European University Institute

DEPARTMENT OF ECONOMICS

EUI WORKING PAPERS

ECO 2020/02
Department of Economics

Impact of a Health Shock on Lifestyle Behaviours

Zoey Verdun

European University Institute
Department of Economics

Impact of a Health Shock on Lifestyle Behaviours

Zoey Verdun

EUI Working Paper **ECO** 2020/02

ISSN 1725-6704

Published in June 2020 by the European University Institute.
Badia Fiesolana, via dei Roccettini 9
I – 50014 San Domenico di Fiesole (FI)
Italy

With the support of the Erasmus+ Programme of the European Union

The European Commission supports the EUI through the European Union budget. This publication reflects the views only of the author(s), and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Impact of a Health Shock on Lifestyle Behaviours

Zoey Verdun*

June 2020

Abstract

Following a healthier lifestyle can improve living quality. Yet mixed evidence exists for whether a health shock induces individuals to change their lifestyle. Panel data from the UK Household Longitudinal Study, *Understanding Society*, is used to estimate the response to a health shock – heart attack or diabetes diagnosis – on a healthy lifestyle index, composed of eight lifestyle behaviours. Using a matching approach, this paper finds a significant positive effect on the index; a large effect is found for a strong shock, but no effect for a weak one. The overall effect is driven by increased fruit and vegetable consumption, decreased number of cigarettes smoked and increased probability to quit drinking alcohol. Among those drivers there is heterogeneity by sex, such as only women increase the probability to quit drinking. Lifestyle changes following a shock suggest updated beliefs about an individual's health status, with heterogeneous costs of change across individuals and behaviours.

Keywords: health shocks; lifestyle behaviours; behavioural change

JEL codes: I12, D83

1 Introduction

Studies have shown that adopting a healthier lifestyle – such as reducing or quitting smoking, improving diet, exercising and reducing alcohol consumption – can improve quality of life by both extending an individual's lifespan and increasing the quality of the years to come (Chou, Hwang, & Wu, 2012; Rizza, Veronese, & Fontana, 2014). In particular, a healthier lifestyle can prevent chronic diseases such as obesity, cardiovascular disease and diabetes; the latter two being among the top ten global causes of death (World Health Organization, 2018). Remarkably, various scholars have found evidence suggesting that the progress of these diseases can be stopped or in some cases even reversed through lifestyle changes.[1]

*zoey.verdun@eui.eu, European University Institute (EUI)

[1]Esselstyn, Ellis, Medendorp, & Crowe (1995); Buttar, Li, & Ravi (2005); McMacken & Shah (2017).

Governments and other institutions have placed a great deal of emphasis on encouraging the adoption of better lifestyles, often through information campaigns. Despite these efforts, the public health literature suggests that information alone is often not very successful in changing lifestyle behaviours (Kelly & Barker, 2016). This appears to be particularly true of diet, whereas the evidence is more mixed in the case of exercise, smoking and alcohol consumption. Nevertheless, the medical literature has some well-documented cases of individuals who *do* make successful lifestyle changes (Esselstyn et al., 1995; Ornish et al., 1998; Lanza et al., 2001).[2] These contrasting findings raise the questions: (why) do some individuals change their lifestyle behaviours while others do not?

The primary goal of this paper is to estimate what impact experiencing a health shock – the diagnosis of a heart attack or diabetes – has on lifestyle behaviours, where the behaviours are related to diet, exercise, smoking and alcohol consumption. The health shock is interpreted as a signal about an individual's health status. For the analysis a healthy lifestyle index is created, which consists of eight behaviours: one for diet[3], two each for exercise and smoking, and three for alcohol consumption.[4] I also explore related questions such as the effect of sex on the main analysis and the impact of the shock strength. The latter is the comparison of the impact of experiencing a weaker shock, the diagnosis of a disease risk factor, with experiencing a stronger shock, the diagnosis of a disease itself.[5] Finally, I explore the heterogeneity in the degree of change across different lifestyle behaviours.

In the economics literature there are only a few studies that investigate what lifestyle changes, if any, an individual undertakes following a health shock. Furthermore, there is no consensus yet as to which lifestyle behaviours change. The three studies closest to this paper are first, Oster (2018) who finds that the healthiness of a household's diet largely does not respond to a diabetes diagnosis. Second, Hut & Oster (2018) building upon Oster (2018), find that neither life events, such as disease diagnosis, nor demographics

[2]Several aspects are often associated with successful lifestyle changes, including if the individual acquires new information or knowledge, has support (e.g. from spouse, family, friends or support groups), has certain personality traits, or has experienced a health shock (see for example: Lanza et al., 2001; Condon & McCarthy, 2006; Durkin, Brennan, & Wakefield, 2012). This paper focuses on the final aspect, the health shock.

[3]Henceforth, 'diet' or 'dietary change' with reference to the analysis only refers to a change in fruit and vegetable consumption. The dataset does not have many diet-related variables and therefore diet is proxied by fruit and vegetable consumption.

[4]The precise behaviours are 1) number of daily servings of fruits and vegetables consumed, 2) number of days in the past four weeks walked at least 10 minutes, 3) same as previous but for at least 30 minutes, 4) if smokes, 5) daily number of cigarettes smoked, 6) if drinks alcohol (within the past 12 months), 7) number of days did not drink alcohol in the past seven days and 8) total number of drinks consumed on the heaviest drinking day in the past seven days.

[5]Throughout this paper I refer to terms such as 'disease risk factor' or a 'disease diagnosis'. Although a heart attack is technically not a 'disease', I still refers to the diagnosis of both diabetes and heart attack as a 'disease diagnosis' for simplicity. Furthermore, a heart attack is nearly always an implicit diagnosis of heart disease. Finally, the UKHLS questions ask about a 'diagnosis' of heart attack.

predict dietary change. Instead, their study shows that baseline diet quality is a good predictor of large dietary change, with dietary concentration being the key driver. Third, Bhalotra, Delavande, Fisher, & James (2020) study the impact of a high blood pressure diagnosis on health behaviours. They find the diagnosis has an impact on decreasing the probability of smoking, has a possible but imprecise impact on improving diets, but has no impact on exercise or alcohol consumption behaviours.

Similarly, there are only a few studies in the medical literature investigating behavioural change in response to a health shock. The two most related studies are first, Chong et al. (2017) who compare the changes in lifestyle behaviours of patients with a recent diagnosis of diabetes to those who have never been diagnosed; they find that only minimal changes are made to lifestyle behaviours following a diagnosis. The recently diagnosed individuals have a lower decrease in vegetable consumption, experience more weight loss and are more likely to quit smoking than those not diagnosed. Second, Fassier et al. (2017) study the impact of a cancer diagnosis on the changes in diet and alcohol consumption; they find that a diagnosis improves some dietary intakes – decreases in alcohol and sweetened beverage consumption – but worsens others, such as a decrease in vegetable consumption.

This paper contributes to the economics literature on perceived risks and subjective probabilities of diseases.[6] It does so by studying *both* the impact of experiencing a health risk factor shock (noisier signal) and the impact of experiencing the actual health shock (clearer signal) on changes in lifestyle behaviour. Additionally, by studying the numerous lifestyle behaviours concurrently, it contributes to the literature on behavioural change and the likely heterogeneous costs of such change. Finally, this paper also contributes to the (economics) behavioural change literature and the findings that often behavioural change is difficult to achieve (e.g. Condon & McCarthy, 2006; Kelly & Barker, 2016; Oster, 2018; Bhalotra et al., 2020). It contributes by discussing how changes in different lifestyle behaviours may or may not occur depending on the costs that an individual faces when considering making one or more changes. By better understanding whether, and if so which, lifestyle behaviours change after a health shock, certain health interventions can take the results into account when trying to encourage changes in behaviour.

The main result of this paper is a positive association between a health shock and a subsequent change in lifestyle. On average, a health shock leads to a statistically significant increase of 0.41 standard deviations in the healthy behaviours lifestyle index. The effect is similar for women (0.46) and men (0.40); however, for about half of the drivers the magnitude of change differs statistically by sex. Both women and men increase their fruit and vegetable consumption, by approximately 0.25 servings per day; only women increase the probability they quit smoking, by 5.4 percentage points; both women and men reduce the number of daily cigarettes smoked (respectively 3.0 versus 4.2, though the

[6]This paper also makes a small contribution to the equivalent medical literature.

difference is not statistically significant); and only women increase the probability they quit drinking alcohol, by 13.1 percentage points.

The second analysis compares the impact of the strength of the shock on lifestyle behaviour. I find the diagnosis of a disease itself (strong shock) has a large effect on improving overall lifestyle, whereas the overall effect of the diagnosis of a disease risk factor (weak shock) has little to no effect. Looking at individual behaviours, the improvement in the lifestyle index affected by experiencing a *strong* shock is driven by an increase in daily fruit and vegetable consumption (0.26 servings) and a decrease of 4.0 cigarettes in the number smoked daily. While the overall effect of the weak shock is not significant, some underlying behaviours are statistically significant; these are decreases in smoking and drinking behaviours, both on the intensive and extensive margins. Specifically, an increase in probability to quit smoking by 2.8 percentage points, an increase in the probability to quit drinking alcohol by 3.9 percentage points, a decrease in the daily number of cigarettes smoked by 2.5 cigarettes and an increase in the number of days abstaining from drinking alcohol by 0.2 days per week.

The rest of the paper is structured as follows: Section 2 provides an overview of the related literature, both in economics and in medicine. Section 3 describes the data, which comes from the household panel survey *Understanding Society*. Section 4 reports the main empirical analysis, the impact of the shock on lifestyle behaviour, and Section 5 the secondary analysis, the impact of a strong versus a weak shock. Section 6 concludes.

2 Related Literature

The majority of the health economics literature on behavioural change builds upon the notion that the provision of (new) information and the subsequent updating of beliefs influences behaviour. However, a growing body of literature, not just in economics but also in health and psychology, finds that usually knowledge or information is necessary but not sufficient for change (Bartiaux, 2008; Mathis & Steffen, 2015; Kelly & Barker, 2016). For example, Kelly & Barker (2016) highlight several common errors made when attempting to foster behavioural change, whether it be changes in diet, alcohol consumption or physical activity, that are related to information.[7]

That being said, when looking across different individual health behaviours the evidence on the impact of information on behavioural change is mixed. For diet, it is the most clear, information is not sufficient (Brownell & Cohen, 1995; Worsley, 2002). There is even evidence that individuals are willing to forgo health benefits to prevent having to modify their diet (Atkin, 2016; Oster, 2018). For exercise, Young, Haskell, Taylor, & Fortmann (1996) find a health education campaign has little to no effect on physical ac-

[7] Two such errors are assuming that 'knowledge and information drive behaviour' or 'it is about getting the message across' (p.111).

tivity levels. In contrast, Craig, Tudor-Locke, & Bauman (2006) find a public-awareness campaign to increase walking among adults, using an objective self-monitoring tool, to be effective. For smoking, two integrative reviews find mass media campaigns effective when it comes to fostering awareness and behavioural change (Flay, 1987; Durkin et al., 2012). However, Strecher et al. (1994) only find positive effects for light to moderate smokers. Finally, for alcohol consumption Fleming, Manwell, Barry, Adams, & Stauffacher (1999) find interventions by a physician can be effective in reducing different outcomes of alcohol use. In contrast, a review of the existing literature on prevention and treatment of college students finds little support for educational or awareness programs, but more support for other (non-informational) interventions in reducing alcohol use and its negative consequences (Larimer & Cronce, 2002).

Compared to the limited success of broader information campaigns there are more example of successful change in both the behavioural change literature and the medical literature; however these examples often involve more than just the general provision of information. Significant behavioural change, when it happens, often occurs in specific subgroups of individuals or individuals in certain situations. For example, in the economics literature, Hut & Oster (2018) find evidence that dietary change occurs in individuals when their baseline diet is made up of a small number of foods. In the medical literature, such specific groups of individuals or situations are usually patients of certain doctors, intensive programs or interventions (Esselstyn et al., 1995; Lanza et al., 2001). For example, in the Lifestyle Heart Trial patients sustained intensive lifestyle changes – diet, exercise, smoking and stress management – over several years (Ornish et al., 1998).

There is also a growing literature in economics on the impact of the diagnosis of risk factors, such as hypertension or high blood pressure, and the provision of (tailored) health status information on changes in dietary behaviour (Zhao, Konishi, & Glewwe, 2013; Carrera, Hasan, & Prina, 2020) and other lifestyle behaviours such as smoking, exercise and alcohol use (Edwards, 2018; Bhalotra et al., 2020). In this literature, the paper closest to mine is Bhalotra et al. (2020). They find a high blood pressure diagnosis leads to a reduction in smoking but it has no impact on either exercise or alcohol consumption; furthermore they find possible but imprecise evidence that diets marginally improve after diagnosis. My paper differs from Bhalotra et al. (2020) in that where they only look at the impact of a diagnosis of a risk factor, high blood pressure, I look at the diagnosis of risk factors *and* the corresponding diagnosis of the disease itself.

A smaller related economics literature is on the impact of disease diagnoses, such as cancer, heart disease or diabetes, on dietary change (Fassier et al., 2017; Oster, 2018; Hut & Oster, 2018). One of the two papers in this area closest to mine, Oster (2018), investigates the impact of a diabetes diagnosis on diet changes and finds a small but significant effect on calorie reduction in the month right after the diagnosis, though the effect is no longer significant in the months following. In the other related paper, Hut &

Oster (2018) also find no effect of a diabetes diagnosis on diet, but rather that baseline diet and dietary concentration are the predictors of dietary change. My paper builds upon Oster (2018) and Hut & Oster (2018) by investigating the impact of disease diagnosis on several lifestyle behaviours, not only diet. By looking at several behaviours, and not just one, I allow for the possibility that individuals change certain behaviours but not others. Finally, this paper differs from Oster (2018) by including heart attack, alongside diabetes, as the health shock being diagnosed; Hut & Oster (2018) are more similar since they look at the diagnosis of three different disease categories, which include diabetes and heart disease.

This paper further contributes to both the above literatures – diagnosis of risk factors and diagnosis of disease – by looking at *both* the receiving of information, a signal, on health status via the diagnosis of risk factors (high blood pressure and angina) and the diagnosis of actual disease (heart attack and diabetes). To the best of my knowledge, this is the first paper to compare the differences in response to a diagnosis of risk factors versus a diagnosis of disease.

Another related literature this paper ties into is on incorrect knowledge and uncertainty about the risks and risk factors for certain health shocks, such as heart attacks, and how they correspond to health status. Individuals may under or overestimate their perceived risk from engaging in certain lifestyle behaviours and hence also their subjective probabilities of having or getting a disease (Belot, James, & Spiteri, 2019).[8] The lack of clear signals on health status is one of the reasons why individuals may not be willing to make any changes to their behaviour (Sanderson, Waller, Jarvis, Humphries, & Wardle, 2009; Logie-MacIver, Piacentini, & Eadie, 2012). In the medical literature, for example, Condon & McCarthy (2006) find that some individuals believed heart attacks only occurred in 'old' people and therefore thought they could postpone changing their lifestyle to a later time. They also find heart attack patients had already been aware of their poor lifestyles – whether it be smoking, stress or poor diet – and yet, for a variety of reasons and beliefs, many did not change their lifestyle prior to the event. Many individuals were waiting for an initial 'warning sign' in order to motivate themselves to improve their lifestyle (Condon & McCarthy, 2006). Similar to the economics literature, this need for a warning sign can be interpreted as a need to receive a clearer signal on an individual's health status before knowing what is the optimal level of a lifestyle (change) to implement.

[8]For examples of incorrectly perceived risks for the different lifestyles analysed in this paper, see: diet (Condon & McCarthy, 2006), smoking (Heikkinen, Patja, & Jallinoja, 2010), and exercise (Fitgerald, Singleton, Neale, Prasad, & Hess, 1994; Crombie et al., 2004). For alcohol consumption, incorrectly perceived risks may stem from conflicting recommendations between some public health associations, such as the American Heart Association (2014) and the current medical literature (Stockwell et al., 2016), a likely source of confusion to the public and doctors alike.

3 Data

The data is from the United Kingdom's Understanding Society longitudinal study (UKHLS), an annually collected representative sample of the United Kingdom (UK) population, that started in 2009. This panel data consists of objective and subjective questions on topics such as health, work, education, income, family and social life. The analysis uses the first five waves to provide pre-, during and post-treatment waves.[9] In the main analysis the health shock is the diagnosis of a heart attack or diabetes and the outcome is a measure of change in lifestyle behaviour. In the secondary analysis the health shock is split by its strength: strong and weak. After applying all inclusion restrictions nearly 16,000 observations remain, of which just over 230 experience a shock.[10]

Dependent Variable The dependent variable is a healthy lifestyle behaviour index, henceforth lifestyle index, that captures four lifestyle-behaviours related to diet, exercise, smoking and alcohol consumption. The lifestyle index is made up of eight components.[11]

The first component relates to diet, the number of servings of fruits and vegetables an individual consumed per day. The second and third relate to exercise, the number of days in the past four weeks that an individual went for a walk of at least 10 minutes and the number of days spent walking at least 30 minutes, respectively. The fourth and fifth relate to smoking, a dummy for if an individual currently smokes (extensive margin) and the number of cigarettes smoked per day (intensive margin). The final three components relate to alcohol consumption. The sixth captures the extensive margin using a dummy for if an individual is a drinker.[12] The seventh and eighth capture the intensive margin with respect to the past seven days: the number of days an individual did not drink alcohol and the total number of drinks consumed on the heaviest drinking day, respectively.

The index is calculated, following Kling, Liebman, & Katz (2007), at the individual level (i) and is the equally-weighted sum of the z-scores of each of the eight behaviour components (j). The z-score is obtained by subtracting the mean of a component (μ_j) from the individual's behaviour amount of that component (x_{ij}) and then dividing it by that component's standard deviation (σ_j):

$$Index_i = \sum_{j=1}^{8} \frac{x_{ij} - \mu_j}{\sigma_j}$$

[9] There are currently nine waves available. The lifestyle behaviour variables are only available in waves 2, 5 and 7; however wave 7 is not used because some questions have been changed compared to those in waves 2 and 5. Wave 1 captures health shocks that may have occurred prior to the start of the UKHLS.
[10] If the shock type is not specified then it always refers to the shock in the main analysis.
[11] There are only seven behaviour variables listed in the pre-analysis plan (see Appendix A.1 for details).
[12] Defined as having had at least one alcoholic drink in the past 12 months.

Independent Variables The shock in the main analysis is the diagnosis of a heart attack or diabetes.[13] Since these diagnosis events are relatively rare, for reasons of power, the two kinds of shocks are pooled and treated as one. The shock variable takes a value of 1 (shock experienced) if the individual was 'newly diagnosed' with at least one of either a heart attack or diabetes or 0 (no shock experienced) if not newly diagnosed with either condition. A medical condition is considered newly diagnosed if the individual responded 'yes' to being diagnosed in waves 3 or 4 and 'no' in the previous waves (1 and 2); a medical condition is considered 'not diagnosed' if the individual responded 'no' in all four waves (1 through 4).[14]

The secondary analysis splits the shock into a strong shock and a weak shock to differentiate possible impacts of the shock strength. This differentiation in shock strength is part of the analysis because a strong shock gives a clear signal of health status; the exact health status from these weak shocks is much less precise. There is a significant share of individuals who receive one or both of the weak shock diagnoses but do not go on to experience either a heart attack or diabetes diagnosis. The shocks are defined as follows: a strong shock is still a diagnosis of a heart attack or diabetes, whereas a weak shock is the diagnosis of angina or high blood pressure. Two analyses are run, each using one of the two shock variables: the 'strong shock only' and the 'weak shock only' variables. The former takes a value of 1 if the individual experienced a strong shock and did not experience a weak shock. It takes a value of 0 if neither shock was experienced. The latter variable takes a value of 1 if the individual experienced a weak shock and did not experience a strong shock. Again, it takes a value of 0 if neither shock was experienced.[15]

Controls An individual's likelihood of a health shock is not uncorrelated to an individual's pre-shock behaviours. Therefore, it is important to account for an individual's initial health shock risk level (i.e. the probability of being diagnosed with either a heart attack or diabetes). This risk level is partly determined by an individual's previous diet, exercise, smoking behaviour and possibly alcohol consumption.[16] Included as part of the analysis are the risk factor variables that make up this initial risk,[17] which in the UKHLS are: age, sex, high blood pressure, smoking (extensive and intensive margins), fruit and vegetable consumption, and physical activity.

[13] It is of course possible that an individual receives both a heart attack and a diabetes diagnosis in the same time frame; and it is quite plausible that the effect of those shocks is as a consequence stronger; however in this study receiving two diagnoses is treated the same as receiving just one.

[14] Individuals are usually surveyed every 12 months. Therefore, the time between the measurement of the pre-shock behaviour and the occurence of the shock is typically 0-24 months; the time between the shock and the measurement of the post-shock behaviour is usually 12-36 months.

[15] For either variable if a person experiences both kinds of shock they are excluded from the analysis.

[16] Alcohol consumption is not included as a risk factor because public health associations, such as the American Heart Association (2014), are less clear on its impact on the risk of heart attack or diabetes.

[17] These risk factors come from risk assessment tools such as "Your Disease Risk" (n.d.) and from organizations such as the American Heart Association (2017).

Other controls include education, ethnicity, employment status and geography. Education is included using the derived variable 'highest education ever reported'. The six original categories are merged into four: bachelor's degree or above, high school completion (A-level), high school completion or equivalent, and no qualifications. Ethnicity is included as a binary variable: white and non-white, where the non-white category consists of the following ethnic groups: mixed; Asian or Asian British; black/African/Caribbean/black British; and other. Employment status is included as a proxy for income: individuals are split into either full-time employed, part-time employed or inactive. An urban dummy is included, which indicates if an individual lives in a rural or urban region. Finally, a categorical variable is included that indicates in which of the twelve UK Government Office Regions (GOR) an individual resides.

Inclusion Restrictions Since the health shock diagnosis (in wave 3 or 4) is the main dependent variable in the analysis, the following two restrictions on inclusion are necessary. Exclude individuals who have been diagnosed with a heart attack or diabetes previously;[18] exclude individuals if information about their diagnosis, or lack thereof, is not available in prior waves, to prevent any confounding effects.[19] Finally, exclude individuals if the shock variable, at least one of the index components, or any of the main controls are missing.[20]

3.1 Descriptive Statistics

Descriptive statistics are shown in Table 1. In the sample, 44% of individuals have a bachelor's degree or above, about half have some form of high school degree and 10% have no qualifications. The sample consists of 89% white individuals, is nearly 60% female and has an average age of 48. In terms of health behaviours, individuals, prior to treatment, on average consume 3.4 daily servings of fruits and vegetables, walk 15.4 days per month at least 10 minutes per day, walk 9.5 days per month at least 30 minutes per day, 81% do not smoke and only 12% do not drink. In the full sample, the average number of daily cigarettes smoked is 2.4, whereas among smokers the average is 11.7. For alcohol consumption, in the full sample, individuals consume 2.8 drinks on their heaviest drinking day in a week, whereas looking only at drinkers, it is 3.9. Finally, in the full sample, individuals abstain from drinking 5 out of 7 days per week, whereas drinkers abstain just over 2 days per week.

[18]Diagnosed at any time prior to wave 3.

[19]This includes individuals for whom it is not known if they were diagnosed prior to the first wave of observation; in wave 1 individuals are asked if they were ever previously diagnosed.

[20]Main controls are: education, age, sex, ethnicity, employment status, urban/rural dummy and GOR.

Table 1: Descriptive Statistics

	count	mean	sd
Demographics			
Education: GCSE or other school qualification	15,974	0.28	0.45
Education: A-level etc	15,974	0.18	0.38
Education: Bachelor's degree or above	15,974	0.44	0.50
Non-white	15,974	0.11	0.31
Female	15,974	0.59	0.49
Age	15,974	47.88	16.09
Health Behaviours (pre-treatment)			
Number of servings of fruit/veg consumed per day	15,974	3.38	1.58
Number of days walked at least 10 minutes, past 4 weeks	15,974	15.42	10.81
Number of days walked at least 30 minutes, past 4 weeks	15,974	9.46	10.12
Does not smoke	15,974	0.81	0.39
Number of cigarettes smoked per day (for all)	15,974	2.39	6.21
Number of cigarettes smoked per day (for smokers)	3,264	11.70	8.93
Does not drink (at least in past 12 months)	15,974	0.12	0.32
Total drinks on heaviest drinking day, past 7 days (for all)	15,974	2.82	3.66
Total drinks on heaviest drinking day, past 7 days (for drinkers)	11,679	3.86	3.78
Number days did not drink, past 7 days (for all)	15,974	5.05	2.08
Number days did not drink, past 7 days (for drinkers)	14,414	2.17	2.08

4 Main Analysis – Impact of Shock

The methodology used for both the main and secondary empirical analysis is kernel matching based on an estimated propensity score, comparing the average treatment effect on the treated (ATT) for treated and control units, using the first-differencing method.[21,22] The propensity score is estimated using the following variables: ethnicity, education, employment status (as a proxy for income), urban/rural dummy, a categorical variable for regions and the previously described initial health risk factors.[23,24] Finally, bootstrapped standard errors are used.[25]

[21] For a discussion on the change from the pre-analysis plan specified regression analysis with matching (using household clustered standard errors) to using kernel matching, see Appendix A.2.

[22] The propensity score estimations satisfy the three necessary conditions: balancing property, unconfoundedness assumption and common support; see Appendix B for details. The matching estimator selection procedure and assessment of the balance are discussed in Appendix C. See Appendix D for an explanation of the different matching approaches, including kernel matching.

[23] Recall, these factors are age, sex, high blood pressure, and the pre-treatment outcomes for fruit and vegetable consumption, smoking and physical activity.

[24] The estimation of the propensity score, aside from being a linear function of these variables, also includes for some of these variables higher order terms and interactions. The higher order terms chosen for inclusion are solely determined by the need for the estimated propensity score to satisfy the balancing property. As such, no behavioural interpretation needs to be given.

[25] Propensity scores are not known, but rather estimated, prior to matching; by default, standard errors from kernel matching do not take this into account. Therefore, the standard errors are bootstrapped.

4.1 Empirical Strategy

The matching analysis uses first-differences for two reasons.[26] First, because the outcome variable is a difference between two periods. Second, the use of fixed effects controls for unobservable time-invariant individual heterogeneity.[27] The first-differences equation takes the form:

$$\Delta Index_{it} = \beta \Delta Shock_{it} + \Delta u_{it}$$

where $\Delta Index_{it} = Index_{i,t=5} - Index_{i,t=2}$ denotes the difference between the lifestyle behaviour index in wave 2 and in wave 5. Similarly, the independent variable, $\Delta Shock_{it} = Shock_{i,t=5} - Shock_{i,t=2}$, is the difference in the shock between wave 2 and wave 5, since in either wave $Shock_i$ is a dummy for whether the health shock has occurred in waves 3 or 4.[28] Finally, $\Delta u_{it} = \Delta u_{i,t=5} - \Delta u_{i,t=2}$ is the differenced error term.

4.2 Results

Table 2 shows the effect of experiencing a shock on changes to the lifestyle index; the shock leads to an increase in the index by 0.41 standard deviations. An increase in the index is interpreted as one or more lifestyle behaviours having become healthier. Therefore, the result suggests that an individual who experiences a health shock improves their lifestyle. A more intuitive understanding of the size of the effect, using a more concrete example, is provided below.

Recall, there are eight lifestyle behaviours where each one is standardized with mean 0 and variance 1 and then combined with equal weight to make up the index. If an individual were to increase the healthiness of one behaviour by one standard deviation, and keep the other behaviours unchanged, then the overall value of the index would increase by one-eighth, 0.125, of a standard deviation. Now, recall the found effect size is 0.41, which can have several interpretations. One interpretation is that approximately three behaviours that make up the index became healthier (the rest remain unchanged), each by about one standard deviation, leading to an index change of about 0.375. Another interpretation, is that less than 3 behaviours improved by, on average, a greater-than-one standard deviation; or vice versa, more than 3 behaviours improved by, on average, a less-than-one standard deviation. Of course, other combinations of small and large (or even negative) 'improvements' in behaviours are also possible.

To get a better understanding for which and how many lifestyle behaviours may be driving the results, an analysis looking at the individual variables is provided in Appendix E.1. The main findings from that decomposition analysis is that the behavioural changes driving the result are an increase in daily fruit and vegetable consumption, a

[26]This paper uses the term 'first-differences' even though the difference is that between waves 2 and 5.
[27]Using first-differences is equivalent to using fixed effects since this analysis only compares two periods.
[28]In wave 2, the dummy takes a value of 0 (no shock) for all individuals; in wave 5 it takes either 0 or 1.

decrease in the daily number of cigarettes smoked and an increase in the probability to quit drinking alcohol. A decomposition by sex is discussed below.

Table 2: ATT of Shock on Change in Lifestyle Index

	Index
Shock	0.408**
	(0.203)
Observations	15,974

Bootstrap standard errors in parentheses, 100 reps
Kernel matching (0.0075 bandwidth)
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Heterogeneous Effects by Sex Table 3 reports the effect of experiencing a shock on the lifestyle index by sex. The effect of a shock for women leads to a 0.46 standard deviation increase in the index, slightly higher than the 0.40 standard deviation increase by men; however, the difference is not statistically significant. A further decomposition of the effect

Table 3: ATT of Shock on Change in Lifestyle Index, By Sex

	Index Female	Index Male
Shock	0.462	0.402
	(0.354)	(0.254)
Observations	9,390	6,584

Bootstrap standard errors in parentheses, 100 reps
Kernel matching, optimal bandwidth for each sex:
0.00046875 for females, 0.005625 for males
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

of the shock on the index's eight lifestyle behaviour components provides insights into the differences in behavioural change between men and women across those components. A short summary of those findings is discussed here (for details see Appendix E.1). Four of the eight lifestyle behaviours have changes that statistically differs from zero – quit smoking, quit drinking alcohol, increase fruit and vegetables consumption and decrease in number of cigarettes smoked. Women improve the healthiness of all four. By contrast, for men the first two behaviour changes (quit smoking and quit drinking) do not differ statistically from zero; the increase in fruit and vegetable consumption is not statistically different between men and women; and similarly the reduction in the number of cigarettes smoked between men and women is not statistically different from zero. Finally, although neither point estimate of number of days not drinking in a week is statistically significant, the point estimate for men is statistically different (in this case larger) than that of the women. The combination of women having larger responses than men for some behaviours

(Smoke and Drink) but smaller for other behaviours (Days) suggests a possible reason for why the overall improvement in the lifestyle index between women and men is not statistically different from zero. It is not immediately clear why women and men differ in their responses across different lifestyle behaviours. However, one reason could be that men and women face different costs when considering changing certain behaviours.

5 Secondary Analysis – Impact of Shock Strength

The secondary analysis follows the same set-up as the main analysis: kernel matching using propensity score estimation with first-differences. It studies the impact of the strength of the shock by analysing two separate cases: 'strong shock only' and 'weak shock only', as defined in Section 3, and then comparing them. The propensity score is estimated separately for each case.[29]

5.1 Empirical Strategy

The analysis of the shock strength is the investigation of the impact of experiencing a shock of a certain strength relative to no shock on the lifestyle index. The strong shock equation takes the form:

$$\Delta Index_{it} = \beta \Delta StrongShock_{it} + \Delta u_{it}$$

where $\Delta Index_{it}$, as defined as in the main analysis, denotes the difference in the lifestyle index. Analogous to the main analysis, the independent variable, $\Delta StrongShock_{it} = StrongShock_{i,t=5} - StrongShock_{i,t=2}$, is the difference in the strong shock outcome between wave 2 and wave 5. Finally, Δu_{it} is also defined as in the main analysis. The weak shock equation takes the form:

$$\Delta Index_{it} = \beta \Delta WeakShock_{it} + \Delta u_{it}$$

where the weak shock equation is the same as the strong shock equation except that the independent variable, $\Delta WeakShock_{it} = WeakShock_{i,t=5} - WeakShock_{i,t=2}$, is the difference in the weak shock outcome between wave 2 and wave 5.

5.2 Results

Strong Shock Only Table 4 shows the positive impact of experiencing a strong shock on the lifestyle index. Recall, a positive increase in the index suggests an overall increase in the healthiness of an individual's lifestyle. The finding that a strong shock leads to an increase in the index by 0.89 standard deviations, statistically significant at the 1% level, suggests that individuals respond to a strong shock by improving their lifestyle.

[29]The matching strategy selection procedure and assessment of the balance are discussed in Appendices C.2 and C.3, respectively. For the strong shock case the optimal kernel bandwidth is 0.00375, for the weak shock case it is 0.00140625. These optimal bandwidths are selected based on what leads to the best balance using the same assessment criteria as the matching strategy selection procedure.

Table 4: ATT of Strong Shock on Change in Lifestyle Index

	Index
Strong Shock	0.892***
	(0.325)
Observations	12,428

Bootstrap standard errors in parentheses, 100 reps
Kernel matching (0.00375 bandwidth)
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Weak Shock Only Table 5 shows the impact of experiencing a weak shock on the lifestyle index. In this case, a weak shock does not lead to a statistically significant increase in the index.[30] This result suggests that individuals do *not* increase their overall lifestyle healthiness after experiencing a weak shock.

Table 5: ATT of Weak Shock on Change in Lifestyle Index

	Index
Weak Shock	0.241
	(0.153)
Observations	12,793

Bootstrap standard errors in parentheses, 100 reps
Kernel matching (0.00140625 bandwidth)
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Comparing Strong and Weak Shocks This section compares the strong and weak shock cases discussed briefly above. The strong shock case only has 94 treated individuals whereas the weak shock case has 458 treated individuals; both cases have over 12,000 control individuals. In the strong shock case, the shock leads to a 0.89 standard deviations increase in the lifestyle index, whereas the weak shock has a point estimate of 0.24, which is not statistically significant. The point estimate of the strong shock is four times larger than that of the weak shock. Although it seems fairly intuitive that the strong shock has a larger effect than the smaller (non-significant) effect found for the weak shock, this paper, to the best of my knowledge, is the first to confirm such intuition empirically. One explanation for this finding is that indeed a strong shock – the diagnosis of a disease – provides a clear(er) signal about an individual's health status compared to a weak shock – the diagnosis of a disease risk factor; therefore an individual is (more) willing or motivated to improve one or more of their lifestyle behaviours. The very large difference in effect sizes found for the index when comparing the shock strengths is reinforced in the decomposition of the index into its component behaviours (see Appendix E.2 for details):

[30] P-value of 0.116.

for all of the behaviours the magnitude of the strong shock point estimate is found to be either greater (i.e. healthier) or not statistically different from the weak one.

6 Conclusion

This paper finds that individuals improve their overall lifestyle, measured using a lifestyle index, after experiencing a health shock by improving some of their lifestyle behaviours. It also empirically confirms the intuition that some individuals make large lifestyle changes when the shock experienced is severe, interpreted as receiving a clearer signal about their health status. Examples of such shocks include experiencing a heart attack or diabetes diagnosis. There is little evidence of lifestyle changes when only experiencing a weak shock, such as being diagnosed with a risk factor such as high blood pressure. Furthermore, although the overall effect on lifestyle changes between men and women is of similar size, the heterogeneity lies in the changes made to the behaviours that make up the overall lifestyle index. One possible reason for these findings is that individuals may face heterogeneous costs when considering behavioural change, both across individuals but also across behaviours within an individual. These findings suggest that policy-makers interested in fostering certain lifestyle changes must take into account both what lifestyle behaviour they are trying to change and which individuals they are targetting, as heterogeneous costs may play a role in the effectiveness of policies or interventions being considered. Future research can explore what are the different costs faced by individuals when considering lifestyle changes, whether they be financial, social or personal, with the latter including personality characteristics.

References

American Heart Association. (2014, August 14). Alcohol and Heart Health. *American Heart Association*. Retrieved from https://www.heart.org/en/healthy-living/healthy-eating/eat-smart/nutrition-basics/alcohol-and-heart-health

American Heart Association. (2017, May 31). What is Cardiovascular Disease? *American Heart Association*. Retrieved from https://www.heart.org/en/health-topics/consumer-healthcare/what-is-cardiovascular-disease

Atkin, D. (2016). The caloric costs of culture: Evidence from Indian migrants. *American Economic Review*, *106*(4), 1144–81.

Bartiaux, F. (2008). Does environmental information overcome practice compartmentalisation and change consumers' behaviours? *Journal of Cleaner Production*, *16*(11), 1170–1180.

Belot, M., James, J., & Spiteri, J. (2019). *What you don't know (or think about) can't hurt you: An experiment on dietary choices.* Mimeo.

Bhalotra, S., Delavande, A., Fisher, P., & James, J. (2020). The impact of a personalised blood pressure warning on health outcomes and behaviours. *Institute of Social and Economic Research (ISER) Working Paper Series; No. 2020-2*. Retrieved from https://www.iser.essex.ac.uk/research/publications/working-papers/iser/2020-02.pdf

Brownell, K. D., & Cohen, L. R. (1995). Adherence to dietary regimens 2: Components of effective interventions. *Behavioral Medicine*, *20*(4), 155–164.

Buttar, H. S., Li, T., & Ravi, N. (2005). Prevention of cardiovascular diseases: Role of exercise, dietary interventions, obesity and smoking cessation. *Experimental & Clinical Cardiology*, *10*(4), 229.

Carrera, M., Hasan, S. A., & Prina, S. (2020). Do health risk assessments change eating habits at the workplace? *Journal of Economic Behaviour and Organization*, *172*, 236–246.

Chong, S., Ding, D., Byun, R., Comino, E., Bauman, A., & Jalaludin, B. (2017). Lifestyle changes after a diagnosis of type 2 diabetes. *Diabetes Spectrum*, *30*(1), 43–50.

Chou, C.-H., Hwang, C.-L., & Wu, Y.-T. (2012). Effect of exercise on physical function, daily living activities, and quality of life in the frail older adults: A meta-analysis. *Archives of Physical Medicine and Rehabilitation*, *93*(2), 237–244.

Condon, C., & McCarthy, G. (2006). Lifestyle changes following acute myocardial infarction: Patients perspectives. *European Journal of Cardiovascular Nursing*, *5*(1), 37–44.

Craig, C., Tudor-Locke, C., & Bauman, A. (2006). Twelve-month effects of canada on the move: A population-wide campaign to promote pedometer use and walking. *Health Education Research*, *22*(3), 406–413.

Crombie, I. K., Irvine, L., Williams, B., McGinnis, A. R., Slane, P. W., Alder, E. M., & McMurdo, M. E. (2004). Why older people do not participate in leisure time physical activity: A survey of activity levels, beliefs and deterrents. *Age and Ageing*, *33*(3), 287–292.

Durkin, S., Brennan, E., & Wakefield, M. (2012). Mass media campaigns to promote smoking cessation among adults: An integrative review. *Tobacco Control*, *21*(2), 127–138.

Edwards, R. D. (2018). If my blood pressure is high, do I take it to heart? Behavioral effects of biomarker collection in the health and retirement study. *Demography*, *55*(2), 403–434.

Esselstyn, C. B., Jr., Ellis, S. G., Medendorp, S. V., & Crowe, T. D. (1995). A strategy to arrest and reverse coronary artery disease: A 5-year longitudinal study of a single physician's practice. *Journal of Family Practice*, *41*(6), 560–568.

Fassier, P., Zelek, L., Lécuyer, L., Bachmann, P., Touillaud, M., Druesne-Pecollo, N., ... Touvier, M. (2017). Modifications in dietary and alcohol intakes between before and after cancer diagnosis: Results from the prospective population-based NutriNet-Santé cohort. *International Journal of Cancer*, *141*(3), 457–470.

Fitgerald, J. T., Singleton, S. P., Neale, A. V., Prasad, A. S., & Hess, J. W. (1994). Activity levels, fitness status, exercise knowledge, and exercise beliefs among healthy, older african american and white women. *Journal of Aging and Health*, *6*(3), 296–313.

Flay, B. R. (1987). Mass media and smoking cessation: A critical review. *American Journal of Public Health*, *77*(2), 153–160.

Fleming, M. F., Manwell, L. B., Barry, K. L., Adams, W., & Stauffacher, E. A. (1999). Brief physician advice for alcohol problems in older adults: A randomized community-based trial. *Journal of Family Practice*, *48*, 378–386.

Garrido, M. M., Kelley, A. S., Paris, J., Roza, K., Meier, D. E., Morrison, R. S., & Aldridge, M. D. (2014). Methods for constructing and assessing propensity scores. *Health Services Research*, *49*(5), 1701–1720.

Heikkinen, H., Patja, K., & Jallinoja, P. (2010). Smokers' accounts on the health risks of smoking: Why is smoking not dangerous for me? *Social Science & Medicine*, *71*(5), 877–883.

Hut, S., & Oster, E. (2018). Changes in household diet: Determinants and predictability (Working Paper 24892). *NBER Working Paper Series*.

Kelly, M. P., & Barker, M. (2016). Why is changing health-related behaviour so difficult? *Public Health*, *136*, 109–116.

Kling, J. R., Liebman, J. B., & Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, *75*(1), 83–119.

Lanza, E., Schatzkin, A., Daston, C., Corle, D., Freedman, L., Ballard-Barbash, R., ... the PPT Study Group (2001). Implementation of a 4-y, high-fiber, high-fruit-and-vegetable, low-fat dietary intervention: Results of dietary changes in the polyp prevention trial. *American Journal of Clinical Nutrition*, *74*(3), 387–401.

Larimer, M. E., & Cronce, J. M. (2002). Identification, prevention and treatment: A review of individual-focused strategies to reduce problematic alcohol consumption by college students. *Journal of Studies on Alcohol. Supplement,*(14), 148–163.

Logie-MacIver, L., Piacentini, M., & Eadie, D. (2012). Using qualitative methodologies to understand behaviour change. *Qualitative Market Research: An International Journal*, *15*(1), 70-86.

Mathis, K., & Steffen, A. D. (2015). From rational choice to behavioural economics. In *European perspectives on behavioural law and economics* (pp. 31–48). Springer.

McMacken, M., & Shah, S. (2017). A plant-based diet for the prevention and treatment of type 2 diabetes. *Journal of Geriatric Cardiology*, *14*(5), 342.

Ornish, D., Scherwitz, L. W., Billings, J. H., Gould, K. L., Merritt, T. A., Sparler, S., ... Brand, R. J. (1998). Intensive lifestyle changes for reversal of coronary heart disease. *Journal of the American Medical Association*, *280*(23), 2001–2007.

Oster, E. (2018). Diabetes and diet: Purchasing behavior change in response to health information. *American Economic Journal: Applied Economics*, *10*(4), 308–48.

Rizza, W., Veronese, N., & Fontana, L. (2014). What are the roles of calorie restriction and diet quality in promoting healthy longevity? *Ageing Research Reviews*, *13*, 38–45.

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, *39*(1), 33–38.

Rubin, D. B. (2001). Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, *2*(3-4), 169–188.

Sanderson, S. C., Waller, J., Jarvis, M. J., Humphries, S. E., & Wardle, J. (2009). Awareness of lifestyle risk factors for cancer and heart disease among adults in the UK. *Patient and Counseling*, *74*(2), 221–227.

Stockwell, T., Zhao, J., Panwar, S., Roemer, A., Naimi, T., & Chikritzhs, T. (2016). Do "moderate" drinkers have reduced mortality risk? A systematic review and meta-analysis of alcohol consumption and all-cause mortality. *Journal of Studies on Alcohol and Drugs*, *77*(2), 185–198.

Strecher, V. J., Kreuter, M., Den Boer, D.-J., Kobrin, S., Hospers, H. J., & Skinner, C. S. (1994). The effects of computer-tailored smoking cessation messages in family practice settings. *Journal of Family Practice*, *39*(3), 262–270.

World Health Organization. (2018, May 24). *The top 10 causes of death.* Retrieved from https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death

Worsley, A. (2002). Nutrition knowledge and food consumption: Can nutrition knowledge change food behaviour. *Asia Pacific Journal of Clinical Nutrition*, *11*(S6), S579–S585.

Young, D. R., Haskell, W. L., Taylor, C. B., & Fortmann, S. P. (1996). Effect of community health education on physical activity knowledge, attitudes, and behavior: The Stanford Five-City Project. *American Journal of Epidemiology*, *144*(3), 264–274.

Your disease risk. (n.d.). *Siteman Cancer Center*. Retrieved 25 March, 2019, from https://siteman.wustl.edu/prevention/ydr/

Zhao, M., Konishi, Y., & 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.

Appendices

A Deviations from Pre-Analysis Plan

A.1 Additional Outcome Variable

My pre-analysis plan mentions seven health behaviour outcome variables, whereas this paper uses eight. The additional outcome variable captures the extensive margin of alcohol consumption: a dummy for if an individual is a drinker or not. Furthermore, the original pre-analysis plan did not include that these behaviours would be summarized into a lifestyle index. An index was introduced to have just a single outcome variable.

The reason for including this additional alcohol consumption variable is the lack of an extensive margin variable for alcohol consumption. In the case of the smoking variables there is one intensive and one extensive margin variable. In the case of alcohol consumption, there were only two intensive margin variables. However, I realized it is important to have both an extensive and an intensive margin variable for this behaviour as it is possible that individuals behave heterogeneously after a shock, changing one of these margins but likely not both.

Finally, I keep both the intensive margin variables for alcohol consumption: the number of days an individual did not drink alcohol in the past seven days and the total number of drinks an individual consumed on the heaviest drinking day in the past seven days. The reason is that the literature on harmful alcohol consumption and behaviours suggests that when measuring the most harmful aspects of alcohol consumption on health (i.e. binge drinking) it is *both* the intensity with which an individual drinks in a given period of time (about two hours) and how often per week an individual drinks that matter.

A.2 Changes to Analysis Approach

The pre-analysis plan specified I would match the sample and then do regression analysis on that sample, either matching using stratification or NN 1:1. However, given that kernel matching provides the best balance, I use it instead. Therefore, instead of matching treated with control units and then running standard regression analyses using the matched sample, I use the matching commands directly. This means that the matching program will match the treated and controls using the selected strategy and then provide the ATT (or ATE, if desired). The consequence of changing the technical approach to regression analysis it is no longer possible to include household clustered standard errors (something I said I would do in my pre-analysis plan).[31] The clustered standard errors

[31] In theory, the inclusion of household clustered standard errors is considered important because there are individuals in the sample that live in the same household together. It is important to take into account that such individuals may not be identically and independently distributed (i.i.d.). One example of this possible dependence is that if one household member experiences a shock another household member

could be included in a regular regression analysis, but cannot be when using matching analysis commands directly. However, the use of a different matching strategy and/or different standard errors have no effect on the findings, as discussed in Appendix D.1.

A.3 No BMI or Blood Cholesterol in Propensity Score

In the pre-analysis plan it was written that if more than half the sample had the BMI and blood cholesterol variables non-missing a robustness check would be done to see if there is any impact of their inclusion on the final results. However, for the blood cholesterol variable only 32% of the sample is non-missing and for BMI only 48% is non-missing. Therefore, these variables are not included in any analysis or robustness check.

A.4 Multiple Hypothesis Correction

The pre-analysis plan stated the use of the Holm-Bonferroni procedure. The reasoning was that the more powerful Hochberg procedure only holds under non-negative dependence. However, that was a mistake, as there is no dependence between any of the outcomes. Therefore, the Benjamin-Hochberg correction procedure, henceforth Hochberg correction, is used instead; it is uniformly more powerful than the Holm-Bonferroni procedure.

A.5 No Splitting of Pooled Shock Variables

The separating of the pooled shocks, as a robustness check, for either the main analysis or secondary analysis is not performed, though it is mentioned in the pre-analysis plan. The reason is insufficient power, which was the main reason to pool to begin with.

B Propensity Score Estimation Conditions

The propensity score is the probability an individual is treated given a set of selected observables; for proper application two lemmas must hold. First, the balancing property: observations with the same propensity score have the same distribution of observable covariates independently of treatment status. Second, the unconfoundedness assumption: the assignment to treatment is independent given the propensity score (Rosenbaum & Rubin, 1985). Finally, there must be common support between treated and control units.

To ensure that the balancing property is satisfied, after the propensity score is estimated, the propensity score estimation program takes the full sample of treated and controls, sorts the individuals by their estimated propensity score, and divides them into bins such that within each bin the mean propensity score is not statistically different between the treated and controls groups. Similarly, the balancing property also requires

may adjust their behaviour, even without experiencing a shock themselves.

that the mean of each covariate used in the estimation of the propensity score is balanced within each bin between treated and control groups. The exact specification for the propensity score estimation is chosen such as to meet these requirements. Finally, common support is confirmed by looking at the overlap in estimated propensity scores between treated and control units.

Main Analysis The balancing property is satisfied using seven bins. It is clear from Figure 1 that there is common support between treated and control units, as none of the treated units are marked as 'off-support'. Common support is further verified by Table 6, which shows the descriptive statistics of the estimated propensity score for the full sample,

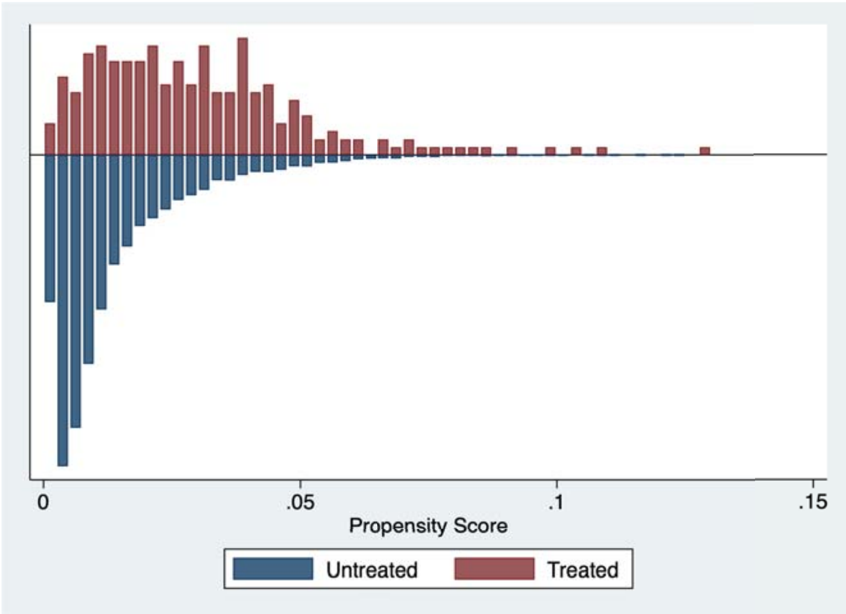

Figure 1: Estimated Propensity Score Histogram (Full Sample)

only the treated and only the controls. It is important to note that both the minimum and the maximum for the treated group falls within the minimum and maximum of the control group. Therefore, common support for all treated units is further verified.

Table 6: Descriptive Statistics of Estimated Propensity Score (Full Sample)

Sample	Count	Mean	SD	Min	Max
Full Sample	22,327	0.0158	0.0155	0.00055	0.1433
Only Treated	354	0.0313	0.0224	0.00113	0.1282
Only Controls	21,973	0.0156	0.0153	0.00055	0.1433

Strong Shock Only The estimated propensity score for the 'strong shock only' case satisfies the balancing property and has common support, shown in Figure 2, since once again there are no treated units marked as off-support. Table 7, analogous to Table 6, shows

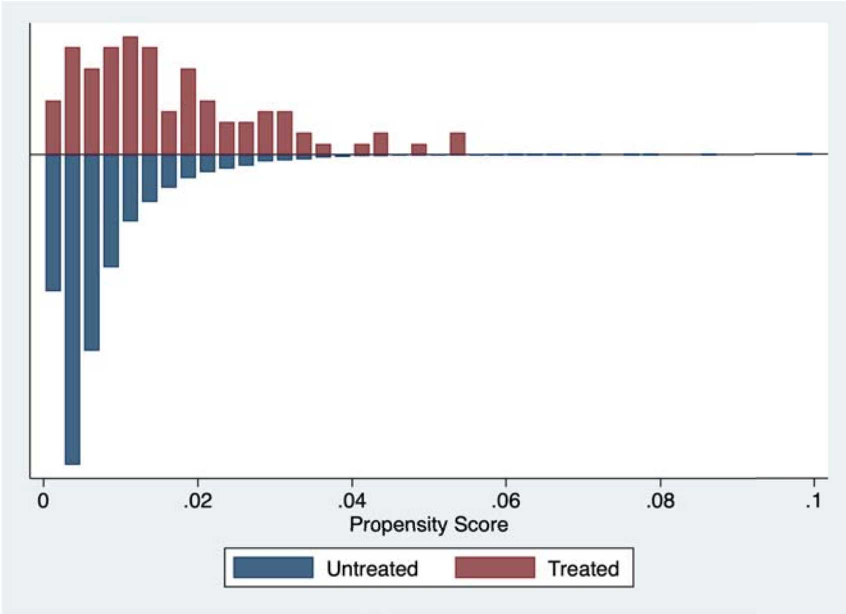

Figure 2: Estimated Propensity Score Histogram (Strong Shock Only)

the descriptive statistics of the estimated propensity score for the three (sub)samples. Here as well, both the minimum and the maximum for the treated group fall within the minimum and maximum of the control group, thereby reinforcing that all treated units fall within the common support.

Table 7: Descriptive Statistics of Estimated Propensity Score (Strong Shock Only)

Sample	Count	Mean	SD	Min	Max
Full Sample	17,285	0.0089	0.0087	0.00062	0.1064
Only Treated	154	0.0175	0.0140	0.00109	0.0782
Only Controls	17,131	0.0088	0.0086	0.00062	0.1064

Weak Shock Only The estimated propensity score for the 'weak shock only' variable satisfies the balancing property; and it has common support for all but one treated observation, as shown in Figure 3. Table 8 shows the descriptive statistics of the estimated propensity score for the full, treated-only, and control-only samples. Again, note that both the minimum and the maximum for the treated group fall within those of the control group, which reinforces that most treated units fall within the common support.

C Matching Strategy – Selection and Assessment

Due to the nature of matching, the propensity score is estimated and the matching performed *prior* to the analysis of the outcome variables. As a result, it is possible and en-

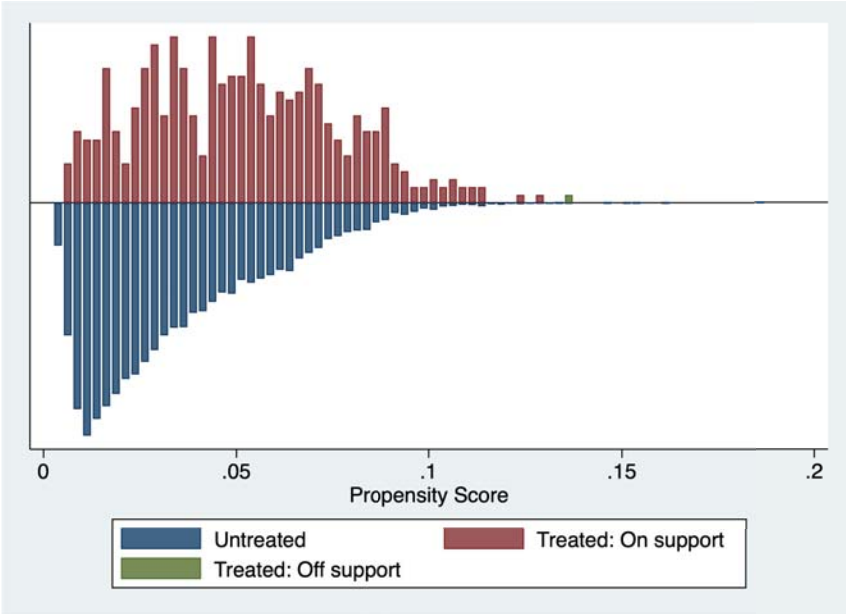

Figure 3: Estimated Propensity Score Histogram (Weak Shock Only)

Table 8: Descriptive Statistics of Estimated Propensity Score (Weak Shock Only)

Sample	Count	Mean	SD	Min	Max
Full Sample	17,758	0.0353	0.0238	0.00240	0.1870
Only Treated	627	0.0514	0.0261	0.00627	0.1488
Only Controls	17,131	0.0347	0.0235	0.00240	0.1870

couraged to try several matching possibilities – such as nearest neighbour with or without replacement and with one or more controls, and radius or kernel with different bandwidths – and to run a range of balance diagnostics to assess which matching strategy leads to the best balance; this strategy is then used for the analysis of the outcomes (Garrido et al., 2014). The goal behind matching is to ensure that treated and control units are as similar as possible across observable covariates. Below are the selection and assessment criteria for both the main and secondary analyses.

C.1 Main Analysis

Selecting Matching Strategy with Best Balance Table 9 reports several possible matching strategies. In all cases, there are 232 treated individuals with a differing number of individuals used as controls depending on the matching strategy.[32] As is shown in Table 9, the kernel matching strategy, and its corresponding bandwidth, reported in the table is the most balanced of the matches among different kernel bandwidths tried. The kernel matching has the lowest mean and median percentage standardized difference in covariates

[32]Pairs are created for those individuals who are not missing the index outcome variable. There are only 232 treated observations and not the 354 reported in Table 6. This is because the 'missing' individuals are missing one or more of the eight lifestyle index outcome components in either wave 2 or 5.

and the lowest Pseudo R^2. Furthermore, both the Rubin's B and Rubin's R fall within their desired cut-offs or ranges: below 25% and between [0.5,2], respectively. Therefore, this kernel matching strategy is used in the analysis. A description of each summary measure of match quality and the characteristics of a good match are discussed next. A comparison of the sensitivity of the results to different matching strategies is provided in Appendix D.1.

Table 9: Summary Measures of Match Quality for Original, Matched and Weighted Samples

Sample Type	Total Obs.	Total Treated	Total Controls	Pseudo R^2	Mean Standardized Difference (%)	Median Standardized Difference (%)	Rubin's B	Rubin's R
Original Sample	15,974	232	15,742	0.078	14.8	8.1	92.9*	0.92
NN 1:1 no replace.	464	232	232	0.032	5.4	3.4	42.4*	1.07
NN 1:1 with replace.	458	232	226	0.031	5.2	3.4	41.6*	1.01
Radius	15,974	232	15,742	0.013	4.3	3.3	26.6*	0.81
Kernel	**15,974**	**232**	**15,742**	**0.011**	**3.6**	**2.6**	**24.3**	**0.78**

Note: NN: Nearest Neighbour matching. Radius method has a caliper of 0.01. NN 1:1 with caliper omitted since match quality results are the same as the no caliper case. Kernel matching uses a bandwidth of 0.0075. If B>25% or R outside [0.5,2], marked with *.

Summary Measures of Match Quality Table 10 provides several summary measures of the overall balance of the variables used to estimate the propensity score and create the match, for both the unmatched and matched samples. The first column, the Pseudo-R^2, is the estimate from a probit of the propensity score equation. The closer the value is to zero, the more the variables used to estimate the propensity score no longer have predictive power for the shock, which implies better balance. Similarly, the second column shows the p-value for the likelihood ratio test that all covariates used for the estimation are jointly insignificant. Both these columns suggest that the match between treated and controls is quite balanced. The third and fourth columns show that the mean and median percentage standardized difference between the treatment and controls groups have been reduced for a large extent by matching, respectively; the mean percentage decreases from 14.8 to 3.6, the median from 8.1 to 2.6. A mean percentage standardized difference of less than 10% is considered a good quality match, which is what is found. Finally, the last two columns are summary measures of matching quality suggested by Rubin (2001). Rubin's B is the absolute standardized difference of the means of the propensity score of the treated and control group. Rubin's R is the variance ratio of the treated and control groups' propensity score. Rubin (2001) specifies that groups are sufficiently balanced when the Rubin's B is less than 25%; similarly, the Rubin's R should be between 0.5 and 2. Both Rubin measures are within the desired range, indicating a good match. Summarizing, overall Table 10 suggests the match is well balanced. A graphical interpretation of balance before and after matching at the individual covariate level is shown and discussed below.

Table 10: Summary Measures of Match Quality

Sample	Pseudo R^2	P	Likelihood ratio test p-value	Mean Standardized Difference (%)	Median Standardized Difference (%)	Rubin's B	Rubin's R
Unmatched	0.078		0.000	14.8	8.1	92.9*	0.92
Matched	0.011		1.000	3.6	2.6	24.3	0.78

Note: * if B>25%, R outside [0.5; 2]

Match Quality – Individual Covariates Figure 4 shows graphically the percent standardized difference (bias) for the covariates used in the propensity score estimation, both before and after matching. Overall, the figure shows that, except for ethnicity, the bias for all variables shown decreases with matching, often quite substantially.[33] In the case of ethnicity, the treated and untreated groups are not statistically different from each other in both the unmatched and matched cases, which is therefore not a concern.

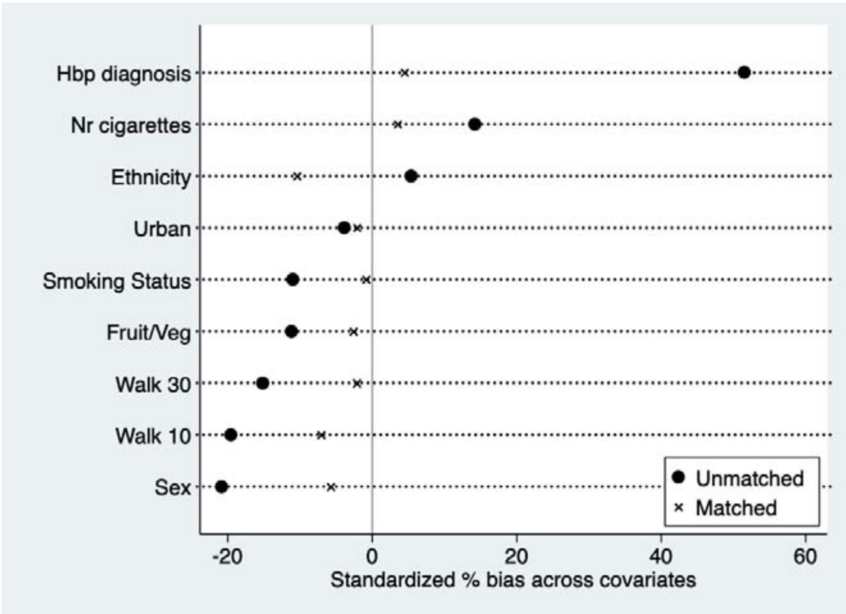

Figure 4: Percent Standardized Difference (Bias) Across Propensity Score Estimation Covariates Before and After Matching

C.2 Secondary Analysis – Strong Shock Only

Selection Matching Strategy with Best Balance Among the reported matching strategies, in Table 11, the kernel matching strategy has the lowest mean and median percentage standardized difference in covariates and the lowest Pseudo R^2. For the Rubin's B, although it does not fall within the desired 25% cut-off, the kernel strategy has the lowest Rubin's B value of all the matching strategies. Finally, the Rubin's R falls within the desired range of [0.5,2]. Although the kernel match is still not very well-balanced —

[33]Non-binary categorical variables are not shown for reasons of readability.

as suggested by both the Rubin's B and the higher values of the other measures compared to both the main analysis match (shown previously) and the weak shock analysis match (shown subsequently) — the kernel strategy is still the best choice of the available strategies and therefore is used for this strong shock analysis.

Table 11: Summary Measures of Match Quality for Original, Matched and Weighted Samples (Strong Shock Only)

Sample Type	Total Obs.	Total Treated	Total Controls	Pseudo R^2	Mean Standardized Difference (%)	Median Standardized Difference (%)	Rubin's B	Rubin's R
Original Sample	12,428	94	12,334	0.063	13.7	8.5	84.9*	1.27
NN 1:1 no replace.	188	94	94	0.079	9.6	9.1	67.7*	0.91
NN 1:1 with replace.	186	94	92	0.081	10.4	9.1	68.8*	0.84
Radius	12,418	94	12,324	0.036	8.4	7.7	44.9*	0.77
Kernel	**12,414**	**94**	**12,320**	**0.025**	**5.7**	**4.7**	**37.2***	**0.65**

Note: NN: Nearest Neighbour matching. Radius method has a caliper of 0.01. NN 1:1 with caliper omitted since match quality results are the same as the no caliper case. Kernel matching uses a bandwidth of 0.00375. If B>25% or R outside [0.5,2], marked with *.

Summary Matching Quality Assessment Table 12 provides, for the strong shock case, several summary measures of the overall balance of the variables used to estimate the propensity score and create the match, for both the unmatched and matched samples. The measures are the same as those described previously for the main analysis case. In this strong shock case, the measures indicate that matching improves the balance between treated and controls. For example, the average percentage standardized differences of the mean and median for the covariates from the propensity score estimation decrease: the mean from 13.7 to 5.7 and the median from 8.5 to 4.7. However, the Rubin's B measure suggests that although the matched sample is an improvement over the unmatched one, it is still not very well-balanced as the value (37.2%) falls above the 25% threshold. This less good match is likely in part attributable to the relatively small sample size of the treated for this strong shock case compared to the other two cases. Nevertheless, the measures still suggest that the matched sample has more balance than the unmatched sample and therefore is used for this 'strong shock only' analysis. A graphical interpretation of balance before and after matching at the individual covariate level is below.

Table 12: Summary Measures of Match Quality (Strong Shock Only)

Sample	Pseudo R^2	Likelihood ratio test p-value	Mean Standardized Difference (%)	Median Standardized Difference (%)	Rubin's B	Rubin's R
Unmatched	0.063	0.000	13.7	8.5	84.9*	1.27
Matched	0.025	1.000	5.7	4.7	37.2*	0.65

Note: * if B>25%, R outside [0.5;2]

Match Quality – Individual Covariates Figure 5 shows graphically the percent standardized difference (bias) for the covariates used in the propensity score estimation for the strong shock, both before and after matching. In the strong shock case, the figure shows the matching is less successful in reducing the difference between the unmatched treated and unmatched control groups (i.e. the bias). For age and the interaction between sex and employment status matching reduces the bias quite drastically. However, for most other variables shown matching does not affect the bias much.[34] However, none of the variables are statistical different between the treated and untreated groups after matching.

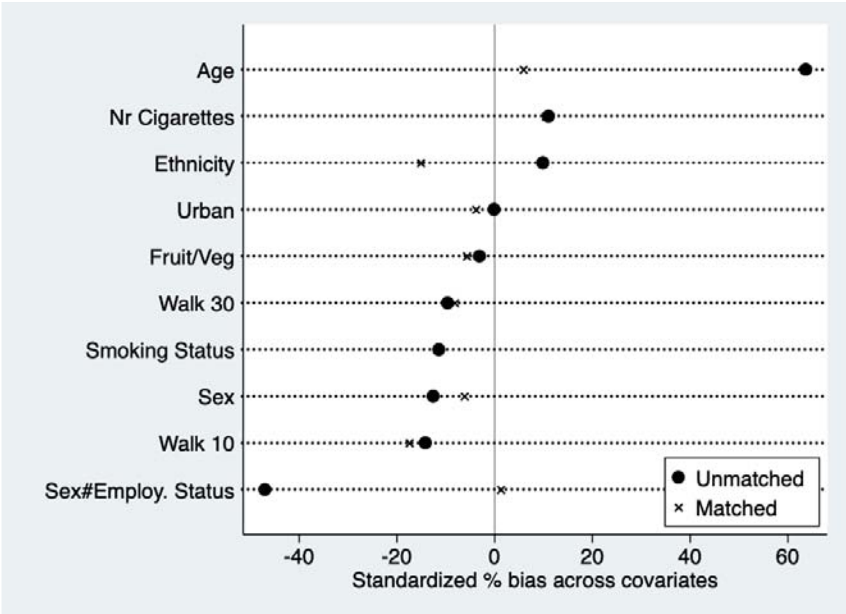

Figure 5: Percent Standardized Difference (Bias) Across Propensity Score Estimation Covariates Before and After Matching (Strong Shock Only)

C.3 Secondary Analysis – Weak Shock Only

Selection Matching Strategy with Best Balance Among the reported matching strategies in Table 13 kernel matching has the lowest mean and median percentage standardized difference in covariates and one of the lowest Pseudo R^2. Furthermore, both the Rubin's B and Rubin's R fall within their desired cut-offs or ranges (below 25% and between [0.5,2], respectively), and the kernel matching strategy has the 'best' value for either measure compared to the other matching strategies. Therefore, this kernel matching strategy is used in the analysis.

Summary Match Quality Assessment Table 14 provides for the weak shock case several summary measures of the overall balance of the variables used to estimate the propensity score and create the match, for both the unmatched and matched samples. The measures are the same as those described previously and all suggest a good match. Summarizing,

[34]Once again, non-binary categorical variables are not shown for readability reasons.

Table 13: Summary Measures of Match Quality for Original, Matched and Weighted Samples (Weak Shock Only)

Sample Type	Total Obs.	Total Treated	Total Controls	Pseudo R^2	Mean Standardized Difference (%)	Median Standardized Difference (%)	Rubin's B	Rubin's R
Original Sample	12,793	459	12,334	0.052	12.0	6.3	71.2*	0.77
NN 1:1 no replace.	918	459	459	0.010	3.3	2.8	23.2	1.31
NN 1:1 with replace.	898	459	439	0.010	3.3	2.9	23.6	1.34
Radius	12,789	458	12,330	0.005	2.9	2.4	17.0	1.08
Kernel	**12,632**	**458**	**12,174**	**0.005**	**2.6**	**2.1**	**16.6**	**1.04**

Note: NN: Nearest Neighbour matching. Radius method has a caliper of 0.01. NN 1:1 with caliper omitted since match quality results are the same as the no caliper case. Kernel matching uses a bandwidth of 0.00140625. If B>25% or R outside [0.5,2], marked with *.

overall Table 14 suggests a well-balanced match. A graphical interpretation of balance before and after matching at the individual covariate level is below.

Table 14: Summary Measures of Match Quality (Weak Shock Only)

Sample	Pseudo R^2	Likelihood ratio test p-value	Mean Standardized Difference (%)	Median Standardized Difference (%)	Rubin's B	Rubin's R
Unmatched	0.052	0.000	12.0	6.3	71.2*	0.77
Matched	0.005	1.000	2.6	2.1	16.6	1.04

Note: * if B>25%, R outside [0.5;2]

Match Quality – Individual Covariates Figure 6 shows graphically the percent standardized difference (bias) for the covariates used in the propensity score estimation for the weak shock, both before and after matching. Overall, the figure shows that except for Fruit/Veg the bias for all variables shown decreases with matching, often quite substantially.[35] In the case of Fruit/Veg, the treated and untreated groups are not statistically different from each other in either the unmatched and matched cases, therefore no concern.

D Matching and Weighting Strategies

Stratification Stratification takes the full sample of treated and controls, sorts the individuals by estimated propensity score and then splits them into bins such that within each bin the mean propensity score is the same for the treated and controls groups. The analysis (the effect calculation) is performed for each bin with the assumption that within each bin individuals are relatively similar.

Nearest Neighbour (NN) 1:1 Matching without Replacement Nearest neighbour 1:1 matching can be implemented with or without replacement.[36] In the case of no replacement,

[35] Once again, non-binary categorical variables are not shown for readability reasons.

[36] The use of a caliper to restrict the distance from which the nearest control is selected has no effect.

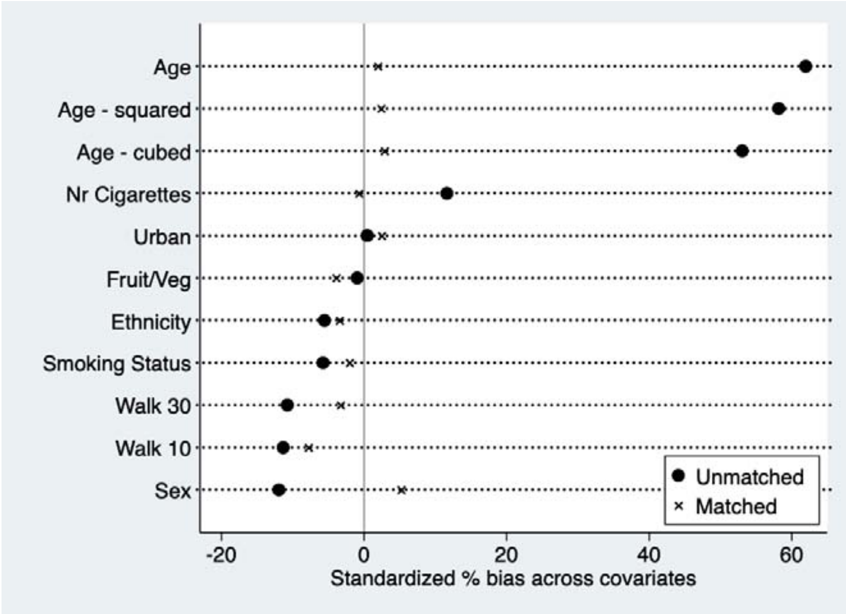

Figure 6: Percent Standardized Difference (Bias) Across Propensity Score Estimation Covariates Before and After Matching (Weak Shock Only)

for each treated individual the control with the closest propensity score is selected, except where that control has already been assigned to another treated individual, in that case the next nearest control unit is selected.

NN 1:1 Matching with Replacement Matching 1:1 with replacement is similar to matching 1:1 without replacement except that the control is returned to the pool each time, such that it can be used as a control for another treated individual. This means that for each treated individual the control with the closest propensity score is always selected.

NN 1:N Matching, N>1 The case of matching more than one control to the treated unit (1:N) is only possible with replacement. It is similar to matching 1:1 with replacement except that instead of matching a treated individual with its single nearest control, the treated individual is matched to its N nearest controls.

Kernel Matching Kernel matching uses all observations within the common support, taking a weighted average of the inverse of the distance between each treated and control unit. In other words, control units nearer to the treated unit receive more weight.

Radius Matching In radius matching for each treated unit, within a specified radius, all control are used and assigned equal weight regardless of their (propensity score) distance to the treated unit.

D.1 Comparing Matching Strategies and Standard Errors

Table 15 compares several matching strategies as well as the different options for standard errors. The main purpose of this table is to show that neither the matching strategy nor the choice for standard errors has a significant impact on the final findings. None of the estimates are statistically significantly different from one another at the 5% level. In each column the estimate reported is the impact of the main analysis shock on the index, first differenced.

Column 1 presents the regression analysis matching using stratification and using regular robust standard errors. Column 2 presents the same as column 1 except that it clusters standard errors at the household level. From these two columns it is clear that such household-level clustering of standard errors has no effect on changing the reported standard errors.

When considering the matching strategy directly – rather than including matching as part of a regression analysis – there are several aspects to decide upon. First, the exact command to execute the matching (*psmatch*, *nnmatch* or *psmatch2*). Second, whether to provide the matching program/command with the propensity score directly (pscore) or whether to provide it with the covariates used to calculate the propensity score (cov) and allow the program to calculate its own propensity score that it will subsequently use to match on.[37] Third, decide on what kind of standard errors to use, both in terms of what the default standard errors are for each matching program/command but also how the standard errors can be adjusted to account for the number of matches used (if using nearest neighbour, N>1) and/or to use bootstrap standard errors, where applicable. For *psmatch* and *nnmatch* the default is robust Abadie-Imbens (A-I) standard errors, which take into account that the propensity score is estimated rather than known.

In columns 3-5 matching is done 1:1 using A-I standard errors. Column 3 uses the *psmatch* command and provides the matching program with the previously calculated propensity score directly. Column 4 uses the *psmatch* command but this time the program is provided with the covariates to calculate its own propensity score prior to matching. Column 5 uses the same approach as column 4, but this time using the *nnmatch* command. The *nnmatch* command only allows the provision of the covariates to calculate its own propensity score prior to matching, it does not allow the input of a previously calculated propensity score.

Finally, columns 6 and 7 use the *psmatch2* command, which allows the use of the kernel matching strategy. Column 6 shows the results using the default standard errors while column 7 shows the same results but using bootstrap standard errors.

[37] One of the main advantages of estimating the propensity score using the *pscore* matching program, over other programs, is that as part of the propensity score estimation this program also checks and requires that the estimated propensity score, as well as the covariates used to do such an estimation, are balanced among propensity score bins of treated and control groups; something that would otherwise have to be checked manually.

Summarizing, from this table neither the exact choice of matching strategy nor the standard error adjustment has any significant effect on the estimates or standard errors displayed. In general the effect has a point estimate of approximately 0.41. The chosen specification is *psmatch2* with previously calculated propensity score, bandwidth of 0.0075 and bootstrap standard errors, corresponding to column 7. The choice of optimal matching strategy was discussed in Appendix C.1.

Table 15: Shock on Index – Several Matching Strategies and Standard Errors

	(1) reg - strat stratification - no hh cl se	(2) reg - strat stratification - hh cl se	(3) psmatch - pscore NN 1:1 - A-I se	(4) psmatch - cov NN 1:1 - A-I se	(5) nnmatch - cov NN 1:1 - A-I se	(6) psmatch2 - pscore kernel 0.0075 bandwidth se	(7) psmatch2 - pscore kernel 0.0075 bandwidth bootstrap se (100reps)
Shock	0.405**	0.405**					
	(0.203)	(0.203)					
ATT			0.421**	0.690***	0.362	0.408**	0.408**
			(0.203)	(0.248)	(0.262)	(0.204)	(0.203)
Constant	0.017	0.017					
	(0.050)	(0.051)					
Observations	15,974	15,974	15,974	15,974	15,974	15,974	15,974

Standard errors in parentheses. See column headers for type of standard error (se). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

E Decomposition of Shock Effect – Separate Behaviours

When analysing the decomposition of the effect of the shock on the lifestyle index's eight components, a correction is needed to control the Type I error rate given the testing of multiple hypotheses. To this end the Hochberg correction is used; it generates a critical value for each outcome to test if that outcome is statistically significant at the 5% significance level. The critical values are calculated using a 0.1 false discovery rate (FDR).

E.1 Main Analysis Results

Table 16 shows the impact of experiencing a shock on the eight components of the index, with and without the Hochberg multiple hypothesis test correction.[38] All the statistically significant point estimates have the expected positive sign, which suggests the shock leads to improvements in lifestyle behaviours. The interpretation of the table and components is as follows, a positive effect is an increase in the quantity of a healthy behaviour, a decrease in the quantity of an unhealthy behaviour, or an increase in the probability of quitting an unhealthy behaviour between wave 2 and wave 5.

The difference in the average number of servings of fruit and vegetables consumed per day (Fruit/Veg) is statistically significant with an increase of 0.24 servings. The number of days per month walked at least 10 minutes per day (Walk 10) or at least 30 minutes per day (Walk 30) are not statistically significant. The negative point estimates

[38]The statistically significant point estimates remain so even after applying the Hochberg correction.

Table 16: ATT of Shock on First-Difference Lifestyle Index Components

	Fruit/Veg	Walk 10	Walk 30	Smoke	Nr Cigs	Drink	Heavy	Days
Shock	0.238**+	-0.837	-1.080	0.0264	3.523***+	0.0657***+	0.209	0.143
	(0.111)	(0.708)	(0.671)	(0.0193)	(1.346)	(0.0247)	(0.217)	(0.121)
Observations	20,048	19,956	19,947	20,116	4,216	16,622	14,610	14,859

Independent variable: difference between shock variable in wave 2 (pre-treatment) and wave 5 (post-treatment). Each dependent variable: difference between its value in wave 2 (pre-treatment) and wave 5 (post-treatment). Kernel matching using 0.0009375 bandwidth. Bootstrap standard errors in parentheses, 100 reps. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, + $p < 0.05$ using a Hochberg correction with FDR of 0.1.

for the two number of days walked variables, though not significant, may hint at possible physical limitations from experiencing a health shock. The probability that an individual quits smoking (Smoke) has a point estimate that, though not statistically significant, suggests a decrease in that probability. The number of cigarettes smoked in a day (Nr Cigs) decreases significantly by 3.52 cigarettes. The probability to quit drinking alcohol (Drink) is statistically significantly decreased by 6.57 percentage points.[39] The reduction in drinks consumed on the heaviest drinking day of the week (Heavy) and the number of days in a week an individual abstains from drinking alcohol (Days), though neither significant, suggests, if anything, a possible decrease in the frequency and intensity of alcohol consumed.

Heterogeneous Effects by Sex There are heterogeneous effects of the decomposed main analysis when split by sex, as shown in Tables 17 and 18. The statistically significant differences between men and women are the probability to quit smoking, the probability to quit drinking alcohol and the number of days individuals abstain from drinking. However, in the latter case, for neither sex is the effect significantly different from zero. Only women increase their probability to quit smoking and their probability to quit drinking alcohol. For both the increase in number of fruits and vegetables consumed and the reduction in the number of cigarettes smoked, both women and men change their consumption, though the difference between them is not statistically significant.[40]

E.2 Secondary Analysis Results

Strong Shock Only Table 19 shows the impact of experiencing a strong shock on the eight index components without multiple hypothesis test correction. Prior to correcting for multiple hypothesis testing the following effects are found statistically significant at the 10% level: increase in fruit and vegetable consumption (0.26 servings) and decrease in

[39] The number of observations for the alcohol consumption components are much lower than for the other components because questions regarding alcohol consumption habits was asked in a self-completion part of the survey. There are quite some individuals who did not do the self-completion survey component.

[40] In the case of number of cigarettes, the larger magnitude for the decrease by men may in part be mechanical, as men already smoke more cigarettes on average, prior to the shock, thus have a greater scope for reduction than women.

Table 17: ATT of Shock on First-Difference Index Components, by sex – first 4 outcomes

	Fruit/Veg Female	Fruit/Veg Male	Walk 10 Female	Walk 10 Male	Walk 30 Female	Walk 30 Male	Smoke Female	Smoke Male
Shock	0.264*	0.245*	-0.856	-0.976	-1.474	-0.686	0.0542*	-0.00174
	(0.145)	(0.145)	(0.987)	(1.072)	(1.019)	(0.900)	(0.0316)	(0.0253)
Observations	11,880	8,168	11,806	8,150	11,801	8,146	11,912	8,204

Bootstrap standard errors in parentheses, 100 reps. Kernel matching (0.0009375 bandwidth)
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 18: ATT of Shock on First-Difference Index Components, by sex – last 4 outcomes

	Nr Cigs Female	Nr Cigs Male	Drink Female	Drink Male	Heavy Female	Heavy Male	Days Female	Days Male
Shock	3.030*	4.233**	0.131***	0.0135	0.234	0.252	0.0339	0.214
	(1.687)	(1.950)	(0.0430)	(0.0268)	(0.285)	(0.330)	(0.185)	(0.184)
Observations	2,391	1,825	9,847	6,775	8,460	6,150	8,655	6,204

Bootstrap standard errors in parentheses, 100 reps. Kernel matching (0.0009375 bandwidth)
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

the number of cigarettes smoked (4.03 cigarettes). Once corrected for multiple hypothesis testing there are no significant results remaining. Nonetheless, except for the walking 30 minutes per day variable, all the point estimates have the expected non-negative sign. Recall, that when combined into an index the effect of the strong shock is statistically significant. Given that there are fewer treated individuals in this strong shock case, it is likely that decomposed there is not enough power to statistically detect the individual effects of the components. Due to this lack of power concerns, and with the aim to

Table 19: ATT of Strong Shock on First-Difference Lifestyle Index Components

	Fruit/Veg	Walk 10	Walk 30	Smoke	Nr Cigs	Drink	Heavy	Days
Strong Shock	0.258*	0.00513	-0.823	0.0528	4.027*	0.0521	0.236	0.254
	(0.142)	(1.042)	(0.900)	(0.0347)	(2.285)	(0.0340)	(0.267)	(0.183)
Observations	15,492	15,464	15,455	15,544	3,394	12,896	11,384	11,567

Bootstrap standard errors in parentheses, 100 reps. Kernel matching (0.00140625 bandwidth)
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

still provide some insights, a more descriptive discussion of the differences in magnitudes between the weak and strong cases is provided below, after the weak shock only case.

Weak Shock Only Table 20 shows the impact of experiencing a weak shock on the eight index components, both with and without the multiple hypothesis test correction. All the statistically significant point estimates, after correction, have the expected positive sign.[41] Although some individual behaviours suggest a weak shock leads to improvements in some lifestyle behaviours, recall that overall, the weak shock does not have a significant effect on lifestyle change. Looking at the individual behaviours, fruit and vegetable consumption

[41] After applying the Hochberg correction, the Walk 10 variable is no longer statistically significant.

Table 20: ATT of Weak Shock on First-Difference Lifestyle Index Components

	Fruit/Veg	Walk 10	Walk 30	Smoke	Nr Cigs	Drink	Heavy	Days
Weak Shock	-0.0141	-0.959*	-0.736	0.0277^{**+}	2.455^{***+}	0.0392^{**+}	0.139	0.230^{***+}
	(0.0703)	(0.516)	(0.477)	(0.0139)	(0.862)	(0.0161)	(0.141)	(0.0825)
Observations	15,922	15,893	15,885	15,975	3,500	13,270	11,711	11,901

Bootstrap standard errors in parentheses, 100 reps. Kernel matching (0.00140625 bandwidth)
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$, $^{+}$ $p < 0.05$ using a Hochberg correction with FDR of 0.1

and the number of days in a month walking at least 10 or 30 minutes are not statistically significant. The following three effects are statistically significant. The probability to quit smoking increases by 2.77 percentage points. The number of cigarettes smoked per day decreases by 2.46 cigarettes. The probability to quit drinking alcohol decreases by 3.92 percentage points. However, the decrease in the number of drinks consumed on the heaviest drinking day of the week is not significant. Finally, the number of days in a week from which alcohol is abstained increases statistically significantly by 0.23 days.

Comparing Strong and Weak Shocks This section compares the strong and weak shock cases, shown in Tables 19 and 20. As mentioned, the concern is given that there are only a select number of individuals who receive a shock, especially a strong shock, some effects may not be detected for reasons of low power. For the strong shock there are at most 141 treated individuals, 572 for the weak shock.[42]

Fruit and vegetable consumption adjusts only upon experiencing a strong shock (difference between strong and weak shock point estimates is significant at the 5% level). There is no significant difference between the shock strength cases for either the walking at least 10 or 30 minutes each day components. The probability to quit smoking is statistically significantly higher with the strong shock, the effect being nearly twice as large. The number of cigarettes smoked per day decreases for both shock strengths, however the difference is not significant. For the probability to quit drinking there is an increase in the probability from the weak shock, and a statistically significant 30% larger increase in the probability from the strong shock. Neither shock strength seems to significantly effect the heaviness of drinking alcohol. Finally, for the number of days abstaining from drinking alcohol per week there is an increase due to a weak shock, there is a similar size point estimate for the strong shock and the difference between the shock strength effects is not statistically significant.

[42]The text says 'at most' because for each outcome variable the number of individuals who received a shock varies. The value reported here is for the smoking status variable, which is the behaviour variable that has the highest number of individuals who receive a shock.