Received: 25 November 2022 | Revised: 5 December 2023 | Accepted: 5 January 2024

DOI: 10.1002/hec.4803

RESEARCH ARTICLE

Health Economics WILEY

The direct and spillover effects of diabetes diagnosis on lifestyle behaviours

Rhys Llewellyn Thomas[1] | **Emmanouil Mentzakis**[2]

[1]St Edmund Hall, University of Oxford, Oxford, UK

[2]Department of Economics, City, University of London, London, UK

Correspondence
Rhys Llewellyn Thomas, University of Oxford, St Edmund Hall, Oxford, UK.
Email: Rhys.Thomas@seh.ox.ac.uk

Funding information
Economic and Social Research Council

Abstract

Using blood sample data we exploit an arbitrary cut-off of diabetes risk and through a fuzzy regression kink design we estimate the effect of a diabetes diagnosis on own and partner health-related behaviours. Diabetes diagnosis increases the probability of exercising, both for those diagnosed with diabetes and their partner. We also conduct mediation analysis which suggests that joint household participation is the channel behind this effect. Our results have significant implications for the understanding of the channels that induce behavioural change, and household decision making, as well as, for the evaluation of diabetes related policies.

KEYWORDS
diabetes, health, household behaviour, regression kink design, spillover

JEL CLASSIFICATION
I12, I18, D1, D83

1 | INTRODUCTION

Priority setting in the budget constrained UK health care system rests upon principles of cost-effectiveness (NICE, 2012a; Wittenberg et al., 2019). However, while the accurate enumeration of technologies' costs is relatively straightforward, the respective calculation of benefits is not, as their full spectrum is often not known (Al-Janabi et al., 2016). This is particularly relevant for non-communicable diseases (NCDs) whose unaccounted spillover effects work through social influence and not pathogen transmission (Schwamm, 2018). Chronic NCDs account for 89% of deaths in the UK and given that many of these are preventable, early public health interventions to prevent these diseases are candidates for achieving substantial improvements in population outcomes (World Health Organization, 2018). To this effect screening programs can not only aid pre-clinical diagnosis and treatment but also raise awareness among households. Better awareness within households can, in principle, promote behavioural change (Fadlon & Nielsen, 2019; Fletcher & Marksteiner, 2017), hence making such screening programs more cost-effective than would be estimated if only direct effects were considered. Indeed, positive spousal correlations have been documented for behaviours such as smoking, alcohol, physical activity, and diet, which are all major modifiable NCD risk factors (Bove et al., 2003; Christakis & Fowler, 2008; Falba & Sindelar, 2008; Farrell & Shields, 2002; Macario & Sorensen, 1998). Hence, the strength of the link between NCD diagnosis, the resulting health behavioural change, and the

presence of spillovers onto other household members clearly determine the cost-effectiveness of condition management programmes.

In this paper, we investigate the effect of diabetes diagnosis on individual and partners' lifestyle behaviours, namely physical activity, diet, alcohol and smoking consumption. Crude diabetes statistics offer tentative evidence for the presence and potential benefit of spillover effects but these have yet to be causally estimated. More than 4.9 million people live with diabetes in the UK (850,000 people living with diabetes but yet to be diagnosed) with the risk of type 2 diabetes significantly raised when a close family member has diabetes. However lifestyle interventions are claimed to reduce such risk by about 50% (Diabetes UK, 2019). The focus on lifestyle behaviours is critical as they are well established risk factors of NCDs (Ezzati & Riboli, 2012, 2013; Willi et al., 2007), as well as constituting the first line of treatment for diabetes (WHO, 2016). Using blood sample data from the Health Survey for England (HSE) we exploit a seemingly arbitrary cut-off of diabetes risk and a fuzzy regression kink design (RKD) to estimate the effect of own diabetes diagnosis on own behaviour, as well as the effects of own diabetes status on partners' behaviour. Through mediation analysis we analyse whether partners are jointly partaking in lifestyle changes or whether they are independently changing their behaviours. Further, we explore heterogeneity in the effect by time-since-diagnosis, which in the absence of panel data, approximates long-term effects or recidivism to pre-diagnosis behaviours.

Several studies have examined the effect of a diabetes diagnosis on health behaviours. Using regression discontinuity designs (RDDs) Kim et al. (2019), Alalouf et al. (2023) and Iizuka et al. (2021) find limited behavioural change or improvement in outcomes in response to disease risk information or diabetes diagnosis, despite an increase in health care spending and utilization. We closely follow these works, in that we also analyse the impact of a diabetes diagnosis on physical activity, and like Iizuka et al. (2021) and Alalouf et al. (2023) we also analyse the impact on smoking and drinking behaviour. Our work is also similar to Oster (2018) and Hut and Oster (2022), who focus on dietary changes of newly diagnosed diabetes patients and find significant but small calorie reductions, which are concentrated in unhealthy foods, suggesting actual efforts to improve diet. Here, we analyse the impact on a narrower set of dietary outcomes. The current study differs from these works as we extend the analysis of the behavioural responses to diabetes diagnosis by estimating the spillover effect of a diagnosis on other household members. We make three key contributions. First, our findings speak to how behaviours are determined, and patients compliance with first line treatments for diabetes. Secondly, we contribute to the literature on marginally ill patients (Kim et al., 2019) which analyses relatively similar individuals in terms of health, and the impact marginally crossing a diagnosis threshold has. We also contribute to the new and growing literature on externalities of health shocks and their pathways. Finally, contrary to past household economics literature that focuses on assortative matching, we provide new evidence of correlated partners' health behaviours being attributed, at least partly, to social learning and joint household decision making.

Our analysis focuses on the impact of a diabetes diagnosis on risk-factors commonly associated with NCDs. Clinical recommendations regarding such risk-factors are clear and well-known to the general population, rendering a priori expectation of the effects straightforward. Increasing physical activity and vegetables consumption, and decreasing tobacco and alcohol consumption mitigate the risk of developing diabetes and are important first-line treatments of the disease (WHO, n.d.). Whereas, despite the established health benefits of fruit intake, recommendations on fruit consumption for patients with diabetes is ambiguous and possibly misunderstood by the general population[1] making priori expectations unclear.

Briefly, we find significant effects of diabetes diagnosis on own physical activity, while partners of individuals with a diabetes diagnosis also increase their physical activity. Spillover effects are mostly driven by partners' behaviour and less so by partners' diabetes status. We find no evidence of heterogeneity in either the effect on own or the spillover by time-since-diagnosis, which suggests persistence in the behavioural changes. All of our falsification tests support our identification strategy and provide evidence towards the robustness of the results. Our results are of particular importance to health policy makers, as the evidence of spillover effects suggests additional health benefits that are currently not accounted for in the evaluation of health policies.

2 | BACKGROUND

2.1 | Diabetes

The World Health Organization (WHO) defines diabetes as "a chronic, metabolic disease characterized by elevated levels of blood glucose (or blood sugar), which leads over time to serious damage to the heart, blood vessels, eyes,

kidneys, and nerves" (WHO, n.d.). Diabetes is classified into two types, type 1 and type 2. Of the 4.9 million people with diabetes in the UK, approximately 8% have type 1, which occurs when insulin production in the body is limited (Diabetes UK, 2019). Although there is limited understanding on its causes, diet and lifestyle are known not to impact the probability of developing type 1 diabetes. Type 2 diabetes affects approximately 90% of those with diabetes, and occurs when the body becomes resistant to insulin and is usually the result of poor diet and lifestyle (Helmrich et al., 1991; Hu et al., 2001).

Glycated haemoglobin (HbA1c) refers to the amount of haemoglobin (i.e., protein within red blood cells) which has been "glycated", which occurs when glucose in the blood attaches to haemoglobin proteins. The red blood cells which contain the haemoglobin proteins usually survive for between 8 and 12 weeks, and therefore HbA1c is considered to be an average blood sugar level over the previous 3 months. HbA1c is considered a useful measure in the diagnosis of diabetes, in that it provides an indication of blood sugar level over a longer duration.[2]

The World Health Organisation recommends a HbA1c of 6.5% as the cut-off point for diagnosing diabetes, while stating that values below 6.5% do not exclude a diabetes diagnosis (WHO, 2011). Levels below 6% are considered normal blood sugar levels while individuals with levels between 6% and 6.5% are considered to be at high risk of developing diabetes, also called pre-diabetes. While the link between HbA1c and the probability of developing diabetes is well-established, the choice of specific cut-off for diabetes and pre-diabetes are relatively arbitrary.[3] Nevertheless, although pre-diabetes usually has no symptoms, NICE[4] recommends that individuals with a HbA1c level between 6.0% and 6.4% should be offered a blood test at least once a year (NICE, 2012b).

Therefore, individuals who have been found to be pre-diabetic have a significantly higher probability of being diagnosed with diabetes due to the annual assessment of their HbA1c level. On the other hand, individuals just below the threshold of 6.0%, while having similar probability of actually having diabetes as those above the threshold, have a much lower probability of being diagnosed because they are not annually tested.

2.2 | Spousal correlation

A diabetes diagnosis transfers two types of health information to the patient: an update of their own health state (i.e., diagnosis of the disease) and information on the disease itself (i.e., causes and consequences of diabetes). In a shared environment, partners make decisions individually based on their preferences but are constrained by shared resources, exposed to common shocks, and share information sets by transferring information between each other (i.e., social learning) (Clark & Etilé, 2006). This gives rise to correlations in spousal behaviours.

Social learning implies that partners share similar information sets, and each partner individually updates their expectations of future risks and uncertainties. Whether this new information indeed promotes behavioural changes is dependent on idiosyncratic individual preferences, structural determinants of health, and their information set pre-diagnosis (Orphanides & Zervos, 1995). Hence, although information sets are shared, realised behaviours are not perfectly correlated. The magnitude of this effect is moderated by the information set prior to diagnosis, where the expectations of well informed individuals will not be substantially impacted by this new piece of information.

The theory of joint household production implies that households jointly produce commodities (goods, activities, capitals) that enter individuals' utility functions (Becker, 1981; Lancaster, 1966). Individuals within the household bargain to allocate resources to the production and consumption of shared commodities, implying a correlation both in behaviour and health. If a diabetes diagnosis changes the optimal consumption of health-related activities of the individual with diabetes, we should expect it to impact the production and consumption decisions of the other productive household members through joint household decision making (Becker, 1973, 1981). Payoffs from producing and subsequently consuming a particular good is a function of own private payoffs, and an externality from their partner consuming the same good. If behaviours or consumption goods are complements, then partners may choose to jointly produce and consume them, which we call joint participation. An obvious example of this is shared diets within households which, at least partly, reflect efficiency in jointly preparing meals.

Finally, assortative matching may also drive spousal correlation in behaviours. Assortative matching views partners' characteristics and preferences as complements, which drives individuals to match with partners they share preferences and characteristics with (Becker, 1973). Although we do not expect individuals to match based on their diabetes status, it is possible that individuals match based on behaviours which cause diabetes. For instance, individuals sharing a dislike for physical activity match in the marriage market. As a result these individuals are more likely to be diagnosed with diabetes.

Our study considers the spillover effect to be explained by the joint household production and social learning. We aim to decompose the spillover effect into the direct impact of the diabetes diagnosis on partners' behaviours, and the indirect impact where the diagnosed partners' induced change in behaviour causes a change is the other partners' behaviours, not the diagnosis itself. The direct effect is the effect caused by the diagnosis itself independent of the diagnosed partners' behavioural response. Presence of this type of spillover effect would suggest that each household member is making private decisions based on the new information set. The indirect effect suggests that partners are jointly partaking in behavioural changes as a result of the diagnosis. This channel does not exclude the possibility that informational transfer is driving the change, and does not exclude the possibility that partners are making private decisions based on the new information set. Instead, evidence of an indirect effect suggests that either household members are choosing to partake in these behavioural changes together, or that the private decisions are leading to the same behavioural outcomes. No evidence of a spillover would be found if partners are privately choosing different behaviours. While we believe that the indirect effect more closely resembles joint household production, whereby change in one partner's behaviour results in reallocation of household resources, we are unable to determine the precise channel in which this happens and acknowledge that there may be other explanations. Therefore we label this channel "joint participation" in behaviours, which better reflects the pathway we are estimating.

Only a handful of studies have explored externalities in the context of health. Fadlon and Nielsen (2019) analyse spillover effects on an extended network of individuals as a result of heart attacks. They find significant and persistent increases in statin consumption of the spouses, children, and co-workers of individuals who had a heart attack. Fletcher and Marksteiner (2017) use experimental data to estimate spillover effects of smoking cessation therapy program and alcoholism treatments. Results suggest significant impact in the behaviours of both partners and their experimental design can reasonably preclude a matching in the marriage market explanation. However, their results are at odds with the conclusions by Clark and Etilé (2006) who show that controlling for individual effects makes smoking behaviours statistically independent between partners. Clark and Etilé (2006) state that spousal correlation in smoking behaviour is the result of correlations in the individual effects, which they interpret as evidence of assortative matching. Finally, Janssen and Parslow (2021) conclude spillover effects exist within a household when looking at the impact of pregnancy on alcohol consumption.

3 | METHODS

3.1 | Data

This paper uses data from the Health Survey for England (HSE) years 2003–2015. HSE is an annual cross-sectional dataset which monitors trends in national health. Approximately 9000 addresses are sampled over the course of the calendar year. Within each household, all individuals are eligible for survey inclusion, however children under 15 years old are asked to complete a different survey. In addition to the individual questionnaire, all respondents are eligible for a nurse visit, in which individuals' physical measurements and a blood sample are taken.

The blood sample is sent to a specialist laboratory to measure, among others, HbA1c. Although 82.4% of individuals (across all years) agreed to be contacted for a nurse visit, only 34.7% of the full sample had samples taken for analysis.[5] Of the 56,245 individuals who had blood taken in the survey, 53,450 individuals had valid HbA1c measurements.[6] We exploit the latter in our econometric design.

We use the HSE Special Licence version which contains household identifiers and the self-reported relationship to each other person in the household. We define partners to be the household member that the respondent identifies as a spouse or cohabitee. In other words, j is defined as a partner to i if i states that j is either a husband/wife or partner/cohabitee. We are unable to access household identifiers in the surveys after 2015, hence we limit our analysis to the years 2003–2015.

In our analysis we use as a treatment variable information from the question "Do you now have, or have you ever had diabetes?" and exclude those that received their diagnosed before the age of 35 and were treated with insulin. We are unable to determine whether the individual has Type 1 or Type 2, however those diagnosed before 35 and treated with insulin were likely to have Type 1 diabetes and therefore we remove them from the analysis.

3.1.1 | Target outcomes: Stated behaviours

Our outcomes (physical activity, diet, tobacco and alcohol) are behaviours that have all been shown to cause diabetes, and are the first line treatment for managing and treating the condition (WHO, 2016). Exercise is a binary response to "did any exercise in the last 4 weeks". Dietary information in the HSE is limited, hence we rely on two relevant variables: binary responses to "any vegetables eaten yesterday?" and "any fruit eaten yesterday?".[7] For drinking, a value of 1 is assigned if individual responded positively to either "whether drinks nowadays" or "whether drinks occasionally", and 0 if stated that they do not currently drink. For smoking, the binary variable takes value of 1 if respondent stated that they are "currently cigarette smoker" and takes value 0 if they stated that they have "never smoked cigarettes at all", they "used to smoke cigarettes occasionally", or "used to smoke cigarettes regularly".

Table 1 presents descriptive statistics of diabetes status, stated behaviours, and observable characteristics for the HSE sample. Further details on these variables and a description of Table 1 are available in Appendix Section A1.

3.2 | Econometric specifications

We aim to estimate the causal impact of own or partners' diabetes diagnosis on lifestyle behaviours. This relationship is described by:

$$Y_i = \theta_0 + \theta_1 EverD_i + \theta_2 EverD_j + e_i \quad (1)$$

where Y_i denotes the health-related lifestyle behaviour, $EverD_i$ denotes whether individual i has ever been diagnosed with diabetes, and $EverD_j$ denotes whether the partner of individual i, person j, has ever been diagnosed with diabetes. An OLS of this form, using survey data, would provide biased estimates of both θ_1 and θ_2.

The most salient source of bias is simultaneity. Individuals with diabetes may behave in a way more damaging to their health than those without diabetes. Such correlation ignores that these individuals would have been diagnosed as having diabetes because they behaved in this way. Indeed, the causes of type 2 diabetes are poor lifestyle factors (Helmrich et al., 1991; Hu et al., 2001). Lifestyle is often determined by environmental and socio-economic factors, and these factors are also likely to influence the probability of a diabetes diagnosis, therefore this bias could also be caused by these omitted confounders. The second source of endogeneity is matching in the marriage market (Dupuy & Galichon, 2014). Individuals selectively marry along similar traits and therefore ignoring this channel through a naive estimation would bias estimates of the spillover effect.

3.2.1 | Regression kink design

To identify the effect of diabetes diagnosis on health-related behaviours we utilise a regression kink design (RKD), where we exploit a slope change in the probability of a binary treatment variable. Figure 1 shows an increasing but low probability of ever being diagnosed with diabetes when plotted against HbA1c, until the kink point of 6%.[8] At 6% there is a sharp increase in the gradient of the probability of being diagnosed. As discussed in Section 2.1, NHS recommends that individuals with a HbA1c level above 6% are offered annual blood tests to monitor their blood sugar levels, and to diagnose diabetes as early as possible. The initial test could be for a variety of reasons, sometimes as part of a regular checkup offered by the NHS, or if an individual shows symptoms that warrant a blood test. Such precise kink in the probability of a diabetes diagnosis is not supported in the medical literature. Yudkin and Montori (2014) state that an inflection point of diabetes risk does not exist, meaning that the assignment of diabetes risk is arbitrary. We will use this threshold of 6% as an exogenous threshold to identify the effect of diabetes diagnosis on behaviour.

Dong (2011) provides the theoretical framework for identification, whereby the RKD identifies the causal effect of a binary treatment when there is no discontinuity in the probability of treatment but rather a kink. When the policy rule is implemented with some error (i.e., the kink is not deterministic) a fuzzy RKD design can be implemented (Card et al., 2015). A fuzzy RKD combines the RKD with a two-stage least squares (2SLS) specification. The first stage identifies the effect of the kink on the probability of treatment:

TABLE 1 Descriptive statistics.

		Blood sample			Blood and partner sample		
	HSE adult sample	**All**	**Below kink**	**Above kink**	**All**	**Below kink**	**Above kink**
Observable characteristics							
Age	49.76	51.31	48.91	63.98	51.90	50.03	62.26
	(18.75)	*(17.54)*	*(17.19)*	*(13.47)*	*(15.16)*	*(14.83)*	*(14.83)*
Males	0.44	0.46	0.45	0.49	0.48	0.47	0.56
	(0.50)	*(0.50)*	*(0.50)*	*(0.50)*	*(0.50)*	*(0.50)*	*(0.50)*
Any Qualifications	0.75	0.77	0.80	0.59	0.79	0.82	0.63
	(0.44)	*(0.42)*	*(0.40)*	*(0.49)*	*(0.41)*	*(0.39)*	*(0.48)*
Degree level education	0.21	0.22	0.24	0.13	0.24	0.26	0.15
	(0.41)	*(0.42)*	*(0.43)*	*(0.36)*	*(0.43)*	*(0.44)*	*(0.36)*
Partner living in household	0.64	0.67	0.67	0.65	–	–	–
	(0.48)	*(0.47)*	*(0.47)*	*(0.47)*			
Household size[a]	2.70	2.62	2.70	2.17	2.92	2.98	2.58
	(1.41)	*(1.33)*	*(1.34)*	*(1.15)*	*(1.18)*	*(1.19)*	*(1.06)*
Employed[a]	0.60	0.61	0.66	0.37	0.67	0.71	0.44
	(0.49)	*(0.49)*	*(0.47)*	*(0.48)*	*(0.47)*	*(0.45)*	*(0.50)*
Equivalised income	30,502.16	31,879.31	32,889.45	26,346.88	33,769.27	34,846.97	27,528.68
	(27,727.47)	*(28,152.77)*	*(28,445.57)*	*(25,801.81)*	*(26,581.32)*	*(26,818.7)*	*(24,241.46)*
Self-assessed general health	2.05	1.98	1.89	2.43	1.93	1.85	2.36
(1 = Very good, 5 = Very poor)	*(0.95)*	*(0.91)*	*(0.87)*	*(0.99)*	*(0.87)*	*(0.83)*	*(0.96*
Glycated hemoglobin (HbA1c)	–	5.59	5.38	6.71	5.59	5.39	6.71
		(0.73)	*(0.33)*	*(1.14)*	*(0.72)*	*(0.32)*	*(1.14)*
Ever diagnosed with diabetes	0.063	0.056	0.011	0.296	0.053	0.011	0.292
	(0.24)	*(0.229)*	*(0.103)*	*(0.457)*	*(0.224)*	*(0.103)*	*(0.455)*
Stated behaviours							
Physical activity[a]	0.45	0.47	0.50	0.26	0.46	0.49	0.27
	(0.50)	*(0.50)*	*(0.50)*	*(0.44)*	*(0.50)*	*(0.50)*	*(0.45)*
Vegetable consumption	0.51	0.52	0.52	0.53	0.53	0.53	0.55
	(0.50)	*(0.50)*	*(0.50)*	*(0.50)*	*(0.50)*	*(0.50)*	*(0.50)*
Fruit consumption	0.59	0.61	0.60	0.65	0.63	0.62	0.66
	(0.49)	*(0.49)*	*(0.49)*	*(0.48)*	*(0.48)*	*(0.48)*	*(0.47)*
Currently a drinker	0.84	0.88	0.90	0.82	0.90	0.91	0.84
	(0.36)	*(0.32)*	*(0.31)*	*(0.38)*	*(0.31)*	*(0.29)*	*(0.37)*
Currently a smoker	0.20	0.19	0.19	0.17	0.16	0.15	0.16
	(0.40)	*(0.39)*	*(0.39)*	*(0.38)*	*(0.36)*	*(0.36)*	*(0.37)*
Ever a drinker	0.90	0.93	0.94	0.90	0.94	0.94	0.91
	(0.30)	*(0.26)*	*(0.25)*	*(0.30)*	*(0.25)*	*(0.24)*	*(0.29)*

TABLE 1 (Continued)

		Blood sample			Blood and partner sample		
	HSE adult sample	All	Below kink	Above kink	All	Below kink	Above kink
Ever a smoker	0.57	0.59	0.58	0.62	0.58	0.57	0.62
	(0.49)	*(0.49)*	*(0.49)*	*(0.48)*	*(0.49)*	*(0.50)*	*(0.49)*
Obs.	92,436	45,063	37,901	7162	30,198	25,565	4633

Note: Table shows the mean, and the standard deviation in parentheses of observable characteristics and stated behaviours. The HSE adult sample column shows the descriptive statistics for the entire Health Survey for England sample. The blood sample column shows only the sub-sample of individuals whom we have valid HbA1c measurements for. Blood and Partner sample represents the sub-sample of individuals who had both valid HbA1c measurements and that we were able to identify partners in the Health Survey for England. Below kink columns represent the sub-sample of individuals with HbA1c levels below 6.0%, and above kink columns represent the sun-sample of individuals with HbA1c levels above 6.0%.

[a]Denotes variables which were not available to us for all years of the survey, and therefore the true number of observations used to calculate them are less than the number of observations denoted at the bottom of the table.

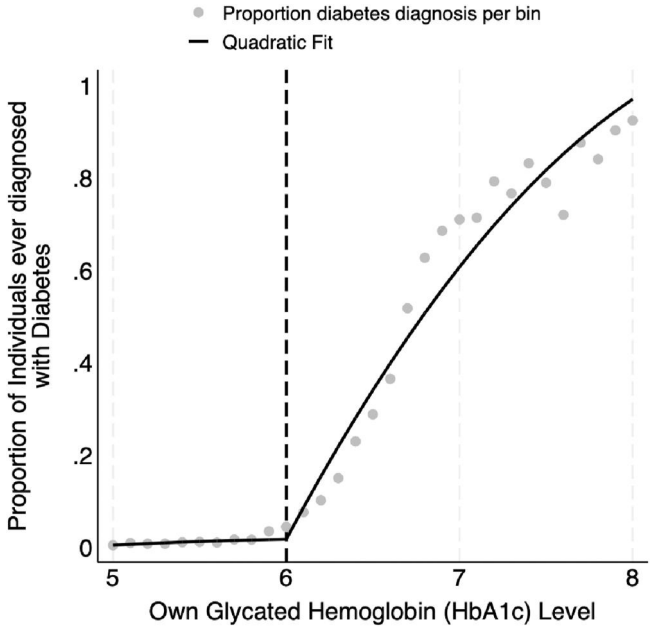

FIGURE 1 Probability of Diabetes Diagnosis by HbA1c Level. Mean of the probability of ever being diagnosed with diabetes per bin. Bin width of 0.1 for glycated hemoglobin levels between 5 and 8. Quadratic fit (solid line) is separately estimated for the left and right hand sides of the kink. Dashed line represents the kink point, where glycated hemoglobin is a value of 6.0.

$$EverD_i = \gamma_0 + \gamma_1(x_i - k)D_i + \left[\sum_{p=1}^{p^*} \nu_p^-(x_i - k)^p\right] + \left[\sum_{p=2}^{p^*} \nu_p^+(x_i - k)^p D_i\right] + \xi_i \quad (2)$$

where $EverD_i$ is a binary variable taking the value of one if i has have ever been diagnosed with diabetes, and zero otherwise. x_i denotes the running variable (HbA1c), and k is the kink point of 6%. $D_i = \mathbb{1}(x_i \geq k)$, takes a value of one if the individual's level of HbA1c is above the kink point, and where $(x_i - k)D_i$ is the excluded instrument for the fuzzy RKD. p^* denotes the highest order of polynomial used, ν_p^- and ν_p^+ are the estimates of the polynomial function below and above the kink point, respectively.

We then estimate the following second stage regression where the kink is used as an instrument for diabetes status:

$$Y_i = \beta_0 + \beta_1 \widehat{EverD}_i + \left[\sum_{p=1}^{p^*} \alpha_p^-(x_i - k)^p\right] + \left[\sum_{p=2}^{p^*} \alpha_p^+(x_i - k)^p D_i\right] + \epsilon_i \quad (3)$$

Y_i denotes the outcome of interest. $\widehat{EverD}_i$ is the predicted probability, from the first stage, of ever being diagnosed with diabetes, while again the terms in the square brackets denote the polynomial function. Standard errors are clustered at

the household level. Under the assumptions outlined by Dong (2011) and Card et al. (2015) (see Section 3.2.2), the coefficient β_1 can be interpreted as the unbiased Marginal Treatment Effect (MTE) of ever having been diagnosed with diabetes. This specification is estimated by 2SLS with standard errors clustered at the household level.

As with RDDs there is a bias-variance trade-off to be made when selecting the estimation sample. A narrow bandwidth will reduce the chances of misspecification error given that around the kink-point the functional form is likely to be closer to linear. Narrower bandwidths also means that the random assortment around the kink-point assumption is more likely to be valid. Narrower bandwidths would include individuals who are more similar in terms of observable and unobservable characteristics, reducing the risk of bias in our estimates. However smaller samples lower power, therefore may not reject a false null hypothesis because of the larger variance in the estimates. Large samples improve precision of the estimates but also increase the chances of misspecifying the functional form, and reduces the likelihood that individuals either side of the cut-off are similar in terms of characteristics, increasing the risk of bias (Cattaneo et al., 2020). HSE includes HbA1c measurements to one decimal place, meaning we have data which looks discrete around the cut-off. Therefore, because we favour a narrow bandwidth for our main estimates, we limit our specification to a linear polynomial to avoid over-fitting. A visual inspection of Figure 1 shows that the true data generating process is close to linear in the narrow interval around the cut-off.

To ensure the robustness of our estimates, we present a number of alternative specifications and bandwidths in sensitivity tests. Given the fewer number of observations on the right hand side of the kink-point we increase that bandwidth and keep the left-hand side bandwidth much narrower where small sample size is less of a problem. For the effect on own behaviour we use a bandwidth of 0.5% on the right hand side of the cut-off and 0.3% on the left hand side (i.e., HbA1c between 5.7% and 6.5%).[9]

3.2.2 | Identification conditions

Identification places three conditions on the running variable: no precise manipulation around the threshold, relevance, and monotonicity. Identification does not require HbA1c to be exogenous. Indeed RKD works precisely because the treatment (i.e., diabetes diagnosis) is endogenous and correlated with the running variable and unobservable characteristics (Dong, 2011).

Lack of precise manipulation of HbA1c implies that individuals do not selectively sort around the threshold. Individuals may coarsely know and influence their blood glucose level but cannot precisely control their HbA1c. Practically, if HbA1c was highly responsive to lifestyle behaviours, then targeting a particular HbA1c level would be impossible because of its high variability. Empirically, easy manipulation of HbA1c would result in greater density on one side of the cut-off and significant HbA1c variation over time. In Appendix Section A2.1 we show that there is not a large mass in the density distribution, while HbA1c transition matrices in Appendix Section A3.2 shows that from initial measurement to follow-up, individuals' HbA1c does vary but appears limited with 80% of individuals falling within two decimal places of their initial measurement.

Relevance requires that our contemporaneously measured instrument (i.e., the kink) is predictive of our treatment variable (i.e., ever diagnosed with diabetes) which may have been sometime in the past, while monotonicity in HbA1c rules out defiers. For these to hold, HbA1c levels need to be sufficiently time-persistent to make contemporaneous HbA1c, or more specifically the kink in HbA1c, predictive of diabetes diagnosis, while movements in the running variable need to be largely monotonic. Empirically we find support for both conditions. HbA1c levels exhibit strong persistence, especially when considering HbA1c categories (see Appendix Section A3.2). We may expect downward trends towards "normal" HbA1c levels, however instead we observe upward monotonic trends, which persists even when we split sample by diabetes status. These findings are supported by the medical literature which confirms that even intensive diabetes treatments lead to modest changes in HbA1c (Andrews et al., 2011; Diabetes Prevention Program Research Group, 2002).[10]

One way in which relevance and monotonicity could be invalidated is if individuals' diabetes treatment was successful enough to reduced their HbA1c levels to below the cut-off. These individuals would be defiers, as past diagnosis results in an increased probability of being below 6.0% HbA1c, which would also make contemporaneous HbA1c a poor predictor of diagnosis. While this argument is theoretically possible, it is not consistent with our data. Figure 1 shows that the proportion of individuals with diabetes to the left of the cut-off is almost zero for each HbA1c bin, just 1% of individuals to the left of the kink have ever been diagnosed with diabetes.[11]

One concern with identification is that nicotine is a metabolic stimulant and appetite suppressant (Pinkowish, 1999), which would invalidate the monotonicity assumption. Nicotine intake may be confounding the relationship between diabetes diagnosis and behaviours. In other words, quitting smoking may lead to increased consumption of unhealthy foods and therefore increasing the chances of a diabetes diagnosis. Indeed, evidence suggests that quitting smoking leads to weight gain (Courtemanche et al., 2018). However, there is also medical evidence of smoking increasing insulin resistance thereby increasing the chances of developing diabetes (Bergman et al., 2012; Chang, 2012; Chiolero et al., 2008; Facchini et al., 1992; Mikhailidis et al., 1998; Śliwińska Mosson & Milnerowicz, 2017), and evidence that smoking cessation improves insulin sensitivity (Bergman et al., 2012; Eliasson, 2003). Therefore, it is unclear whether quitting smoking increases the risk of developing diabetes (Bajaj, 2012; Yeh et al., 2010). In Appendix Section A11 we confirm that our results are robust to excluding smokers that quit before their diagnosis to deal with this concern.

Finally, to identify the MTE of diabetes diagnosis, two observable implications must hold (Card et al., 2015). The first is the smooth density of the assignment variable and tests the assumption of no deterministic sorting. The second is no discountinuity or kinks in the pre-determined covariates and tests the assumption that the marginal effect of the assignment variable on the outcome is smooth. Both assumption are validated in Appendix Section A2.

3.3 | Identification of spillover effects

To handle the endogeneity in the effect of partner's diabetes diagnosis on own behaviour, we use partners' kink as an instrument for partners' diabetes status. The first stage of the 2SLS is:

$$EverD_j = \lambda_0 + \lambda_1 (x_j - k) D_j + \left[\sum_{p=1}^{p^*} \rho_p^- (x_j - k)^p \right] + \left[\sum_{p=2}^{p^*} \rho_p^+ (x_j - k)^p D_j \right] + u_i \tag{4}$$

where j denotes the partner, $EverD_j$ is partners' diabetes status, and x_j denotes the partners' HbA1c level. The second stage is:

$$Y_i = \delta_0 + \delta_1 \widehat{EverD_j} + \left[\sum_{p=1}^{p^*} \tau_p^- (x_j - k)^p \right] + \left[\sum_{p=2}^{p^*} \tau_p^+ (x_j - k)^p D_j \right] + \varepsilon_i \tag{5}$$

Y_i denotes the behavioural outcome. $\widehat{EverD_j}$ is the predicted probability, from the first stage, of partner ever being diagnosed with diabetes. The terms in the square brackets denote the polynomial function below and above the kink point. The estimating sample are those who have partners, and those partners having HbA1c levels within the bandwidths. Therefore, for the spillover effect we increase the bandwidths to 0.4% on the left-hand side and 0.9% on the right-hand side, this makes the own and spillover estimation samples similar in size. Once again, this specification is estimated by 2SLS and standard errors are clustered by household.

We interpret results from the above specification as spillover effects. We do not believe that the assortative matching channel explains the spillover as it would require individuals to be aware enough of their own HbA1c level at the time of matching, and to selectively match based on being either side of the arbitrary kink-point. It is reasonable to assume that individuals match based on their relative position in the HbA1c distribution, or some unobservable variable correlated with HbA1c, but doing so does not violate the identifying assumption. To further assure ourselves that assortative matching does not explain our results, a robustness check is run where we estimate a specification including both own and partners' diabetes status using the same instruments. If the spillover was the result of partners' diagnosis increasing the probability of own diagnosis then we would find no evidence of a spillover effect if we were to control for own diagnosis. This specification (presented in Appendix Section A6) would allow us to confirm that our results are a true spillover effect, as opposed to assortative matching or due to an increase in probability of own diabetes diagnosis. Without tracking individuals across time and observing their matching decisions we are unable to claim with certainty that our estimates are not explained by assortative matching.

A further nuance is that the cross-sectional nature of our data might lead us to make inferences with only couples that did not separate after diagnosis. If this were the case we might be overestimating the true effect because couples that did not separate are ones that are willing to jointly participate in behavioural changes. Again, we check this in Appendix Section A2.2.1.

4 | RESULTS

4.1 | Effect of own diagnosis

Appendix Figure A6 shows the reduced form estimates graphically.[12] Physical activity outcome is the only reduced form that has a statistically significant kink. This graph shows a decreasing trend across the entire range, as one may have expected. The probability displays a kink at 6% where the magnitude of the slope decreases, implying that there is a positive impact on physical activity. All the other behaviours have statistically insignificant slope changes.

The first row of Table 2 presents OLS estimates of the impact of ever being diagnosed with diabetes, while fuzzy RKD estimates of coefficient β_1 from Equation (3) are given below. The relevance of the kink as an instrument is given in the first stage coefficients in Appendix Table A6 with results suggesting statistically significant positive effect of the kink on diabetes status. Romano-Wolf multiple hypothesis correction *p*-values are presented in square brackets. OLS estimates show that diabetes diagnosis is associated with a decrease in the probability of physical activity by 22% points (p.p.), smoking by 5.7 p.p. and drinking by 11.7 p.p. and an increase in fruit consumption by 8.6 p.p. and vegetable consumption by 1.8 p.p. RKD estimates show that a diabetes diagnosis increases the probability of physical activity by 84 p.p. We find no evidence of an impact on consumption of fruit or vegetables, or evidence of changes to drinking or smoking behaviour.

The OLS estimates show a strong correlation between diabetes status and behaviours, whereas there are far fewer significant effects and a change in sign for physical activity for the RKD estimator. We do not consider the OLS estimates causal, because individuals that received a diabetes diagnosis would have been diagnosed because of their poor lifestyle choices. Whereas our RKD estimation strategy aims to remove the reverse causality from the estimates, and instead reflects the causal effect of receiving a diabetes diagnosis on these behaviours.

Aside from differences in signs and significance of effects, RKD estimates are much larger in magnitude. Given that our measure of physical activity is "Whether did any exercise in the last 4 weeks", the large magnitudes do not necessarily imply wholesale changes in lifestyle but a more moderate change in which those diagnosed with diabetes now do some physical activity, whereas before they may have done very little or none at all.

TABLE 2 Fuzzy RKD estimates of change in own behaviour as a result own diabetes diagnosis.

	Exercise	Vegetable consumption	Fruit consumption	Currently a smoker	Currently a drinker
OLS estimates					
Effect of own diabetes	−0.121***	0.0297	0.0568***	−0.0540***	−0.0530***
	(0.0275)	(0.0199)	(0.0175)	(0.0135)	(0.0154)
Romano and wolf *p*-values	[0.000]	[0.126]	[0.002]	[0.000]	[0.000]
Obs.	4385	11,089	11,098	13,148	13,150
RKD estimates					
Effect of own diabetes	0.836**	0.148	0.00148	−0.213	0.0131
	(0.389)	(0.184)	(0.169)	(0.149)	(0.138)
Romano and wolf *p*-values	[0.084]	[0.816]	[1.000]	[0.518]	[1.000]
First stage $F-Statistic$	21.24	80.65	80.64	85.02	84.37
Obs.	4385	11,089	11,098	13,148	13,150

Note: RKD coefficients are estimated using a linear specification each side of the kink point, equally weighted observations, and with bounds of 0.5 on the right hand tail, and 0.3 on the left hand side. Parentheses includes cluster-robust standard errors, clustered at the household level. Each specification include year fixed-effects. OLS coefficients estimated using equation $Y_i = \theta_0 + \theta_1 EverD_i + \theta_2 W_i + e_i$, using the same sample as the RKD estimating sample, and the same year dummies W_i as the RKD estimates. Square brackets include the Romano-Wolf multiple hypothesis correction *p*-values.

*** denotes *p*-value of 0.01 or less, ** denotes *p*-value of 0.05 or less, * denotes *p*-value of 0.10 or less.

4.2 | Spillover effect

Appendix Figures A7 shows the educed form effect of partners' Hb1Ac level graphically.[13] Once again physical activity is the only statistically significant slope change at the 10% level. The shape of the curve is similar to the effect of own Hb1Ac level, except that it is approximately 60% of the magnitude.

First stage estimates are presented in Appendix Table A7, again providing evidence of its relevance as an instrument. The spillover estimates that is, parameter δ_1 in Equation (5), are presented in Table 3 below the OLS estimates in the first row, with findings suggesting very similar patterns to those of own diabetes diagnosis. Romano-Wolf multiple hypothesis *p*-values are shown in square brackets. We find significant positive effects on physical activity. The magnitude of these effects are smaller than those estimated for the direct effect, which we may have expected a priori.

Comparing OLS and RKD estimates for physical activity, we find that the direction of the effect changes. As discussed in Section 2.2, having a partner with a diabetes diagnosis may be negatively correlated with physical activity due to assortative matching. Partners that do not participate in physical activity are both more likely to be diagnosed with diabetes and match with partners that do not participate in physical activity. OLS estimates will capture this effect, whereas our RKD approach aims to remove this from our estimates. RKD estimates show that a partner receiving a diabetes diagnosis increases the probability of participating in physical activity by 39% points, which is approximately 47% the magnitude of the effect on own.

4.3 | Heterogeneity

In Appendix Section A9 we explore effect heterogeneity by: whether living with a partner, education, sex, and time-since-diagnosis. We explore heterogeneities for the effect of own diagnosis and the spillover effect. Here we only interpret the coefficients for exercise, given that it is the only outcome we find significant effects for. By whether partner lives in the household and sex we find that there is no heterogeneity. There is no evidence that the effect on own behaviour is heterogeneous by education, however the spillover effect seems to be driven by those with degree-level education. We find that the "ever diagnosed with diabetes" coefficient is not statistically significant while the ever diagnosed with diabetes - degree-level interaction is. We find no evidence that there is any heterogeneity by time-since-diagnosis, either for the effect on own or the spillover effect. This suggests a habit formation type of behaviour, where

TABLE 3 Fuzzy RKD estimates of change in own behaviour as a result of partner's diabetes diagnosis.

	Exercise	Vegetable consumption	Fruit consumption	Currently a smoker	Currently a drinker
OLS estimates					
Effect of Partner's diabetess	−0.132***	0.00131	0.00926	−0.0187	−0.0431***
	(0.0293)	(0.0212)	(0.0193)	(0.0130)	(0.0155)
Romano and wolf *p*-values	[0.000]	[0.960]	[0.876]	[0.452]	[0.012]
Obs.	3451	8202	8203	9584	9585
RKD estimates					
Partner's diabetes	0.433**	0.0375	−0.0150	0.0437	0.0709
	(0.187)	(0.0990)	(0.0910)	(0.0663)	(0.0710)
Romano and wolf *p*-values	[0.088]	[0.892]	[0.892]	[0.864]	[0.740]
First stage $F-Statistic$	65.26	239.39	239.43	254.48	254.47
Obs.	3451	8202	8203	9584	9585

Note: RKD coefficients are estimated using a linear specification each side of the kink point, equally weighted observations, and with bounds of 0.9 on the right hand tail, and 0.4 on the left hand side. Parentheses includes cluster-robust standard errors, clustered at the household level. Each specification include year fixed-effects. OLS coefficients estimated using equation $Y_i = \theta_0 + \theta_1 EverD_j + \theta_2 W_i + e_i$, using the same sample as the RKD estimating sample, and the same year dummies controls (W_j) as the RKD estimates. Square brackets include the Romano-Wolf multiple hypothesis correction *p*-values.

*** denotes *p*-value of 0.01 or less, ** denotes *p*-value of 0.05 or less, * denotes *p*-value of 0.10 or less.

individuals change their behaviour immediately after a diagnosis, and this new equilibrium is persistently maintained over the long-term. Our data does not allow us to analyse the short-term dynamics of the behavioural changes due to the diagnosis, however this result suggests that the diagnosis induces a change in behaviours upon diagnosis and this is maintained over the long-term. Further details and discussion of the heterogeneity analysis is available in Appendix Section A9.

4.4 | Robustness checks

In the appendix of this paper we include a number of additional results and tests to check the robustness of our results, the location are provided in the parentheses. We check the sensitivity of our results to alternative bandwidths and polynomials (A8), estimate own and partners' diabetes status simultaneously (A6), explore placebo and predetermined outcomes (A7 and A2.2.1), replicate our results using English Longitudinal Study for Ageing data (A3), estimate our results using weights to match the blood sample to the entire HSE sample (A10), explore the possibility of reverse causality in our smoking estimates (A11), and test for the location of the kink point (A12). These results confirm the findings from the main text and validate the robustness of our results.

5 | CAUSAL PATHWAYS OF THE SPILLOVER EFFECTS

As discussed in Section 2.2, the correlation between spouses can theoretically be attributed to assortative matching, shared environment and joint household decision making. Our identification strategy allows us to plausibly exclude attributing spillover effects to assortative matching. Therefore, the results we present in Table 3 are the combined effect of the latter two pathways and we seek to decompose the spillover effect into two pathways: changes in own behaviour that are the result of partner's diagnosis (i.e., direct effect), and changes in own behaviour that are the result of the induced change in partner's behaviours (i.e., indirect effect), see Figure 2 for illustration.

Given the endogenous nature of Y_i and $EverD_i$ and the presence of only a single instrument (i.e., kink in the fuzzy RKD), we follow the framework outlined by Dippel et al. (2020) which requires the additional assumption that the confounding variable that jointly affects $EverD_i$ and Y_i is independent of the confounding variable that jointly causes Y_i and Y_j.

Four equations are estimated:

$$EverD_i = \beta_T^Z (x_i - k)D_i + f(x_i - k) + \epsilon^T \quad (6)$$

$$Y_i = \beta_M^T \widehat{EverD}_i + f(x_i - k) + \epsilon^M \quad (7)$$

$$Y_i = \gamma_M^Z EverD_i + \gamma_M^T (x_i - k)D_i + f(x_i - k) + \xi^M \quad (8)$$

$$Y_j = \beta_Y^M \widehat{Y}_i + \beta_Y^T EverD_i + f(x_i - k) + \epsilon^Y \quad (9)$$

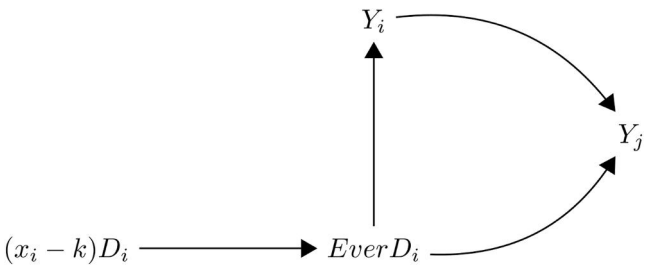

FIGURE 2 Causal Pathway of the spillover effect. $(x_i - k)D_i$ denotes the kink, which we use as the instrument in the fuzzy RKD specification. $EverD_i$ is the diabetes status of individual i. Y_i is the health-related behaviour of individual i, and Y_j is the health-related behaviour of individual j. The pathway $EverD_i \rightarrow Y_j$ is considered to be the direct effect of individual i's diabetes diagnosis on the behaviours of individual j. The pathway $EverD_i \rightarrow Y_i \rightarrow Y_j$ is the indirect effect, where the diagnosis of i causes a change in j's behaviours which is the result of the induced change in i's behaviours. In other words, the effect of the diagnosis $EverD_i$ on Y_j, through the mediator Y_i.

$EverD_i$ is whether i has ever been diagnosed with diabetes, $\widehat{EverD_i}$ is the predicted probability from Equation (6). Y_i denotes the behavioural outcome of interest, $\widehat{Y_i}$ is the predicted equivalent from 8. x_i is HbA1c-level, and k is the kink-point (6%). $D_i = \mathbb{1}(x_i \geq k)$ takes value one if the individual's level of HbA1c is above the kink point. $f(x)$ represents the polynomial function used throughout the analysis in this paper: $\left[\sum_{p=1}^{p^*} \nu_p^-(x_i - k)^p\right] + \left[\sum_{p=2}^{p^*} \nu_p^+(x_i - k)^p D_i\right]$. β_Y^T is the direct effect, and the indirect effect is $\beta_M^T \times \beta_Y^M$. Equations (6) and (7) are the same specifications as Equations (2) and (3).

The system of equations is estimated using 2SLS estimators (Dippel et al., 2020). Equations (6) and (7) are the first and second stage of one 2SLS estimation, respectively, while Equations (8) and (9) are the first- and second-stages of a separate 2SLS estimation. The standard errors are then corrected using the usual 2SLS correction procedure.

In addition to the usual exclusion restrictions for the instrument (i.e., the kink $(x_i - k)D_i$) in the Y_i and Y_j outcome equations (see Section 3.2.1), estimation of direct and indirect effects additionally requires that the confounder in Y_i and Y_j outcome equations be independent. More formally, we require that $\epsilon^M \perp\!\!\!\perp \epsilon^Y$, which is akin to stating that the confounding variable that jointly affects $EverD_i$ and Y_i is independent of the confounding variable that jointly causes Y_i and Y_j (Dippel et al., 2020). The implication of this assumption is that an additional exclusion restriction is required, such that our instrument can be used as an instrument for the mediator Y_i when conditioned on $EverD_i$ in the Y_j outcome equation ($(x_i - k)D_i \perp Y_j(Y_i) \mid EverD_i$). It is important to note that this assumption does not assume away the endogeneity of $EverD_i$ in the Y_i outcome equation.

This assumption is reasonable in our setting, as the unobserved confounder which causes bias in the Y_i outcome equation when estimating the impact of $EverD_i$, is different to the one that causes the bias in Y_i in the Y_j outcome equation. As discussed in Section 3.2.2, when estimating the effect of $EverD_i$ on Y_i we are concerned with bias arising from simultaneity. Whereas when estimating the impact of Y_i on Y_j the source of bias is assortative matching. One way in which this assumption may be violated is if own diabetes diagnosis impacts partners' behaviour through increasing their probability of being diagnosed with diabetes. In other words, if own diabetes status impacts partners' diabetes status directly and it is this that induces the changes in partners' behaviour. If the spillover effect worked through this channel we would expect the magnitude of the spillover effect to fall when controlling for own and partners' diabetes status in the same regression. However, as we show in Appendix Section A6 the magnitude of the spillover effect is similar in this model, making such causal channel unlikely. In further support in Table 4 we directly estimate the probability of partners diabetes status as a result of own diabetes status. We find no evidence that probability of partners diagnosis increases as a result of own ever having had a diabetes diagnosis. We also show that there is no effect of own diabetes diagnosis on partners' probability of taking anti-diabetic medication (Appendix Section A7). However, to account for this possibility, we repeated estimation directly controlling for partner's diabetes status (and using the appropriate instrument), and find that magnitudes remain comparable to those in the main results, again offering evidence against this channel. Cross-tabulation of partners' diabetes status is presented in Table 5.

Testing the additional requirement that the instrument is relevant for the mediator Y_i when conditioned on $EverD_i$, we present F-statistics for Equation (8) in Table 6. F-statistics values, as expected, are much smaller than for Equation (6), however, for the outcome in which we find evidence of a spillover effect, they suggest our instrument is valid.

β_Y^T is an estimate of the effect of a change in partner j's behaviour that is a result of partner i's diagnosis itself. The indirect effect $\beta_M^T \times \beta_Y^M$ captures the change in own behaviour that is caused by the induced change in partner's behaviours. The theoretical difference between these two channels is that the indirect effect reflects joint participation in

TABLE 4 Fuzzy RKD estimates of effect of own diagnosis on partner's diabetes status.

	Partner's diabetes status
Effect of own diabetes	−0.0327
	(0.0612)
First stage $F - Statistic$	263.53
Obs.	9474

Note: Coefficients are estimated using a linear specification each side of the kink point, equally weighted observations, and with bounds of 0.9 on the right hand tail, and 0.4 on the left hand side. Parentheses includes cluster-robust standard errors, clustered at the household level. Estimates include year fixed-effects.

*** denotes p-value of 0.01 or less, ** denotes p-value of 0.05 or less, * denotes p-value of 0.10 or less.

TABLE 5 Cross-tab of own and partners' diabetes diagnosis status.

Partners' diabetes diagnosis status	Own diabetes diagnosis status		
	Never diagnosed	Diagnosed with diabetes	Total
Never diagnosed	19,834	1089	20,923
Diagnosed with diabetes	1263	204	1467
Total	21,097	1293	22,390

Note: These descriptive statistics are taken from sample of individuals used to estimate Table 3 but without a constraint on bandwidth.

TABLE 6 Total, direct and indirect effect estimates from mediation analysis.

	Exercise	Vegetable consumption	Fruit consumption	Currently a smoker	Currently a drinker
Total effect					
Total effect of Partner's diabetes	0.433**	0.037	−0.015	0.0436	0.071
	(0.187)	(0.0990)	(0.0910)	(0.0663)	(0.0709)
Direct effect					
Partner's diagnosis ($EverD_i$)	−0.048	0.013	0.020	−0.040	−0.002
	(0.032)	(0.0319)	(0.0432)	(0.0344)	(0.0180)
Indirect effect					
Partner's behaviour (Y_i)	0.481**	0.024	−0.035	0.084	0.072
	(0.2139)	(0.110)	(0.1325)	(0.1034)	(0.0850)
First stage $F-Statistic$	65.26	239.39	239.43	254.48	254.47
Equation (6): $(x_i - k)D_i$ on $EverD_i$					
First stage $F-Statistic$	16.70	0.86	2.57	6.814	1.92
Equation (8): $(x_i - k)D_i$ on Y_i| $EverD_i$					
Obs.	3451	8201	8202	9583	9583

Note: The total effect corresponds to the coefficient δ_1 from Equation (5), albeit for a slightly smaller sample in some cases. The direct effect corresponds to β_Y^T in Equation (9), and the indirect effect corresponds to $\beta_M^T \times \beta_Y^M$ in Equations (7) and (9). Each stage uses a linear specification each side of the kink point, equally weighted observations, and with bounds of 0.9 on the right hand tail, and 0.4 on the left hand side. Parentheses includes cluster-robust standard errors, clustered at the household level. Each specifications include year fixed effects.

*** denotes p-value of 0.01 or less, ** denotes p-value of 0.05 or less, * denotes p-value of 0.10 or less.

these behaviours, whereas finding evidence of a direct effect would suggest that partners independently change their behaviours due to their partners diagnosis.

In the case of the direct effect, partner i receives new health information about their diagnosed condition which they then share with the non-diagnosed partner j. The transfer of information from partner i to j therefore provides j with a new information set which they use to *privately* re-evaluate their optimal behaviour and *independently* change their behaviours.

The indirect effect $\beta_M^T \times \beta_Y^M$ captures the change in own behaviour that is caused by the induced change in partner's behaviours, which suggest that couples are jointly participating in these behaviours. As discussed in the introduction, jointly participating in behaviours may be the result of complementaries in behaviours. In terms of physical activity, individual j may get utility or dis-utility from exercising, however joint time with their partner may provide sufficient utility to render exercising a utility increasing choice. The indirect effect may also reflect other channels, namely it could be the case that the new health information induces changes in both partners simultaneously.

Table 6 provides estimates of the direct and indirect effects from the mediation analysis. For physical activity we find that the spillover effect is driven by partner's behaviour Y_i, and we find limited evidence that the diagnosis itself is

causing a change in behaviours of *j*. These results suggest that couples are jointly participating in these behaviours in response to the diagnosis. For the remaining outcomes, as with the total effects, we find no evidence of direct or indirect effects.

6 | DISCUSSION AND CONCLUSION

Diabetes is a unique condition in that changes in lifestyle and behaviour is both the first line treatment and the recommended method of preventing the disease. By jointly partaking in diabetes treatment partners of people with diabetes could substantially benefit from their partners' diabetes diagnosis. In this paper we estimate the effect of own and partners' diabetes status on own lifestyle behaviours, namely exercising, diet, smoking and drinking. Exploiting guidelines around the levels of sugar in the blood and recommendation for annual testing for those above a specific threshold, a fuzzy kink regression design is implemented using data on blood samples and behaviours from the HSE dataset.

Findings show that ever been diagnosed with diabetes significantly increase physical activity, suggesting compliance with first line treatment. In analysis included in the appendix of this paper, we find that there is no evidence of heterogeneous effects by time-since-diagnosis, suggesting persistence in the effect. Most importantly, we uncover substantial spillover effects from diabetes diagnosis in the form of an increase in physical activity of partners. Although the spillover effects are smaller (approximately 50%) in magnitude compared to the impact of own diabetes diagnosis, the effect are substantial and represents a significant change in behaviours. We find no evidence of changes in fruit or vegetable consumption. However, we acknowledge that our results may not offer a complete picture of the dietary changes made by diagnosed individuals, especially since treatment recommendations often focus on reducing fat, salt, and sugar intake (Hut & Oster, 2022; Oster, 2018). Analysing the channels these effects work through, we find that partners are jointly partaking in these behaviours, rather than responding independently to the diagnosis. Overall, our findings suggest that non-communicable disease diagnoses possess unaccounted spillover effects that should feature in cost-benefit evaluations to establish the best value-for-money interventions. The range of spillover effects are potentially limited and disease specific but, nevertheless, raise an important issue for economic evaluation studies that take the societal and system perspective.

Comparing our results of the own effect to previous studies, our estimated impact on diet differs to Hut and Oster (2022) and Oster (2018). Hut and Oster (2022) estimated there to be significant and positive changes in diet post-diagnosis, and found that increased fruit purchases was the fourth largest contributor to these dietary changes. However, their results suggest that the improvements in diet begin to fade over time. Oster (2018) found that purchases of fruit and vegetables increase in the month post-diagnosis, however the effect also decreases over time, and between months 2–12 post-diagnosis the effect becomes insignificant. Although our results do differ to Hut and Oster (2022) and Oster (2018), we note the difference in time-since-diagnosis between the studies and suggest that our findings follow the same temporal pattern. Given that the average time-since-diagnosis in our sample is over 10 years, and that Hut and Oster (2022) and Oster (2018) both find decreasing effects over time, it might be expected that the effects reduce to zero in the long-run.

Kim et al. (2019), Iizuka et al. (2021), and Alalouf et al. (2023) analyse the impact of diabetes diagnosis on physical activity. Kim et al. (2019) find no evidence of an increase in physical activity as a result of a diabetes diagnosis in either the short-run (1 or 2 years) or the long-run (3 or 4 years). Alalouf et al. (2023) find no evidence of a change in the probability of "engaging in physical activity 3+ times per week" or "recently made changes to increase physical activity". Iizuka et al. (2021) find that crossing the diabetes diagnosis threshold increases an individual's probability of "exercising enough to work up a sweat for 30 min or more per day [for] 2 ~7 days per week" by 4%. Whereas we find large and persistent effects on the probability of "any physical activity in the past 4 weeks". Our seemingly large point estimate does not imply that individuals make wholesale changes to their physical activity, but rather that they now partake in some physical activity. The threshold for responding positively to our question is much lower than those used by Kim et al. (2019) and Iizuka et al. (2021). The impact on the latent amount of physical activity may be marginal, and may be the same in magnitude as Iizuka et al. (2021). Finally in terms of the own effect on smoking and drinking behaviour, our results concur with Iizuka et al. (2021) and Alalouf et al. (2023), who estimate a null effect of a diabetes diagnosis.

Unfortunately, there are no studies to directly compare our spillover estimates to, albeit our broader conclusions do concur with previous studies. Clark and Etilé (2002) conclude that partners' health plays no role in the decision to quit smoking. Our estimates also show that a partners' diabetes diagnosis does not lead to a change in the smoking

behaviour of their partner. Our results may also support the conclusions of Clark and Etilé (2006) who found that the correlation between partners' smoking behaviour was driven mainly by matching in the marriage market, indeed we do not find any evidence to the contrary in terms of smoking behaviour. Comparing our results to those of Fletcher and Marksteiner (2017) is challenging, as they estimate the spillovers from a smoking cessation and alcoholism programme, not a health shock. Given that we find no evidence that the diagnosed individual changes their smoking or drinking behaviour, we may also not expect to find evidence of a spillover. Finally, although we cannot directly compare our results to Fadlon and Nielsen (2019), both studies find significant health-related behavioural spillovers.

As with previous work that analyses marginally-ill patients, there are some limitations to our approach (Alalouf et al., 2023; Iizuka et al., 2021; Kim et al., 2019; Rodríguez-Lesmes, 2021). Clearly, our methodology uses only individuals close to the diagnosis cut-off. This has several implications. Firstly, there is a bias-variance trade-off when choosing bandwidths for RDD or RKD estimators. Narrow bandwidths reduce the chances of misspecification errors, however smaller samples have lower power, which is of particular concern for our spillover estimates given the small initial sample we have. Large samples improve precision of estimates but also increase the chances of misspecifing the functional form, therefore increasing the risk of bias (Cattaneo et al., 2020). To ensure that these issues are handled, we present a number of alternative bandwidths and specifications in the Appendix. Focusing on marginally-ill patients also means that we are unable to make any claims regarding the impact a diabetes diagnosis has on individuals away from the cut-off. Our conclusion are specifically for those patients that are close to the diabetes risk cut-off we investigate. However, marginally-ill patients are likely those individuals who are of interest in terms of policy analysis, and the ones that most likely benefit from health checks (Rodríguez-Lesmes, 2021). Another limitation is that we do not have longitudinal data in which we observe a health shock and then the subsequent changes in behaviours. Indeed, we analyse the impact of a diabetes diagnosis using past diabetes diagnosis and contemporaneous HbA1c levels, which is a limitation of our approach. This may lead to potential concerns regarding the existence of "defiers" in our 2SLS set-up. To ensure that we are identifying a causal effect, we outline the necessary conditions for identification, and provide ample evidence along with robustness checks to validate our approach. Further, in Appendix Section A3 we replicate our analysis and confirm our main findings using the English Longitudinal Survey for Ageing (ELSA) which includes contemporaneous values of HbA1c and diabetes status.

From a public health perspective, confirmation of long-term compliance of those diagnosed with diabetes to first line treatments and necessary lifestyle changes is reassuring, at least in relation to physical activity. However, further work is required on how to induce behavioural changes in terms of diet, tobacco and alcohol consumption. From a policy perspective, our findings suggest that benefit evaluation of diabetes interventions needs to be revisited in the presence of spillover effects, as their current benefit-cost ratio is likely to be substantially underestimated, especially in relation to physical activity.

ACKNOWLEDGMENTS

We would like to thank Thomas Gall, Grant Gibson, Brendon McConnell, Christopher Millet, Carmine Ornaghi, Matt Sutton, Hans van Kippersluis and participants of the European Economic Association Congress, the Royal Economic Society Conference, the Virtual Economics of Risky Behaviours Seminar, and the PhD - Economics Virtual Seminar, for their feedback on various version of this paper. We would also like to thank the editor and two anonymous referees for their valuable comments which substantially improved this paper. Some of this work was done while Rhys Llewellyn Thomas was a PhD Student at the University of Southampton, which was funded by an ESRC 1 + 3 PhD Studentship.

CONFLICT OF INTEREST STATEMENT

Authors have no conflicts of interest to disclose.

DATA AVAILABILITY STATEMENT

Data used is the Special License version of the Health Survey for England, years 2003–2015. The data that support the findings of this study are available at the UK Data Service (ukdataservice.ac.uk). Restrictions apply to the availability of this data, which were used under license for this study. Data are available from https://doi.org/10.5255/UKDA-Series-2000021 with the permission of NatCen and NHS Digital.

ETHICS STATEMENT

This work received ethical approval from the Social Sciences Ethics Committee University of Southampton (52974) in 2019.

ORCID

Rhys Llewellyn Thomas https://orcid.org/0000-0003-0124-0391

ENDNOTES

[1] On one hand, experts encourage fruit consumption due to their low energy density, and high content of vitamins, minerals, phytochemicals and dietary fibre. Others argue that fruit should be limited due to the high carbohydrate content which raises blood sugar, which is problematic in those with diabetes (Forouhi et al., 2018). NHS advice states that those with diabetes should "eat a wide range of foods - including fruit", the advice also states that individuals should "keep sugar, fat and salt to a minimum" (NHS, 2018), which can potentially cause confusion due to the high sugar content of fruit. Indeed, there are a number of ongoing campaigns to resolve understanding of the guidelines (Diabetes UK, n.d.). However, confusion is present both among healthcare professionals and patients with 25% and 57% respectively, stating that "fresh fruit can be eaten freely with little effect on blood glucose levels" (Speight & Bradley, 2001). Forouhi et al. (2018) state that "consumption of fruits should be guided within the overall dietary pattern of an individual, their taste and other preferences and by their glycaemic control and need for antidiabetic medication, supported by healthcare professionals".

[2] An alternative measure, blood glucose level, is the concentration of sugar in the blood at a single point in time and is highly variable within individuals, and more dependent on very recent consumption than persistent behaviour.

[3] Yudkin and Montori (2014) state that "glycaemia are continuous, with no inflections to provide obvious cut-off points. Cut-offs for the diagnosis of diabetes are based on thresholds for risk of retinopathy. Lesser degrees of hyperglycaemia increase the risk of developing diabetes and maybe arterial disease. But in both cases the risk is graded, making any choice of cut-off point purely arbitrary." This claim is also supported by NICE (2011, 2012b).

[4] The National Institute for Health and Care Excellence (NICE) is an executive non-departmental public body of the UK Department of Health and Social Care which publishes guidelines for clinical practice and the use of healthcare technologies in the National Health Service.

[5] Given that only approximately a third of survey participants had a blood sample taken, we may be concerned about selection into blood sample. We explore this further in Appendix Section A10 in which we show balance tables for the blood sample, partner and blood sample, and those excluded from both. We also replicate our main results using weights so that our estimation sample matches the entire HSE sample in terms of descriptive statistics.

[6] A change in calibration of the equipment used for analysing HbA1c was made on 19th of September 2013, which resulted in a slight change in result for equivalent blood samples. Throughout the analysis we use "valid HbA1c result", as recommend in the HSE documentation, which adjusts the results post-2013 to be equivalent to pre-2013 results for the same blood samples.

[7] Given our data, we are unable to explore effects on other foods/nutrients such as sugar, fat, and salt for which guidelines suggest reduced intake.

[8] Although the NHS recommends a diabetes diagnosis at a HbA1c level of 6.5% there is no discontinuity or kink-point at this level in our data, and therefore we do not exploit this threshold in our analysis. We explore the possibility of an alternative jump or kink point in the Appendix Section A12, by testing the fit of alternative kink/jump-points and specifications.

[9] As discussed in Iizuka et al. (2021) there is not a clear method for determining the optimal bandwidth, given the discrete nature of the running variable. However, we transparently present a number of bandwidths in the appendix of this paper to ensure robustness.

[10] Although several studies find significant changes in HbA1c from a range of interventions, these changes are typically small in magnitude. The estimated impact of lifestyle interventions on long-term HbA1c levels in the literature is typically a drop by 0.5% (Andrews et al., 2011; Diabetes Prevention Program Research Group, 2002; Kazeminezhad et al., 2018; Rothman et al., 2005; Tshiananga et al., 2012) and these studies are usually conducted with individuals with high levels of HbA1c (i.e., not at the margin). Therefore, it would require extremely drastic and persistent interventions for diagnosed individuals to not only go into remission (below 6.5%), but to also cross our cut-off here of 6.0%.

[11] Of these approximately 40% discovered their diabetes during pregnancy. Further, one may be concerned that taking insulin might lead to the same result, but in our data only 23% of those with diabetes take insulin, and of the small number of people below the cut-off that have diabetes, 97% of those are not taking insulin.

[12] The reduced form estimates are estimated using: $Y_i = \chi_0 + \chi_1(x_i - k)D_i + \left[\sum_{p=1}^{p^*} \psi_p^-(x_i - k)^p\right] + \left[\sum_{p=2}^{p^*} \psi_p^+(x_i - k)^p D_i\right] + \mu_i$.

[13] Reduced form RKD estimates are estimated using $Y_i = \chi_0 + \chi_1(x_i - k)D_i + \left[\sum_{p=1}^{p^*} \psi_p^-(x_i - k)^p\right] + \left[\sum_{p=2}^{p^*} \psi_p^+(x_i - k)^p D_i\right] + \mu_i$.

REFERENCES

Alalouf, M., Miller, S., & Wherry, L. R. (2023). What difference does a diagnosis make? Evidence from marginal patients. *American Journal of Health Economics*. https://www.journals.uchicago.edu/doi/abs/10.1086/724415

Al-Janabi, H., Van Exel, J., Brouwer, W., Trotter, C., Glennie, L., Hannigan, L., & Coast, J. (2016). Measuring health spillovers for economic evaluation: A case study in meningitis. *Health Economics*, *25*(12), 1529–1544. https://doi.org/10.1002/hec.3259

Andrews, R., Cooper, A., Montgomery, A., Norcross, A., Peters, T., Sharp, D., Jackson, N., Fitzsimons, K., Bright, J., Coulman, K., England, C., Gorton, J., McLenaghan, A., Paxton, E., Polet, A., Thompson, C., & Dayan, C. (2011). Diet or diet plus physical activity versus usual

care in patients with newly diagnosed type 2 diabetes: The early ACTID randomised controlled trial. *The Lancet*, *378*(9786), 129–139. https://doi.org/10.1016/s0140-6736(11)60442-x

Bajaj, M. (2012). Nicotine and insulin resistance: When the smoke clears. *Diabetes*, *61*(12), 3078–3080. https://doi.org/10.2337/db12-1100

Becker, G. S. (1973). A theory of marriage: Part I. *Journal of Political Economy*, *81*(4), 813–846. https://doi.org/10.1086/260084

Becker, G. S. (1981). A treatise on the family. Retrieved from https://www.nber.org/books/beck81-1

Bergman, B. C., Perreault, L., Hunerdosse, D., Kerege, A., Playdon, M., Samek, A. M., & Eckel, R. H. (2012). Novel and reversible mechanisms of smoking-induced insulin resistance in humans. *Diabetes*, *61*(12), 3156–3166. https://doi.org/10.2337/db12-0418

Bove, C. F., Sobal, J., & Rauschenbach, B. S. (2003). Food choices among newly married couples: Convergence, conflict, individualism, and projects. *Appetite*, *40*(1), 25–41. https://doi.org/10.1016/s0195-6663(02)00147-2

Card, D., Lee, D. S., Pei, Z., & Weber, A. (2015). Inference on causal effects in a generalized regression kink design. *Econometrica*, *83*(6), 2453–2483. https://doi.org/10.3982/ecta11224

Cattaneo, M. D., Titiunik, R., & Vazquez-Bare, G. (2020). The regression discontinuity design. In *Handbook of Research methods in political science and international relations* (pp. 835–857). Sage Publications. Retrieved from http://dx.doi.org/10.4135/9781526486387.n47

Chang, S. A. (2012). Smoking and type 2 diabetes mellitus. *Diabetes & Metabolism Journal*, *36*(6), 399–403. https://doi.org/10.4093/dmj.2012.36.6.399

Chiolero, A., Faeh, D., Paccaud, F., & Cornuz, J. (2008). Consequences of smoking for body weight, body fat distribution, and insulin resistance. *The American Journal of Clinical Nutrition*, *87*(4), 801–809. https://doi.org/10.1093/ajcn/87.4.801

Christakis, N. A., & Fowler, J. H. (2008). The collective dynamics of smoking in a large social network. *New England Journal of Medicine*, *358*(21), 2249–2258. https://doi.org/10.1056/nejmsa0706154

Clark, A., & Etilé, F. (2002). Do health changes affect smoking? Evidence from British panel data. *Journal of Health Economics*, *21*(4), 533–562. https://doi.org/10.1016/s0167-6296(01)00140-0

Clark, A. E., & Etilé, F. (2006). Don't give up on me baby: Spousal correlation in smoking behaviour. *Journal of Health Economics*, *25*(5), 958–978. https://doi.org/10.1016/j.jhealeco.2006.02.002

Courtemanche, C., Tchernis, R., & Ukert, B. (2018). The effect of smoking on obesity: Evidence from a randomized trial. *Journal of Health Economics*, *57*, 31–44. https://doi.org/10.1016/j.jhealeco.2017.10.006

Diabetes, U. K. (n.d.). Myth: I can't eat fruit if I have diabetes. Retrieved from https://www.diabetes.org.uk/guide-to-diabetes/enjoy-food/eating-with-diabetes/diabetes-food-myths/myth-fruit-diabetes

Diabetes, U. K. (2019). Facts and stats 2019. Retrieved from https://www.diabetes.org.uk/professionals/position-statements-reports/statistics

Diabetes Prevention Program Research Group. (2002). Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *New England Journal of Medicine*, *346*(6), 393–403.

Dippel, C., Gold, R., Heblich, S., & Pinto, R. (2020). Mediation analysis in IV settings with a single instrument, working paper.41.

Dong, Y. (2011). *Jumpy or kinky? Regression discontinuity without the discontinuity*. University of California-Irvine, Department of Economics. Retrieved from https://ideas.repec.org/p/irv/wpaper/111207.html.55.

Dupuy, A., & Galichon, A. (2014). Personality traits and the marriage market. *Journal of Political Economy*, *122*(6), 1271–1319. https://doi.org/10.1086/677191

Eliasson, B. (2003). Cigarette smoking and diabetes. *Progress in Cardiovascular Diseases*, *45*(5), 405–413. https://doi.org/10.1053/pcad.2003.00103

Ezzati, M., & Riboli, E. (2012). Can noncommunicable diseases be prevented? Lessons from studies of populations and individuals. *Science (New York, N.Y.)*, *337*(6101), 1482–1487. https://doi.org/10.1126/science.1227001

Ezzati, M., & Riboli, E. (2013). Behavioral and dietary risk factors for noncommunicable diseases. *New England Journal of Medicine*, *369*(10), 954–964. https://doi.org/10.1056/NEJMra1203528

Facchini, F. S., Hollenbeck, C. B., Jeppesen, J., Ida Chen, Y. D., & Reaven, G. M. (1992). Insulin resistance and cigarette smoking. *The Lancet*, *339*(8802), 1128–1130. https://doi.org/10.1016/0140-6736(92)90730-q

Fadlon, I., & Nielsen, T. H. (2019). Family health behaviors. *The American Economic Review*, *109*(9), 3162–3191. https://doi.org/10.1257/aer.20171993

Falba, T. A., & Sindelar, J. L. (2008). Spousal concordance in health behavior change. *Health Services Research*, *43*(1p1), 96–116. https://doi.org/10.1111/j.1475-6773.2007.00754.x

Farrell, L., & Shields, M. A. (2002). Investigating the economic and demographic determinants of sporting participation in England. *Journal of the Royal Statistical Society: Series A*, *165*(2), 335–348. https://doi.org/10.1111/1467-985x.00626

Fletcher, J., & Marksteiner, R. (2017). Causal spousal health spillover effects and implications for program evaluation. *American Economic Journal: Economic Policy*, *9*(4), 144–166. https://doi.org/10.1257/pol.20150573

Forouhi, N. G., Misra, A., Mohan, V., Taylor, R., & Yancy, W. (2018). Dietary and nutritional approaches for prevention and management of type 2 diabetes. *BMJ*, k2234. https://doi.org/10.1136/bmj.k2234

Helmrich, S. P., Ragland, D. R., Leung, R. W. and Paffenbarger, R. S. (1991). Physical activity and reduced occurrence of non-insulin-dependent diabetes mellitus, *New England Journal of Medicine 325*(3): 147–152.oi: https://doi.org/10.1056/NEJM199107183250302

Hu, F. B., Manson, J. E., Stampfer, M. J., Colditz, G., Liu, S., Solomon, C. G., & Willett, W. C. (2001). Diet, lifestyle, and the risk of type 2 diabetes mellitus in women. *New England Journal of Medicine*, *345*(11), 790–797. https://doi.org/10.1056/nejmoa010492

Hut, S., & Oster, E. (2022). Changes in household diet: Determinants and predictability. *Journal of Public Economics*, *208*, 104620. https://doi.org/10.1016/j.jpubeco.2022.104620

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

Janssen, A., & Parslow, E. (2021). Pregnancy persistently reduces alcohol purchases: Causal evidence from scanner data. *Health Economics*, *30*(2), 231–247. https://doi.org/10.1002/hec.4188

Kazeminezhad, B., Taghinejad, H., Borji, M., & Tarjoman, A. (2018). The effect of self-care on glycated hemoglobin and fasting blood sugar levels on adolescents with diabetes. *Journal of Comprehensive Pediatrics*, *9*(2). https://doi.org/10.5812/compreped.62661

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. *Journal of Health Economics*, *65*, 1–14. https://doi.org/10.1016/j.jhealeco.2019.01.003

Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy*, *74*(2), 132–157. https://doi.org/10.1086/259131

Macario, E., & Sorensen, G. (1998). Spousal similarities in fruit and vegetable consumption. *American Journal of Health Promotion: AJHP*, *12*(6), 369–377. https://doi.org/10.4278/0890-1171-12.6.369

Mikhailidis, D., Papadakis, J., & Ganotakis, E. (1998). Smoking, diabetes and hyperlipidaemia. *Journal of the Royal Society of Health*, *118*(2), 91–93.

NHS. (2018). *Type 2 diabetes - food and keeping active*. Section: conditions. Retrieved from https://www.nhs.uk/conditions/type-2-diabetes/food-and-keeping-active/

NICE. (2011). *Type 2 diabetes prevention: Population and community-level interventions*. Technical report. NICE Guideline. Publisher: NICE. Retrieved from https://www.nice.org.uk/guidance/ph35

NICE. (2012a). *The guidelines manual: Assessing cost effectiveness, Technical Report Process and methods [PMG6]*. NICE. Publisher. Retrieved from https://www.nice.org.uk/process/pmg6/chapter/assessing-cost-effectiveness

NICE. (2012b). Type 2 diabetes: Prevention in people at high risk. Retrieved from https://www.nice.org.uk/guidance/ph38/chapter/Recommendations

Orphanides, A., & Zervos, D. (1995). Rational addiction with learning and regret. *Journal of Political Economy*, *103*(4), 739–758. https://doi.org/10.1086/262001

Oster, E. (2018). Diabetes and diet: Purchasing behavior change in response to health information. *American Economic Journal: Applied Economics*, *10*(4), 308–348. https://doi.org/10.1257/app.20160232

Pinkowish, M. D. (1999). Hand in glove: Smoking cessation and weight gain. *Patient Care*, *33*(2), 134–135.

Rodríguez-Lesmes, P. (2021). Estimating the gains of early detection of hypertension over the marginal patient. *PLoS One*, *16*(7), e0254260.

Rothman, R. L., Malone, R., Bryant, B., Shintani, A. K., Crigler, B., Dewalt, D. A., Dittus, R. S., Weinberger, M., & Pignone, M. P. (2005). A randomized trial of a primary care-based disease management program to improve cardiovascular risk factors and glycated hemoglobin levels in patients with diabetes. *The American Journal of Medicine*, *118*(3), 276–284. https://doi.org/10.1016/j.amjmed.2004.09.017

Schwamm, L. H. (2018). The communicable nature of non-communicable diseases. *The Lancet Neurology*, *17*(8), 665. https://doi.org/10.1016/s1474-4422(18)30216-3

Śliwińska Mossoń, M., & Milnerowicz, H. (2017). The impact of smoking on the development of diabetes and its complications. *Diabetes and Vascular Disease Research*, *14*(4), 265–276. https://doi.org/10.1177/1479164117701876

Speight, J., & Bradley, C. (2001). The ADKnowl: Identifying knowledge deficits in diabetes care. *Diabetic Medicine*, *18*(8), 626–633. https://doi.org/10.1046/j.1464-5491.2001.00537.x

Tshiananga, J. K. T., Kocher, S., Weber, C., Erny-Albrecht, K., Berndt, K., & Neeser, K. (2012). The effect of nurse-led diabetes self-management education on glycosylated hemoglobin and cardiovascular risk factors: A meta-analysis. *The Diabetes Educator*, *38*(1), 108–123. https://doi.org/10.1177/0145721711423978

WHO. (n.d.). Diabetes. Retrieved from https://www.who.int/westernpacific/health-topics/diabetes

WHO. (2011). Use of glycated haemoglobin (HbA1c) in the diagnosis of diabetes mellitus, Technical report. Retrieved from https://www.who.int/diabetes/publications/diagnosis_diabetes2011/en/

WHO. (2016). Global report on diabetes, Technical report. Retrieved from https://www.who.int/publications-detail/global-report-on-diabetes

Willi, C., Bodenmann, P., Ghali, W. A., Faris, P. D., & Cornuz, J. (2007). Active smoking and the risk of type 2 diabetes: A systematic review and meta-analysis. *JAMA*, *298*(22), 2654–2664. https://doi.org/10.1001/jama.298.22.2654

Wittenberg, E., James, L. P., & Prosser, L. A. (2019). Spillover effects on caregivers' and family members' utility: A systematic review of the literature. *PharmacoEconomics*, *37*(4), 475–499. https://doi.org/10.1007/s40273-019-00768-7

World Health Organization. (2018). *Noncommunicable diseases country profiles 2018*. World Health Organization.

Yeh, H.-C., Duncan, B. B., Schmidt, M. I., Wang, N.-Y., & Brancati, F. L. (2010). Smoking, smoking cessation, and risk for type 2 diabetes mellitus. *Annals of Internal Medicine*, *152*(1), 10–17.

Yudkin, J. S., & Montori, V. M. (2014). The epidemic of pre-diabetes: The medicine and the politics. *BMJ*, *349*(jul15 24), g4485. https://doi.org/10.1136/bmj.g4485

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Thomas, R. L., & Mentzakis, E. (2024). The direct and spillover effects of diabetes diagnosis on lifestyle behaviours. *Health Economics*, 1–19. https://doi.org/10.1002/hec.4803