. This is also likely to be relevant policy-wise as type-2 diabetes is currently affecting 537 million individuals, with a higher prevalence among men, and its burden of disease is projected to increase in both developing and developed countries (IDF, 2021).

2. Previous Literature

Several recent studies in the fields of medicine and economics have explored the effects of health information. Previous medical papers focused either on the effects of the diagnosis of specific types of cancers (e.g. Jazieh et al., 2006; Burris et al., 2015) or the role of the information provided by portable devices (e.g. Patel et al., 2015; Jo et al., 2019) on clinical outcomes as well as risky health-behaviours including smoking, alcohol drinking and dietary changes. However, these findings are mostly based on standard statistical associations and are either mixed or observed only among specific sub-groups of individuals.

The economics literature has also started exploring the role of health information in influencing health and health-behaviours. This is highly relevant as standard economic models assume that individuals have complete knowledge about their health, and they can perfectly and rationally process it when making health investment decisions (e.g. Grossman, 1972; Cawley and Ruhm, 2011). However, this assumption has been recently re-assessed by empirical and experimental studies (e.g. Bhargava et al., 2017; Kettlewell, 2020; Arni et al., 2021). Early economic studies focus on the effects of public health information campaigns (e.g. Brown and Schrader, 1990; Chern at al., 1995; Kim and Chern, 1999; Roosen et al., 2009) or nutritional labels (e.g. Alleis et al., 2015; Fichera and von Hinke, 2020) while more recent contributions attempt to identify the causal impact of health information via the diagnosis of chronic conditions, including T2DM, on lifestyle behaviours (e.g., Zhao et al., 2013; Kim et al., 2019; Gaggero, 2020; Iizuka et al., 2021; Gaggero et al., 2021) and cardiovascular risk factors (Fukuma et al., 2020). The latter studies tend to increasingly find significant causal impacts of a T2DM diagnosis mostly on weight loss or fat intake.

Moreover, the dual relationship between diabetes and mental health has been investigated, mainly in the medical literature. Several of these studies have shown a significant association between the diagnosis of T2DM and the deterioration of mental health (e.g. Pan et al., 2010; Mezuk et al., 2013; Deschênes et al., 2015; Feng and Astell-Burt, 2017) with negative effects on quality of life, social contacts (Saito et al., 2006; Feng and Astell-Burt, 2017), and increases in the consumption of antidepressants (Mezuk et al., 2013). In addition, a consistent finding in the medical literature is that major depressive disorders increase the risk of developing T2DM and subsequent complications (Eaton et al., 1996; Estrodi and Kenardy, 2006; Pan et al., 2010). Gaggero (2019) appears to be among the very few economics studies attempting to identify the effect of a diabetes diagnosis on a measure of mental health. However, he only employs self-reported information on mental health together with a less reliable biomarker to detect diabetes on a sample of older UK individuals, finding no statistically significant effects of a T2DM diagnosis.

Importantly, most medical evidence tends to rely on self-reported information of key variables of interest and the statistical models previously used are often only capable of identifying standard statistical correlations, while overlooking potential endogeneity issues. As a result, the literature has not yet established whether health information may have a *causal* impact on developing a diagnosed mental health condition and, more specifically, whether a T2DM diagnosis might causally affect clinical depression.

3. Data and Key Variables

Data

We employ administrative and longitudinal individual-level data drawn from six GP practices and two hospitals located in Badalona, Spain, an EU country with a universal health care system free at the point of delivery (Bernal et al., 2018). The initial sample includes patients aged 16+ who had at least one contact with those hospitals and centres between 1 January 2004 and 31 December 2010.[5]

The dataset contains detailed information about patients' clinical measurements of height and weight and any diagnosed health condition, including clinical or major depression. More specifically, clinical depression is identified by a binary variable taking value 1 if the patient is diagnosed with clinical/major depression, corresponding to the code/registry P76 of the International Classification of Primary Care, second edition (ICPC-2), 0 otherwise. The information used by physicians to diagnose depression is based on a series of psychometrically validated measures of clinical depression, including the Golberg and Hamilton depression scales as well as the Geriatric Depression Scale (GDS) (Hamilton, 1967; Goldberg, 1993; Yesavage et al, 1983). All three measures are collected by physicians through interviews with patients on the basis of a series of items identifying several symptoms of depression experienced by the patients during the previous week (15 items in the short form of the GDS index or 30 in the longer version; 17 in the Hamilton and 18 in the Goldberg depression scales, respectively). While the Goldberg and the Hamilton scales can be used to detect depression in the general population, the GDS is an instrument specifically designed to diagnose depression among older patients.

Other key variables and descriptive statistics

Our medical records also include data on glycated haemoglobin (henceforth, HbA1c), a biomarker providing a reliable measure of a patient's average blood sugar level in the previous 8-12 weeks that is commonly used to diagnose T2DM (IEC, 2009). In our setting, physicians follow standard national and international medical guidelines for patients with T2DM (Mata et al., 2013) and use the threshold value of HbA1c $\geq$ 6.5 percent to diagnose T2DM.[6] This test is

[5] We exclude from the analysis those patients who were transferred or moved to other centres. However, since movements across centres are rare events, this exclusion is unlikely to influence our results.

[6] Note that while HbA1c was universally considered the main tool to diagnose diabetes only from 2009 (IEC, 2009), its use was already widespread, and the Spanish national health care system was routinely using it as one of the key measures to diagnose diabetes during the years used for this analysis.

administered as part of routine health checks to all individuals presenting relevant risk factors or symptoms of hyperglycaemia (high blood sugar levels).

HbA1c measurements are endorsed by the International Expert Committee (IEC) and the American Diabetes Association (ADA) as they are more reliable if compared to other measures of blood sugar such as the ones based on Fasting Plasma Glucose (FPG). The latter appears to have a substantially shorter time validity; to be sensitive to short-term lifestyle changes and stress; and tend to systematically underestimate the prevalence of diabetes (IEC, 2009; ADA 2020; Ho-Pham et al., 2017).[7] Relevant to this study, upon a T2DM diagnosis, patients of the Spanish health care system are normally recommended to follow a non-pharmacological treatment consisting of educational training sessions for diabetes self-management aimed at improving their lifestyle though dietary changes and regular exercise.

In addition, the dataset includes a rich set of demographic and socioeconomic characteristics such as age; gender; employment status (active/retired); marital status (married/cohabiting vs living alone); immigration status (EU vs non-EU), that we use as control variables. For the purpose of our analysis, we include in our estimating sample individuals with at least one biomarker measurement per year. This effectively includes any individual either diagnosed with diabetes; at risk of diabetes (including pre-diabetics, i.e. patients with a HbA1c value between 5.7-6.4 percent); or any other individual with relevant risk factors or symptoms that may lead to high blood sugar levels. This leads to a sample of 39994 individuals.

Table 1 reports the summary statistics of the main variables of interest. The Table shows that around 18 percent of the patients are diagnosed with clinical depression. We next report statistics on individuals diagnosed with T2DM via the corresponding ICPC-2 code informed by the HbA1c values. The average HbA1c for the patients in our sample is around 6.6 percent and on average patients have been diagnosed for a little over 3 years (see the variable labelled "onset"). With respect to other demographic variables, the average age of the sample is around 65, and the sample is almost evenly split by gender. Furthermore, 87, 27 and 2 percent of the sample are, respectively, living with a partner; active in the labour market; and were born outside the EU. Finally, the Table also reports that 59 and 53 percent of the patients are also diagnosed with hypertension and dyslipidaemia (the presence of high amounts of lipids,

[7] Notice patients have usually several HbA1c measurements per year. Since one single measurement per patient and year is needed, we opted for calculating both a within yearly mean HbA1c value and the last year observed measurement. In addition, given the differences between type-1 and type-2 diabetes and related treatments, we dropped all individuals with type-1 diabetes.

including cholesterol, in blood), respectively; while 4, 7 and 5 percent of the patients are affected by asthma, neoplasms/cancers and COPD respectively.

4. Empirical Approach

We employ a *fuzzy* regression discontinuity design (RDD) exploiting the discontinuity offered by the exogenous cut-off value of 6.5 of the biomarker (HbA1c) used to diagnose T2DM. We choose to follow a fuzzy design as physicians may not base their diagnosis solely on HbA1c values. For instance, they could potentially look at further patients' characteristics as well as family history around T2DM or whether individuals may suffer from other metabolic conditions, such as hypertension or dyslipidaemia. For instance, it might be the case that some physicians may diagnose with T2DM individuals with several metabolic conditions and a value of HbA1c just below 6.5 percent. Our fuzzy RDD approach would allow accounting for such cases.

A fuzzy RDD estimation is akin to a two-stage least squares instrumental variable (2SLS-IV) specification (Hahn et al., 2001). Accordingly, the first stage equation can be represented as follows:

$$D_{i,t} = \mu + \rho Above_{i,t} + f(HbA1c - 6.5) + \boldsymbol{X}'_{\boldsymbol{i,t}}\Omega + \epsilon_{i,t}; \qquad (1)$$

where $D_{i,t}$ is a dummy indicator for whether individual i was diagnosed with T2DM at time t. $Above_{it}$ is a binary variable that takes the value of 1 when the HbA1c of an individual is above the predetermined cut-off value of 6.5 and acts as the instrument for the T2DM diagnosis. $f(HbA1c - 6.5)$ is a function of the centred HbA1c, i.e. our running variable allowed to vary around the cut-off.[8] Following Gelman and Imbens (2018) our main specification considers a linear polynomial of the running variable. However, we examine the robustness of our results to higher order polynomials and non-parametric estimations based on a local randomization approach (Cattaneo et al., 2020). All estimations control for a vector of covariates $\boldsymbol{X}'_{\boldsymbol{i,t}}$ including sociodemographic characteristics (age, gender, employment, marital and immigrant status). We also control for several pre-diagnosed conditions, including hypertension, dyslipidaemia, asthma, neoplasms/cancers, and chronic obstructive pulmonary

[8] In practice, we allow the function form to vary on either side of the cut-off by including an interaction term between the binary variable $Above_{it}$ and the running variable.

disease (COPD) as well as time elapsed from the diagnosis. Additionally, our econometric specifications include time-, health area- and GP fixed-effects (FE). This allows controlling for any systematic (time-invariant) differences across physicians that might affect the diagnoses of T2DM and major depression.

The second stage equation can be written as:

$$Y_{i,t+1} = \alpha + \beta D_{i,t} + g(HbA1c - 6.5) + \boldsymbol{X'_{i,t}}\gamma + \varepsilon_{i,t}. \quad (2)$$

$Y_{i,t+1}$ denotes the outcome of interest and measures whether individual i is diagnosed with clinical depression at time $t+1$, that is *after* the T2DM diagnosis, conditional on not having been diagnosed at time t. As above, the function $g(HbA1c - 6.5)$ is a flexible polynomial of the recentred running variable, and it is also allowed to vary around the cut-off. $\boldsymbol{X'_{it}}$ is the same vector of covariates discussed above. Finally, $\varepsilon_{i,t}$ is a random error term. We cluster standard errors on the running variable based on the recommendation of Lee and Card (2008).[9] The main term of interest is β as it measures the change in the probability of being diagnosed with depression following a T2DM diagnosis. To make full use of the sample size, we estimate Equations (1) and (2) parametrically. However, we further explore the robustness of our results using the non-parametric approach mentioned above as part of our sensitivity analysis.

4.1 RDD Validity

The main identifying assumption of our RDD approach leading to the estimation of an unbiased causal estimate is that, conditional on covariates, the only difference between individuals above and below the cut-off is the treatment assignment (Hahn et al., 2001). Accordingly, the average outcome of those just below the cut-off can be used as the counterfactual for those above the cut-off (Lee and Lemieux, 2010). Although this assumption cannot be tested directly (as we cannot directly observe counterfactuals), we provide two indirect tests suggesting the overall credibility of our RDD application.

First, we examine whether there are any significant differences in pre-determined characteristics at the cut-off point. Ideally, we should find null effects of the diagnosis on these characteristics. The results of this exercise are presented in Figure 1. Each graph presents the

[9] Note that following Kolesár and Rothe (2018), we also estimate our models using Eicker-Huber-White (EHW) heteroskedasticity-robust standard errors. These are normally recommended when the number of support points around the cut-off is sufficiently large and are based on a smaller bandwidth. Results are very similar and available upon request.

local polynomial smoothing (LPS) for pre-diagnosis major depression and covariates (age, gender, active, immigrant, hypertension, dyslipidaemia, asthma, neoplasm/cancer, and COPD) as a function of the HbA1c. This confirms the validity of our design by revealing non-significant jumps at the cut-off for any of these variables.

Second, for the RDD to be valid it is also critical that individuals would not be capable of manipulating their diabetes diagnosis (McCrary, 2008). In our case, it seems highly unlikely that patients could manipulate their HbA1c scores, as this measure is based on a blood test administered by physicians and refers to the average glucose concentration over the previous 8-12 weeks. Yet, to exclude this, in Figure 2 we provide the distribution of the density function of the (centred) HbA1c around the cut-off, suggesting the absence of any discontinuity, as expected.

5. Results

5.1 Main Results

Figures 3a-b examine the impact of having a HbA1c level above 6.5 percent. Specifically, Figure 3a shows a sizeable discontinuity in the probability of being diagnosed with T2DM around the HbA1c cut-off as for our first stage. Similarly, the plot in Figure 3b implies that patients with (normalised) HbA1c just above the cut-off are more likely to be diagnosed with depression than their counterparts. We next test the relevance of these findings in a regression framework while controlling for potential confounding factors.

Table 2 reports the RDD estimates of a T2DM diagnosis, as outlined in Equation (1). Specifically, column (1) includes the basic estimates with no covariates; column (2) adds a set of socioeconomic characteristics; column (3) further accounts for dummies indicating the presence of pre-existing medical conditions; column (4) includes time and area FE; and column (5) finally adds GP FE. Estimates appear statistically significant across all specifications. According to the most comprehensive specification in column (5), the estimated coefficient implies that patients with an HbA1c above the 6.5 percent cut-off are around 9 percentage points more likely to be diagnosed with T2DM than their counterparts.

Table 3 reports fuzzy RDD estimates of a T2DM diagnosis of being diagnosed with major or clinical depression. Similarly to Table 2, in columns (1)-(5) we report findings of different specifications including an incremental number of covariates. Here, we observe positive and statistically significant effect of a T2DM diagnosis on clinical depression. In particular, according to our preferred model (column 5), receiving a T2DM diagnosis raises by 1.6 percent the probability of being diagnosed with depression.

5.2 Heterogeneity and potential mechanisms

Table 4 presents results split by gender and lifestyle changes induced by the T2DM diagnosis, namely whether individuals lose weight following the diagnosis, as this is often one of the main lifestyle changes recommended by physicians according to the recent literature (e.g. Kim et al., 2019; Gaggero et al., 2021; Seuring et al., 2021). We do this to explore potential heterogeneity and mechanisms driving the observed impact on clinical depression.

Columns (1)-(3) report baseline RDD estimates for the full sample, males and females, respectively. These estimates clearly suggest that the increase in the probability of developing clinical depression appears to be driven by women: women present an increase of around 3.2 percentage points in the probability of being diagnosed with clinical depression following a diabetes diagnosis, while the corresponding estimate for men is not statistically significant. This appears to be in line with previous findings indicating that male patients with diabetes present higher levels of subjective well-being (Siddiqui et al., 2013, Deischinger et al., 2020).

We also explore whether losing weight after a T2DM diagnosis might also play a role in explaining the effects on clinical depression. Columns (4)-(6) and (7)-(9) in Table 4 present results for individuals who did versus who did not lose weight, respectively, for the whole sample and for subsamples of men and women. Importantly, size and statistical significance of these estimates suggest that the observed increase in depression is mostly concentrated among female patients, particularly among those who did not lose weight following a diabetes diagnosis (5.8 percentage points). Yet, men who lost weight after the T2DM diagnosis, are less likely to be diagnosed with major depression (8.3 percentage points). Overall, this appears to suggest the presence of a potentially relevant "protective" effect with respect to the probability of being diagnosed with clinical depression but only among male individuals.

5.3 Sensitivity Analysis

In order to check the robustness of our main results, we also present a series of sensitivity tests. Table A1 shows RDD estimates based on a placebo test consisting of alternative cut-off values of the biomarker. This test should not produce a statistically significant effect for our outcome of interest (clinical depression) at values below the 6.5 percent of the biomarker. Indeed, our placebo test using 5.5, 5 and 4.5 percent cut-off values confirm that the impact on major depression is not statistically different from zero. At the 6 percent value of the biomarker we find statistically significant effects, although smaller in magnitude. Yet, this is expected as prediabetic patients (HbA1c ranging from 5.7-6.4 percent) are normally recommended similar

non-pharmacological treatments (dietary changes and regular exercise) (Mata et al., 2013). At the 7 percent theshold we find slightly larger effects of a T2DM diagnosis on mental health, and this is also in line with the lifestyle changes recommended by doctors to patients with uncontrolled diabetes.

Table A2 further shows the robustness of our results by estimating the fuzzy RDD parametrically by means of different polynomial orders, ranging from a polynomial of order 1, column (1), to a polynomial of order 4, column (4). Importantly, the corresponding estimates obtained are qualitatively similar.

Finally, Table A3 reports RDD estimates produced using a non-parametric local randomization approach focusing on observations within a small neighbourhood around the cut-off (Cattaneo et al., 2020). For comparative purposes, column (1) of Table A3 reports the baseline RDD estimates obtained using the parametric approach. Columns (2)-(6) show non-parametric RDD estimates based on different choices of bandwidth, starting from a bandwidth of 2 and then gradually decreasing it until reaching a bandwidth of 1, around the threshold of 6.5 percent. Results confirm that all findings are consistent, similar in magnitude to our baseline estimates, and still statistically significant.

6. Conclusions and Discussion

We contribute to the literature by identifying the causal impact of health information on mental health via the effect of a type-2 diabetes diagnosis on clinical depression. We exploit the exogenous cut-off value in the diagnosis of type-2 diabetes provided by a well-established biomarker (glycated haemoglobin) and information on diagnosed clinical depression drawn from rich administrative longitudinal data from Spain by employing a fuzzy regression discontinuity design. In addition, we explore heterogeneity in the effects on mental health by gender as well as the role played by weight losses as a potential mechanism leading to changes in the probability of developing clinical depression, following a diabetes diagnosis.

Our results suggest a statically significant impact of health information on diagnosed clinical depression. However, this appears to be driven by gender as well as weight losses eventually occurring after the diabetes diagnosis. More specifically, the overall increase in the risk of developing clinical depression following a diabetes diagnosis appears to be mostly influenced by women. Moreover, the occurrence of weight losses could be one of the possible mechanisms governing the relationship between a diabetes diagnosis and depression. Here differences by gender are also present: whereas a diabetes diagnosis increases the probability of depression among women who did not lose weight, it substantially decreases the risk of clinical depression

among men who lost weight. Interestingly, this may suggest a somewhat protective effect of health information via weight losses for male patients with diabetes.

In general, the finding that individuals with healthier behavioural patterns present lower levels of mental disorders is also supported by the medical literature (e.g. Conway et al., 2016; Conry et al. (2011); Scott and Happell, 2011; Jacka et al., 2010; Luppino et al., 2010; Stathopoulou et al., 2006; Lasser et al., 2000). As for the different effects by gender, these could be explained on several grounds. For instance, the literature suggests that women might have a higher propensity to clinical depression driven by biological factors (Deischinger, 2020). Second, evidence also suggests that depression tend to be underdiagnosed among men (Deischinger, 2020). Third, the mental health of individuals, especially that of women, with a high baseline weight may not be significantly affected by weight losses (e.g. Simon et al. 2010). This might be the case in our data as well, where women present higher rates of obesity as well as a higher average baseline (i.e. pre-diabetes diagnosis) BMI if compared to male patients (30.9 vs 29.35, respectively, with the difference being statistically significant as suggested by a pairwise t-test). However, further evidence might be needed to firmly establish the reasons behind heterogeneity in the effects by gender.

As usual, this study may have some potential limitations. First, as our dataset does not include comprehensive information on the medications prescribed by physicians, we might not be able to account for those pharmaceuticals potentially affecting mental health. Second, the administrative records used here only includes a relatively limited number of variables proxying socioeconomic status. While this might not necessarily be a major issue given that the Spanish health care system is universal and free at the point of delivery (i.e. access does not depend on income or the purchase of an insurance policy), we might not be able to identify potentially informative socioeconomic gradients. Despite such limitations, our study provides novel empirical evidence on the causal impact of health information on mental health, shedding light on gender-based differences in such effects and potential mechanisms through changes in lifestyle behaviours.

References

American Diabetes Association (ADA) January 01 2020; volume 43 issue Supplement 1.

Arni, P., Dragone, D., Goette, L., Ziebarth, N. R., 2021. Biased Health Perceptions and Risky Health Behaviors—Theory and Evidence. *Journal of Health Economics*, 21(76) 102415.

Bernal-Delgado, E., García-Armesto, S., Oliva, J., Martínez, F., Repullo, J., Peña-Longobardo, L., ... & Hernández-Quevedo, C. (2018). *Spain-health System Review*.

Bhargava, S., Loewenstein, G., Sydnor, J., 2017. Choose to lose: health plan choices from a menu with dominated option. *The Quarterly Journal of Economics*, 132 (3),1319–1372.

Brown, D.J. and Schrader, L.F., 1990. Cholesterol information and shell egg consumption. *American Journal of Agricultural Economics*, *72*(3), pp.548-555.

Burris, J.L., Studts, J.L., DeRosa, A.P. and Ostroff, J.S., 2015. Systematic review of tobacco use after lung or head/neck cancer diagnosis: results and recommendations for future research. *Cancer Epidemiology and Prevention Biomarkers*, *24*(10), pp.1450-1461.

Cawley, J. and Ruhm, C.J., 2011. The economics of risky health behaviors. In *Handbook of Health Economics* (Vol. 2, pp. 95-199). Elsevier.

Conry, M.C., Morgan, K., Curry, P., McGee, H., Harrington, J., Ward, M. and Shelley, E., 2011. The clustering of health behaviours in Ireland and their relationship with mental health, self-rated health and quality of life. *BMC Public Health*,*11*(1), pp.1-10.

Conway, K. P., Swendsen, J., Husky, M. M., He, J. P., & Merikangas, K. R. (2016). Association of lifetime mental disorders and subsequent alcohol and illicit drug use: results from the National Comorbidity Survey–Adolescent Supplement. *Journal of the American Academy of Child & Adolescent Psychiatry*, *55*(4), 280-288.

Cattaneo M.D., Idrobo N., Titiunik R., 2020. A Practical Introduction to Regression Discontinuity Designs (Elements in Quantitative and Computational Methods for the Social Sciences). *Cambridge University Press*.

Deschênes, S. S., Burns, R. J., & Schmitz, N. (2015). Associations between diabetes, major depressive disorder and generalized anxiety disorder comorbidity, and disability: Findings from the 2012 Canadian Community Health Survey—Mental Health (CCHS-MH). *Journal of Psychosomatic Research*, *78*(2), 137-142.

Deischinger, C., Dervic, E., Leutner, M., Kosi-Trebotic, L., Klimek, P., Kautzky, A., & Kautzky-Willer, A. (2020). Diabetes mellitus is associated with a higher risk for major depressive disorder in women than in men. *BMJ Open Diabetes Research and Care*, 8(1), e001430.

Eaton, W.W., et al. (1996). Depression and risk for onset of type II diabetes. A prospective population-based study. *Diabetes Care* 19(10): 1097-1102.

Estrodi and Kenardy (2006). Psychosocial and non-psychosocial risk factors for the new diagnosis of diabetes in elderly women. *Diabetes Research and Clinical Practice* 74(1): 57-65.

Feng, X., & Astell-Burt, T. (2017). Impact of a type 2 diabetes diagnosis on mental health, quality of life, and social contacts: a longitudinal study. *BMJ Open Diabetes Research and Care*, *5*(1), e000198.

Fichera, E. and von Hinke, S., 2020. The response to nutritional labels: Evidence from a quasi-experiment. *Journal of Health Economics*, 72.102326.

Fukuma, S., Iizuka, T., Ikenoue, T., Tsugawa, Y. 2020. Association of the National Health Guidance Intervention for obesity and cardiovascular risks with health outcomes among Japanese men. *JAMA Internal Medicine* 2020; 180(12): 1630-1637.

Gaggero, A., 2020. The effect of type 2 diabetes diagnosis in the elderly. *Economics & Human Biology*, 37, pp. 1-23.

Gaggero, A., Gil, J., Jiménez-Rubio, D., Zucchelli, E. 2021. Health information and lifestyle behaviours: the impact of a diabetes diagnosis. *IZA Discussion Paper Series No. 14106*.

Gelman, A., Imbens, G., 2018. Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs. *Journal of Business & Economic Statistics* 37(3), pp. 447-456.

Goldberg, I.K. 1993. Questions and answers about depression and its treatment: a consultation with a leading psychiatrist. *Charles Press Publishers*, 1993.

Grossman, M. 1972. On the concept of health capital and the demand for health. *Journal of Political Economy* 80 (2), 223–255.

Hahn, J., Todd, P., van der Klaauw, W., 2001. Identification and Estimation of Treatment Effects with Regression-Discontinuity Design. *Econometrica*, 69(1), pp. 201-209.

Hamilton, M. (1967). Development of a rating scale for primary depressive illness. *British Journal of Clinical Psychology*, 6, 278–296.

Ho-Pham, L.T., Nguyen, U.D., Tran, T.X. and Nguyen, T.V., 2017. Discordance in the diagnosis of diabetes: Comparison between HbA1c and fasting plasma glucose. *PLoS One*, *12*(8), p.e0182192.

Iizuka, T., Nishiyama, T., Chen, B., Eggleston, K., 2021. False alarm? Estimating the marginal value of health signals, *Journal of Public Economics*, 195, 104368.

International Diabetes Federation (2021) IDF diabetes atlas: tenth edition 2021. International Diabetes Federation, Brussels, Belgium. Available at IDF Diabetes Atlas 2021 | IDF Diabetes Atlas.

International Expert Committee, 2009. International Expert Committee Report on the Role of the A1C Assay in the Diagnosis of Diabetes. *Diabetes Care*, 32(7), 1327-1334.

Jacka, F. N., Pasco, J. A., Mykletun, A., Williams, L. J., Hodge, A. M., O'Reilly, S. L., ... & Berk, M. (2010). Association of Western and traditional diets with depression and anxiety in women. *American Journal of Psychiatry*, *167*(3), 305-311.

Jazieh, A.R., Foraida, M., Ghouse, M., Khalil, M.M., Kopp, M. and Savidge, M., 2006. The impact of cancer diagnosis on the lifestyle and habits of patients served at a Veterans Administration Hospital. *Journal of Cancer Education*, *21*(3), pp.147-150.

Jo, A., Coronel, B.D., Coakes, C.E. and Mainous III, A.G., 2019. Is there a benefit to patients using wearable devices such as Fitbit or Health Apps on Mobiles? A systematic review. *The American Journal of Medicine*, *132*(12), pp.1394-1400.

Kettlewell, N., 2020. Policy choice and product bundling in a complicatedhealth insurance market: do people get it right? J. Hum. Resour. 55(2), 566–610.

Khan, M.A.B., Hashim, M.J., King, J.K., Govender, R.D., Mustafa, H. and Al Kaabi, J., 2020. Epidemiology of Type 2 Diabetes–Global Burden of Disease and Forecasted Trends. *Journal of Epidemiology and Global Health*, *10*(1), p.107.

Kim, S.R. and Chern, W.S., 1999. Alternative Measures of Health Information and Demand for Fats and Oils in Japan. *Journal of Consumer Affairs*, 33: 92-109.

Kim, H.B., Lee, S.A. and Lim, W., 2019. Knowing is not half the battle: impacts of information from the National health screening program in Korea. *Journal of Health Economics*, *65*, pp.1-14.

Lasser, K., Boyd, J. W., Woolhandler, S., Himmelstein, D. U., McCormick, D., & Bor, D. H. (2000). Smoking and mental illness: a population-based prevalence study. *JAMA*, *284*(20), 2606-2610.

Lee, D.S., Card, D., 2008. Regression discontinuity inference with specification error. *Journal of Econometrics*, 142(2), pp. 655-674.

Lee, D.S., Lemieux, T., 2010. Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, 48(2), pp. 281-355.

Luppino, F. S., de Wit, L. M., Bouvy, P. F., Stijnen, T., Cuijpers, P., Penninx, B. W., & Zitman, F. G. (2010). Overweight, obesity, and depression: a systematic review and meta-analysis of longitudinal studies. *Archives of general psychiatry*, *67*(3), 220-229.

Mata, M., Cos, F.X., Morros, R., Diego, L., Barrot, J., Berengué, M., et al., 2013. Abordatge de la Diabetes Mellitus Tipus 2. 2a. Edició. Institut Català de la Salut 2013 (accessible at Abordatge de la diabetis mellitus tipus 2: guies de pràctica clínica).

Mezuk, B., Johnson-Lawrence, V., Lee, H., Rafferty, J. A., Abdou, C. M., Uzogara, E. E., & Jackson, J. S. (2013). Is ignorance bliss? Depression, antidepressants, and the diagnosis of prediabetes and type 2 diabetes. *Health Psychology*, *32*(3), 254.

McCrary, J., 2008. Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test. *Journal of Econometrics*, 142(2), pp. 698-714.

Pan, A., Lucas, M., Sun, Q., van Dam, R. M., Franco, O. H., Manson, J. E., ... & Hu, F. B. (2010). Bidirectional association between depression and type 2 diabetes mellitus in women. *Archives of Internal medicine*, *170*(21), 1884-1891.

Patel, M.S., Asch, D.A. and Volpp, K.G., 2015. Wearable devices as facilitators, not drivers, of health behavior change. *JAMA*, *313*(5), pp.459-460.

Robinson, D.J., Coons, M., Haensel, H., Vallis, M., Yale, J.F. 2018. Diabetes and mental health. *Canadian Journal of diabetes*, 42, Suppl. 1, S130-S141.

Roosen, J., Marette, S., Blanchemanche, S. and Verger, P., 2009. Does health information matter for modifying consumption? A field experiment measuring the impact of risk information on fish consumption. *Review of Agricultural Economics*, *31*(1), pp. 2-20.

Scott, D., & Happell, B. (2011). The high prevalence of poor physical health and unhealthy lifestyle behaviours in individuals with severe mental illness. *Issues in Mental Health Nursing*, *32*(9), 589-597.

Stathopoulou, G., Powers, M. B., Berry, A. C., Smits, J. A., & Otto, M. W. (2006). Exercise interventions for mental health: a quantitative and qualitative review. *Clinical psychology: Science and Practice*, *13*(2), 179.

Seuring, T., Serneels, P., Suhrcke, M., Bachmann, M., 2020. Diabetes, employment and behavioural risk factors in China: Marginal structural models versus fixed effects models. *Economics & Human Biology*, 39.

Siddiqui, M. A., Khan, M. F., & Carline, T. E. (2013). Gender differences in living with diabetes mellitus. *Materia Socio Medica*, *25*(2), 140-142.

Simon, G. E., Rohde, P., Ludman, E. J., Jeffery, R. W., Linde, J. A., Operskalski, B. H., & Arterburn, D. (2010). Association between change in depression and change in weight among women enrolled in weight loss treatment. *General Hospital Psychiatry*, *32*(6), 583-589.

World Health Organization (2021). Depression. Retrieved from Depression (who.int).

Yesavage, J. A. et al. (1982). Development and validation of a geriatric depression screening scale: a preliminary report. *Journal of Psychiatric Research*, 17, 37–49.

Zhao, M., Konishi, Y. and Glewwe, P., 2013. Does information on health status lead to a healthier lifestyle? Evidence from China on the effect of hypertension diagnosis on food consumption. *Journal of Health Economics*, *32*(2), 367-385.

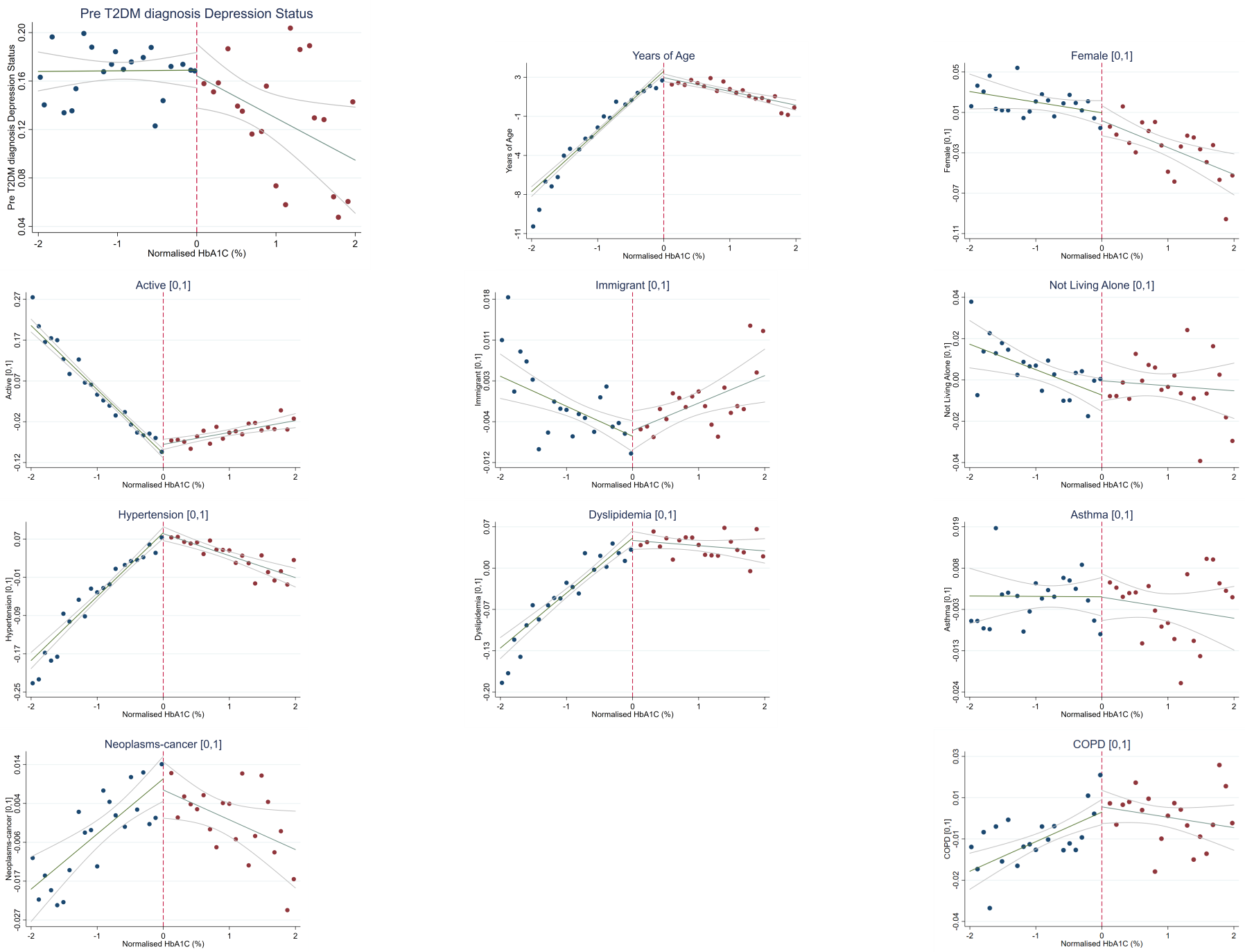
Tables and Figures

Table 1: Summary Statistics

	Mean	S.D.	Min	Max	Obs
Outcome Variable:					
Depression [0,1]	0.18	0.38	0	1	39688
T2DM Variables:					
T2DM Diagnosis [0,1]	0.67	0.47	0	1	39688
Onset of T2DM	3.13	3.72	0	39	34741
HbA1C (%)	6.60	1.43	0	20	39994
Demographics:					
Years of Age	65.10	12.62	16	106	39688
Female [0,1]	0.52	0.50	0	1	39688
Not Living Alone [0,1]	0.87	0.33	0	1	39688
Active [0,1]	0.27	0.44	0	1	39594
Immigrant [0,1l	0.02	0.13	0	1	39688
Other Conditions:					
Hypertension [0,1]	0.59	0.49	0	1	39688
Dyslipedimia [0,1]	0.53	0.50	0	1	39688
Asthma	0.04	0.21	0	1	39688
Neoplasms-cancer [0,1]	0.07	0.25	0	1	39688
COPD	0.05	0.23	0	1	39688
Observations	39994				

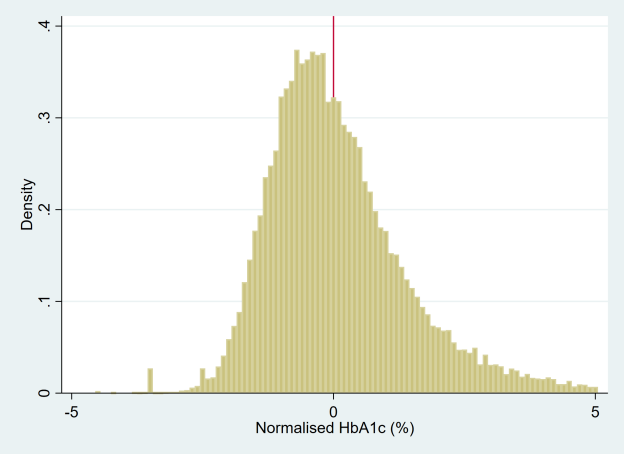
Note: The Table reports summary statistics of the main variables of interest.

Figure 1: Continuity Test



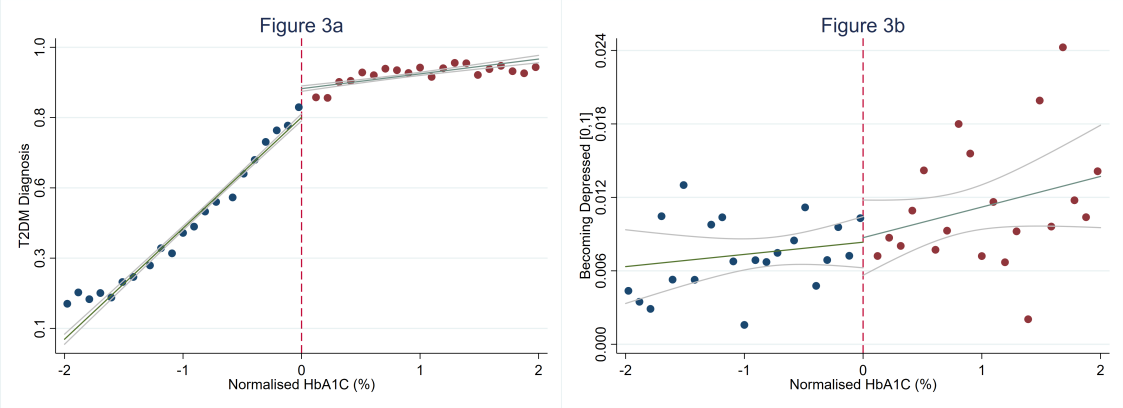
Note: The Figure shows local polynomial estimates of a number of covariates as a function of the running variable.

Figure 2: Density of the Running variable



Note: The Figure shows evidence of no manipulation of the running variable. Bin size = 0.1. The bin size has been selected by means of the McCrary test Stata routine, i.e. `DCdensity`.

Figure 3: RD Graphical Evidence



Note: Figure 3A shows local polynomial estimates of the probability of being diagnosed with T2DM as a function of the (centered) HbA1c, our first stage. Similarly, Figure 3b shows local polynomial estimates of the probability of being diagnosed with depression as a function of the (centered) HbA1c.

Table 2: RDD Estimates of a T2DM Diagnosis

	(1)	(2)	(3)	(4)	(5)
Above [0,1]	0.196***	0.193***	0.110***	0.094***	0.088***
	(0.032)	(0.031)	(0.025)	(0.025)	(0.025)
Above * HbA1c	-0.236***	-0.213***	-0.162***	-0.190***	-0.187***
	(0.025)	(0.024)	(0.017)	(0.019)	(0.019)
Running Variable:					
HbA1c (%)	0.247***	0.231***	0.163***	0.188***	0.185***
	(0.025)	(0.024)	(0.017)	(0.019)	(0.018)
Attributes:					
Years of Age		0.001**	-0.004***	-0.001***	-0.001***
		(0.000)	(0.000)	(0.000)	(0.000)
Female [0,1]		-0.024***	-0.043***	-0.044***	-0.044***
		(0.005)	(0.005)	(0.004)	(0.004)
Not Living Alone [0,1]		0.004	-0.021***	-0.016***	-0.018***
		(0.005)	(0.005)	(0.005)	(0.005)
Active [0,1]		-0.106***	-0.075***	-0.011**	-0.012**
		(0.009)	(0.007)	(0.005)	(0.005)
Immigrant [0,1l		-0.148***	-0.067***	-0.029*	-0.041***
		(0.020)	(0.018)	(0.015)	(0.015)
Onset of T2DM			0.066***	0.075***	0.075***
			(0.006)	(0.005)	(0.005)
Pre-existing Conditions:					
Hypertension [0,1]			0.084***	0.056***	0.054***
			(0.005)	(0.004)	(0.004)
Dyslipedimia [0,1]			0.054***	0.037***	0.038***
			(0.003)	(0.004)	(0.004)
Asthma			0.004	0.008	0.011*
			(0.007)	(0.007)	(0.007)
Neoplasies-cancer			0.020***	0.019***	0.017***
			(0.007)	(0.006)	(0.006)
COPD			-0.005	-0.016***	-0.019***
			(0.007)	(0.006)	(0.007)
Observations	39688	39594	34359	34356	34319

Note: The Table reports RDD estimates of a T2DM diagnosis. Although not shown in the Table, the presented estimates are conditional on time, area, and GP fixed effects. Robust standard errors are clustered on the running variable. ** $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.*

Table 3: Fuzzy RDD estimates of the Impact of a T2DM Diagnosis on Depression

	(1)	(2)	(3)	(4)	(5)
T2DM Diagnosis [0,1]	0.006**	0.011*	0.011*	0.016***	0.016***
	(0.003)	(0.006)	(0.006)	(0.005)	(0.005)
Running Variable:					
HbA1c (%)	0.000	0.001	0.001	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Attributes:					
Years of Age		-0.000	-0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)	(0.000)
Female [0,1]		0.006***	0.006***	0.006***	0.006***
		(0.001)	(0.001)	(0.001)	(0.001)
Not Living Alone [0,1]		0.002	0.002	0.002	0.002
		(0.002)	(0.002)	(0.002)	(0.002)
Active [0,1]		-0.004**	-0.004**	-0.001	-0.002
		(0.002)	(0.002)	(0.002)	(0.002)
Immigrant [0,1l		-0.006***	-0.006***	-0.004*	-0.004*
		(0.002)	(0.002)	(0.002)	(0.002)
Onset of T2DM		-0.001***	-0.001***	-0.001***	-0.001***
		(0.000)	(0.000)	(0.000)	(0.000)
Pre-existing Conditions:					
Hypertension [0,1]			-0.000	-0.002	-0.002*
			(0.001)	(0.001)	(0.001)
Dyslipedimia [0,1]			-0.001	-0.002	-0.002
			(0.001)	(0.001)	(0.001)
Asthma			-0.001	-0.001	-0.001
			(0.003)	(0.003)	(0.003)
Neoplasms-cancer			-0.003*	-0.003*	-0.004**
			(0.002)	(0.002)	(0.002)
COPD			0.000	0.001	0.001
			(0.002)	(0.002)	(0.002)
Time FE				✓	✓
Area FE				✓	✓
GP FE					✓
Observations	39688	34359	34359	34356	34319

Note: The Table reports RDD estimates of the impact of a T2DM diagnosis on depression. Although not shown in the Table, estimates are conditional on a set of covariates, time, areas and GP fixed effects. Robust standard errors are clustered on the running variable. ** p < 0.1, ** p < 0.05, *** p < 0.01.*

Table 4: Fuzzy RDD estimates of the Impact of a T2DM Diagnosis on Depression - Heterogeneity Analysis

	Full-Sample			Weight Loss			NO Weight Loss		
	(1) Full Sample	(2) Men	(3) Women	(4) Full Sample	(5) Men	(6) Women	(7) Full Sample	(8) Men	(9) Women
T2DM Diagnosis [0,1]	0.016***	-0.001	0.032***	-0.032	-0.083**	0.001	0.032**	0.015	0.058**
	(0.005)	(0.006)	(0.008)	(0.022)	(0.037)	(0.025)	(0.015)	(0.016)	(0.025)
Observations	34319	16483	17836	6990	3111	3879	6862	3242	3620

Note: The Table reports RDD estimates of the impact of a T2DM diagnosis on depression. Although not shown in the Table, estimates are conditional on a set of covariates, time, areas and GP fixed effects. Robust standard errors are clustered on the running variable. ** p < 0.1, ** p < 0.05, *** p < 0.01.*

Appendix

Table A.1: RDD Estimates of the Impact of a T2DM Diagnosis on Depression - Different Cut-offs

	(1) **Cut-off 4.5**	(2) **Cut-off 5**	(3) **Cut-off 5.5**	(4) **Cut-off 6**	(5) **Cut-off 6.5**	(6) **Cut-off 7**	(7) **Cut-off 7.5**	(8) **Cut-off 8**
T2DM Diagnosis [0,1]	0.015	0.015	0.002	0.014**	0.016***	0.017***	0.020***	0.015**
	(0.026)	(0.010)	(0.007)	(0.006)	(0.005)	(0.005)	(0.007)	(0.007)
Observations	34319	34319	34319	34319	34319	34319	34319	34319

Note: The Table reports RDD estimates on the outcomes of interest, when using a different cut-offs. The alternative cut-off value are 4.5, 5, 5.5, 6, 6.5, 7, 7.5, and 8. Although not shown in the Table, estimates are conditional on a set of covariates, time, areas and GP fixed effects (FE). Robust standard errors are clustered on the running variable. ** p < 0.1, ** p < 0.05, *** p < 0.01.*

Table A.2: RDD Estimates of the Impact of a T2DM Diagnosis on Depression - Different Polynomials

	(1) **Linear**	(2) **Quadratic**	(3) **Cubic**	(4) **Quartic**
T2DM Diagnosis [0,1]	0.016***	0.022***	0.024***	0.023***
	(0.005)	(0.005)	(0.006)	(0.006)
Observations	34319	34319	34319	34319

Note: The Table reports parametric RDD estimates when using different polynomials of the running variable. Although not shown in the Table, estimates are conditional on a set of covariates, time, areas and GP fixed effects (FE). Robust standard errors are clustered on the running variable. ** p < 0.1, ** p < 0.05, *** p < 0.01.*

Table A.3: Fuzzy RDD Estimates - Non-Parametric Approach

	(1) **Benchmark**	(2) **Bandwidth 2**	(3) **Bandwidth 1.75**	(4) **Bandwidth 1.5**	(5) **Bandwidth 1.25**	(6) **Bandwidth 1**
T2DM Diagnosis [0,1]	0.016***	0.008***	0.008***	0.007**	0.009***	0.011***
	(0.005)	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)
Observations	34319	35401	33626	31793	27938	24741

Note: The Table reports non-parametric RDD estimates of the impact of a T2DM diagnosis on lifestyle behaviours. Each coefficient in the table report the effect of being diagnosed with T2DM on lifestyle behaviours. Although not shown in the Table, estimates are conditional on a set of covariates, time, areas and GP fixed effects (FE). Robust standard errors are clustered on the running variable. ** p < 0.1, ** p < 0.05, *** p < 0.01.*