A Work Project, presented as part of the requirements for the Award of a Master's degree in Economics from the Nova School of Business and Economics.

Nurturing Lifestyles: The Determinants of Health Prevention Behaviors in Portugal

Andreia Raquel Ferreira Novais
42804

Work project carried out under the supervision of:

Professor Pedro Pita Barros

15/01/2024

Abstract

The Portuguese National Health Surveys were used to answer the question "What are the determinants of health prevention behaviors?". The health behaviors considered were exercise, healthy diet, no smoking and prudent alcohol consumption. The variables of interest included demographic information, socio-economic factors and healthcare related variables. A maximum likelihood estimation was performed using a multivariate probit model. The results point to the importance of secondary education, while also highlighting the importance of income for a healthy diet. Differences were found between the NUTS II regions, which, alongside the influence of General Practitioners, should be considered when developing healthcare policies.

Keywords: Health Economics; Health Behaviors; National Health Survey; Multivariate Probit; Instrumental Variable

This work used infrastructure and resources funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209).

Introduction

One of the main causes of death in developed societies are chronic non-communicable diseases. These are also responsible for many cases of disabilities and, consequently, for the loss in life quality. At the origin of this type of diseases one can find risky behaviors such as tobacco consumption, a poor diet, excessive alcohol consumption and the lack of physical activity. In Portugal, around 30% of all the deaths registered in 2019 were associated with diseases influenced by risky behaviors (OECD/European Observatory on Health Systems and Policies 2021). Changing lifestyles can thus lead to a reduction in chronic diseases and in the individual and social economic costs that these entail.

WHO has defined lifestyle as a way of living which involves the interplay between living conditions and the individuals' behaviors, marked by sociocultural determinants and personality traits (Contoyannis and Jones 2004). If these lifestyles can promote or decrease the health of the individual, they can be considered health behaviors.

Hence, to prevent diseases associated with risky behaviors, one needs informed policies and actions targeting lifestyles, which can only be achieved by identifying the factors influencing health behaviors. This study aims precisely at answering this question: "What are the main determinants of health prevention behaviors?". To achieve this, the author looks at socio-economic factors as well as at variables related to healthcare in the Portuguese context, by using the National Health Surveys carried out in 2014 and 2019.

Balia and Jones (2008) and Contoyannis and Jones (2004) have already proven that there is an association between social class and health behaviors, denoting that healthy behaviors are clustered in specific social groups, often with higher education. On a different note, Páscoa et al. (2021) looked at the importance of family doctors in influencing health behaviors, concluding that patients recognize the importance of the family doctor's advice on a number of health behaviors, but that the doctor most often advises on diet.

Following the same empirical strategy used in Balia and Jones (2008) and Contoyannis and Jones (2004), the author has used a multivariate probit model to identify the main determinants of the following health behaviors: exercise, healthy diet, non-smoking, and prudent alcohol consumption. This study fills a gap in the existing literature by consistently estimating an association between socio-economic, demographic, and healthcare related variables with these health preventive behaviors. Besides this, to the best of the author's knowledge, this is the first analysis of the sort done in the context of Portugal, in a repeated cross-sectional setting.

The main findings point to a strong positive association between education and behaviors like exercise and prudent alcohol consumption. Moreover, they indicate that being out of the labor force is associated positively with healthy behaviors, while being employed and income present dichotomic results. The results also highlight differences among the NUTS II regions and that the General Practitioners' influence is more striking on smoking habits.

The report is divided as follows: section 1 reviews the existing literature; section 2 describes the data and the sample used in the analysis; section 3 delves into the methodology; section 4 describes the main results of the analysis and the robustness checks and discusses them; section 5 presents the limitations of the analysis performed, while section 6 concludes.

1. Literature Review

The healthy lifestyles' role has been widely reported in several epidemiological studies, systematic reviews, and meta-analyses. Balia and Jones (2008) delve into the determinants of premature mortality risk in Great Britain and into how lifestyle choices explain the socio-economic inequality in health. To perform this analysis, the authors used a recursive model of mortality, morbidity and lifestyles, specified as a multivariate probit model and estimated by maximum simulated likelihood. It is assumed that lifestyle is a set of behaviors that influence health and whose choice depends on the individuals. The results of this paper highlight how

inequalities in health can be partially explained by differences in lifestyle and living conditions. It is important to mention that lifestyle depends on economic circumstances and, thus, is different among groups. More educated individuals have more healthy behaviors, while people who did not attain any education level have more unhealthy behaviors. In fact, Balia and Jones state that health behaviors tend to cluster in certain categories of individuals and highlight the importance of accounting for the relationship between lifestyle and social class and education.

These results are aligned with those found by Contoyannis and Jones (2004). In this paper, a recursive specification, which included a health production function and reduced form equations for lifestyle, was developed by also using maximum simulated likelihood for a multivariate probit model. Moreover, the authors estimated a static model to analyze the interactions between health behaviors and the self-assessed health status, by taking into account observable factors, like socio-economic status, and unobservable heterogeneity. From simple descriptive statistics and correlations, Contoyannis and Jones (2004) concluded that lifestyle choices can be related to some observable characteristics. For instance, when analyzing the move from completely unhealthy to completely healthy lifestyles, it was evident that the number of individuals in higher social groups increased, while the opposite happened with the lower social class individuals. In a similar way, the proportion of more educated individuals increased with the move from completely unhealthy lifestyles to completely healthy ones. The contrary happened for less educated individuals. Regarding employment status variables, some differences in the probability of having more or less healthy lifestyles were also found. As expected, unemployed individuals were more likely to have an unhealthy lifestyle. Contoyannis and Jones corroborated the conclusions from these simple correlations in their empirical analysis, after controlling for other variables. When looking at the relationship between social class and some health behaviors, the authors found that individuals who belong to lower social classes had a lower probability of eating breakfast. As for the likelihood of being a smoker, the

results pointed to a strong social class gradient. In what concerns other behaviors, such as prudent alcohol consumption and exercise, the same trend was observed, with the higher social classes being more likely to make healthier choices. Regarding education, more educated individuals had a lower probability of being smokers and those with less educational qualifications were less likely to consume alcohol prudently and to exercise.

Both Balia and Jones (2008) and Contoyannis and Jones (2004) point to strong results regarding the relationship between social class, and the socio-economic characteristics that compose it, and health behaviors. Pill et al. (1995) looked precisely at this relationship between social class and preventive health behaviors by using a British national representative sample. They also studied what were the socio-economic factors underlying the observed relationship. Pill and co-authors highlight the fact that social class is a summary descriptive measure which is indicative of differences in income, education, access to resources, attitudes and other factors influencing behavior. The results clearly demonstrated a connection between social class and a health practices index for women. The factors noted by the authors which explain the relationship between social class and behavior include education, tenure, level of crowding, and the measure of recognition that lifestyle influenced health status.

On a different note, Páscoa et al. (2021) look at patients' perspectives regarding lifestyle behaviors and health in a family medicine context in Portugal. The authors perform descriptive statistics and non-parametric tests to study differences between personal beliefs and self-assessed behavior, and between the importance attributed to the family doctor to address health behaviors and the actual approach adopted by the doctor, and its association with bio-demographic variables. The importance of this paper lies in the goal to better understand patients' beliefs and perspectives about lifestyle interventions, such as diet, physical activity, alcohol consumption, tobacco use, illicit drugs, sleep habits, screen activities, stress and sedentarism. According to the survey performed in the context of this study, participants

highlighted the importance of the family doctor's advice on diet, sleep habits, stress, and being sedentary. The respondents reported that the doctor most often asks and advises on diet.

The literature reviewed above sheds some light on some of the determinants of health prevention behaviors. However, it does not consistently estimate a causal relationship between the socio-economic, demographic, and healthcare related variables and health preventive behaviors (such as healthy diet, physical exercise, prudent alcohol consumption and non-smoking). To the best of the author's knowledge, the present study fills a gap in the existing literature on the topic, by using empirical methods to estimate this relationship in the context of Portugal. Given the nature of the data used, repeated cross-sections, only associations were established and not a causal effect.

2. Data and Descriptive Statistics

2.1 Data

To build the dataset for this analysis, the National Health Surveys (NHS) that took place in Portugal in 2014 and 2019, carried out by INE, were used.[1] These surveys targeted all the individuals who were 15 or more years old, who were living in national territory at the time of reference. A stratified sampling design was used so that the sample was representative of the general Portuguese population who were eligible to take the survey. The National Health Survey provides information regarding the following areas: health and disease status; healthcare, namely preventive behaviors and treatment; and health determinants, related to lifestyle and habits which can be classified as disease preventive. The first step when building the dataset was to transform the variables needed in binary variables, given that the data provided was mostly categorical, and to create new variables with the collected data. The data originated by the 2019 NHS was appended to the data from the 2014 survey. A pooled cross-

[1] INE: Instituto Nacional de Estatística ("National Statistics Institute").

section dataset was created from this, given that in each year the sample was randomly selected, hence these surveys are not following the same individuals over the years (panel data setting).[2]

The outcome variables considered refer to health behaviors that are rooted in the lifestyle of the individual. A binary variable to evaluate the physical activity of the individual was created: *exercise*. This variable is equal to one if the individual exerts any physical exercise at least once a week, and zero if this is not the case. WHO has developed guidelines for the different age groups and specific population groups on what is the amount of physical activity needed for good health. However, given the data available and the complications that would be created in the empirical strategy, only one general variable was created for physical exercise, as "some physical activity is better than none" (WHO 2022).

Moreover, a binary variable was created to evaluate whether the individual's diet can be considered healthy and balanced: *healthydiet*. WHO also provides guidelines for the consumption of food. These include basic principles such as consuming fruit and vegetables every day, and consuming legumes, nuts and whole grains. Conversely, the consumption of foods and drinks containing high amounts of sugar should be avoided (WHO 2020). Given the data available, the author has considered a healthy diet when the individual eats three meals per day, and if the individual consumed certain foods in the day before the interview, such as: milk, yoghurt or cheese; bread; meat or fish; potatoes, rice or pasta; and legumes. A healthy diet also considered the non-consumption of cakes, chocolates, desserts and soda on the day before the interview. Notably, the individual should consume fruit and vegetables every day.

Tobacco is one of the leading causes of death and illness (WHO 2023). Hence, to account for the preventive behavior related to tobacco consumption, the author inverted the smoking

[2] The variables definition and sample statistics can be found in Table A.1 in the Appendix.

variable, originating *nosmoke*, that is equal to one if the individual does not smoke daily or occasionally and equal to zero when the individual smokes.

Finally, an outcome variable for prudent alcohol consumption was also created: *safealcohol*. According to WHO, there is not an amount of alcohol that can be considered safe for human health. Therefore, the less alcohol an individual drinks, the safer it is health wise (WHO 2023). Several factors interplay when evaluating the impact of alcohol consumption on health. Frequency, the quantity and the circumstances are aspects that should be taken into account (WHO 2022). Similar to Contoyannis and Jones (2004), abstinence was not chosen as the preventive health behavior. The binary variable *safealcohol* equals one if the individual does not consume alcohol on a weekly basis, that is the frequency of alcohol consumption in the twelve months prior to the survey is inferior to once a week and equals zero when the individual reported to have consumed alcohol at least once a week.[3]

The variables whose impact the author is interested in are the factors which are likely to influence health behaviors, and which can lead to policy implications. With this intent, binary variables were created for each of the NUTS II regions, to account both for the differences in the functioning of the regional health systems, as well as the differences in life conditions and habits between the regions. In fact, in Portugal, the national health system is composed by health regional administrations, one for each of the NUTS II regions in the continent. These have the goal of securing regional resource planning, by promoting territorial cohesion in healthcare and by developing public health activities (Lei Orgânica do Ministério da Saúde aprovada pelo Decreto-Lei n.º 124/2011, de 29 de dezembro).

Other possible determinants of health behaviors include variables more related to the socio-economic status of the individual. Binary variables were created for each of the education

[3] This measure does not distinguish between a week of moderate drinking, and one composed of periodic abstinence and heavier drinking, which can have different health effects (Contoyannis and Jones 2004).

levels, which correspond to the maximum education level attained by the individual. Similarly, the author has used binary variables to characterize the employment status of the individual. Regarding income, a continuous variable was used with only five categories, each corresponding to an income quintile. Binary variables were also created for each income quintile when proceeding to a more detailed analysis of the impact of income on the health preventive behaviors.

Finally, the other subset of determinants considered are related to healthcare. Firstly, a binary variable was used to evaluate whether the individual had access to a healthcare subsystem or not. A binary variable for the appointments with a General Practitioner was also considered, which equals one if the individual had an appointment with a General Practitioner in the twelve months prior to the survey and equals zero if the last appointment was held more than one year before the survey. This variable was included to shed some light on the impact of primary healthcare in incentivizing health behaviors. One of the goals of primary healthcare is to promote health and disease prevention (WHO 2023). General Practitioners and, particularly, family doctors in Portugal, play an important role in primary healthcare, since they are in a position which allows them to deal with lifestyle preventive behaviors (Páscoa et al. 2021).

The control variables used in the model can be grouped into categories: demographic information about the individual; main job; the health status of the individual; the self-reported satisfaction with life conditions; and year dummies. However, it is important to highlight how some of the variables were built. Binary variables for the skill level of the main job reported by the individual were created, according to the ISCO-08 structure developed by the International Labor Organization. In the first skill level, only non-qualified work was considered. As for the second skill level, the following categories were included: administrative personnel; personal, protection and security services and sellers; qualified workers in agriculture, fishery and forestry; qualified workers in industry and construction; and, finally, plant and machine

operators. Regarding the third skill level only intermediate level jobs were considered. Finally, in the fourth skill level one can find the following categories: army; representatives of the legislative branch and executive bodies, directors and executive managers; and specialists in intellectual and scientific activities. The binary variable regarding chronic diseases (*chronicdis*) is equal to one if the individual reported suffering or having suffered in the twelve months prior to the survey from at least one of the following health problems: hypertension; arthrosis; back or neck chronic pain; diabetes; allergies; and depression. Not all the possible chronic diseases reported by the surveyed individuals were considered, but only the six most common ones, according to the National Health Institute Doutor Ricardo Jorge. To evaluate the impact of mental health, a group of binary variables was used, according to the Patient Health Questionnaire-9 (PHQ9).[4] The author equalized the binary variables to one if the individual reported having a behavior almost every day or more than half of the days in the two weeks prior to the survey, indicating a moderate, moderately severe or severe depression. Only two behaviors which are part of the PHQ9 evaluation and were reported in the NHS were not considered in the analysis: difficulties in concentrating and feeling tired. These were excluded given the high probability of suffering from reverse causality with health behaviors and their lack of relevance for the analysis, given the other control variables.

All observations which included non-responses to any of the survey questions used in the study were dropped. This meant that in total 2036 observations were dropped (1302 out of 14617 from the 2019 survey and 734 out of 18204 from the 2014 survey), not evidencing self-selection bias. Hence, both the descriptive statistics and the empirical analysis are conditional upon this restriction.

[4] The PHQ9 is an instrument used in healthcare to screen, diagnose, monitor and measure the severity of depression.

2.2 Descriptive Statistics

Table 1 ***Sample Characterization (n=30785)***

	Absolute Frequency	Relative Frequency (%)
Region (NUTS II)		
Norte	4539	14.74
Algarve	4037	13.11
Centro	6418	20.85
Lisboa	3069	9.97
Alentejo	4970	16.14
Região Autónoma dos Açores	3783	12.29
Região Autónoma da Madeira	3969	12.89
Gender		
Male	13354	43.38
Female	17431	56.62
Education Level		
No Education Level	3607	11.72
Basic Education	17557	57.03
Secondary Education	4782	15.53
Post-secondary/Higher Education	4839	15.72
Employment Status		
Employed	13279	43.13
Unemployed	2936	9.54
Out of the Labor Force	14570	47.33
Skill Level of Main Job (ISCO-08)		
Skill Level 1	1762	5.72
Skill Level 2	6791	22.06
Skill Level 3	1410	4.58
Skill Level 4	3316	10.77
Healthcare Subsystem	5652	18.36
Health Insurance	5621	18.26
Chronic Disease	19866	64.53
Year		
2014	17470	56.75
2019	13051	42.39
2020	264	0.86

The full sample is composed of 30785 observations. Of these, 17470 were registered in 2014, resulting from the survey developed in that same year. From the survey taking place in 2019, 13051 observations were collected in that year, while 264 were collected in the following year, 2020.

In Table 1 there is a simple characterization of the sample. Overall, the sample was evenly distributed among the NUTS II regions. The regions which registered more observations were "Centro" (20,85% of the sample), "Alentejo" (16,14% of the sample), and "Norte" (14,74% of the sample). Notably, 57,03% of the individuals reported only having the basic education level.

As for employment status, only 9,54% of the individuals reported being unemployed, with 43,13% of the sample being employed and 47,33% being out of the labor force. Furthermore, 18,36% of the individuals reported having access to a healthcare subsystem, while 18,26% reported having private health insurance.

The health behaviors of the surveyed individuals are characterized in Table 2. It is important to note that 70,18% of the respondents reported not exercising one single day in a week. Besides this, only 4,60% of the individuals have a diet which can be classified as a healthy one according to the criteria used. Regarding the proportion of respondents who reported smoking daily or occasionally, this amounts to 18,13% of the sampled individuals. Finally, when asked about the frequency of alcohol consumption in the 12 months prior to the survey, 39,23% of the respondents reported having consumed alcohol at least once a week. When restricting the sample to women aged 20-40, this value decreases to 16,83%.

Table 2 ***Health Prevention Behaviors Characterization (full sample)***

	Absolute Frequency	**Relative Frequency (%)**
Frequency of physical exercise in a week (number of days)		
Zero	21604	70.18
One	1203	3.91
Two	2447	7.95
Three	1810	5.88
Four	819	2.66
Five	1071	3.48
Six	365	1.19
Seven	1466	4.76
Healthy Diet	1417	4.60
Smoking	5581	18.13
Alcohol consumption (weekly basis)	12078	39.23

When dividing the sample between individuals who reported having health insurance and those who reported not having, some striking differences can be observed. Regarding the education level, one concludes that the individuals who reported having insurance have a higher education level than those with no health insurance. Notably, 69,51% of the surveyed individuals who reported having health insurance are employed. Finally, and illustrating the income effect when getting insurance, the fifth income quintile is the one with the highest

proportion of individuals with health insurance.[5]

Regarding the health behaviors adopted by individuals with health insurance and individuals without insurance, one can observe that 56,63% of individuals with insurance reported not exercising at least once a week, while this percentage was higher for the individuals with no health insurance (73,20%). In both sub-samples the proportion of respondents with a healthy diet is similar. However, the proportion of individuals with harmful behaviors like smoking and alcohol consumption is slightly higher in the sub-sample of insured respondents.[6]

According to simple correlations, one can already observe some associations between the explanatory variables and the healthy behaviors. For instance, being out of the labor force seems to have a positive relationship with being a non-smoker and a prudent alcohol consumer. Regarding education, the basic education level is only positively correlated with having a healthy diet, while the secondary education level displays a strong positive correlation with exercising and a positive correlation with consuming alcohol prudently. Having a post-secondary/higher education level is positively correlated with exercising as well. Moreover, having an appointment with a General Practitioner only has a negative correlation with exercising, while having a strong positive correlation with being a non-smoker. Finally, when looking at the partial correlations between the NUTS II regions and the health behaviors, one observes that the region "Norte" is the only region with a positive correlation with healthy diet, while this region and "Centro" are the only ones with a positive correlation with non-smoking. Nevertheless, the region "Norte" is the only region with a negative correlation with prudent alcohol consumption.

On a different note, when analyzing the healthy behaviors according to the education levels, one can see some non-linearity. While the proportion of individuals who exercise at least

[5] This information is detailed in Table A.6 in the Appendix.
[6] This information is detailed in Table A.8 of the Appendix.

once a week increases as the education level increases, this does not happen for the other three behaviors. In fact, the proportion of individuals who have a healthy diet and don't smoke is fairly similar between the basic and the post-secondary/higher education levels, while for the secondary education level the proportion is lower. As for the prudent consumption of alcohol, the opposite happens, that is the proportion of individuals with secondary education is higher.[7]

3. Methodology

To perform the analysis the following equation was used:

$$y_i = \beta_0 + \beta_1 Determinants_i + \beta_2 X_i + \beta_3 T_i + \varepsilon_i \quad (1)$$

Subscript *i* refers to all the individuals in the sample across all the variables considered, including the year dummies (stacked observations). The vector of outcome variables is y_i and includes the binary variables corresponding to the health behaviors: exercise, healthy diet, no smoking and prudent alcohol consumption. The vector $Determinants_i$ corresponds to the independent variables whose impact on health behaviors the author is interested in: the NUTS II regions, the individual's education level, the individual's employment status, household income, having access to a healthcare subsystem and having had an appointment with a General Practitioner in the twelve months prior to the survey. β_1 is the coefficient of interest. X_i represents the vector of control variables used in the regressions for *healthydiet*, *nosmoke*, and *safealcohol*, and can be divided into the following categories: individual specific and demographic characteristics; information regarding the individual's main job; the respondent's physical and mental health status; and the self-reported satisfaction with life conditions. The vector X_i was also used in the regression *exercise* and includes all the covariates mentioned for the other health behaviors, but also a dummy variable for whether the individual has any limitation with impact on daily activity, the level of effort exerted on the individual's main activity, and whether the individual smoked in the past or not. Finally, T_i corresponds to a vector

[7] This information is detailed in Table A.9 of the Appendix.

of year dummies to control for the year in which the data was collected.

The empirical strategy applied was a Multivariate Probit Model (MVP), with a recursive system of equations for the four health behaviors analyzed. Since the behaviors are being registered at the same point in time, there is a high probability that unobservable individual characteristics are influencing all the individual's health behaviors. The main advantage of this method, when compared to a standard univariate probit model, is that the random components of each health behavior equation are allowed to be correlated between themselves, rendering the estimates more efficient. However, there is the need to assume that the error terms have a multivariate normal distribution.

Nevertheless, some econometric issues have arisen. Given the model used and the nature of the data collected, there is the need to consider possible endogeneity problems. Firstly, there are individual specific characteristics which cannot be observed by the author, such as risk attitude. This can lead to omitted variable bias, since it can impact the health behaviors considered, as well as the decision to have private health insurance. Therefore, whether the individual has health insurance or not was not controlled for in the main specification model. To account for this factor, the main sample was divided into two sub-samples according to the health insurance criterium. The same model was then applied. Moreover, reverse causality can also arise. One of the variables used to control for the individual's mental health status was *feeldep,* to evaluate if the respondent felt depressed in the two weeks prior to the survey. However, this feeling can be instigated or mitigated by health behaviors like exercise and diet. Therefore, to deal with the potential endogeneity, an instrumental variable was used: a binary variable (*lifeunsat*) equal to one if the individual was not satisfied with life, and equal to zero otherwise. The model used for this specification was a univariate probit with an IV. Two assumptions had to be made: the instrument exogeneity and the instrument relevance. For the first assumption, it is considered that the instrumental variable is uncorrelated with the error

term for all the four health behaviors, since it seems unlikely that the self-reported unsatisfaction with life can impact the behaviors through any other variables. To verify the second assumption, instrument relevance, the correlation between *feeldep* and *lifeunsat* was computed, rendering a value of 0.3059. Thus, the instrument relevance assumption is also valid, even if the instrument is classified as a weak instrument.[8]

4. Results

4.1 Multivariate Probit Results for the Main Model

Table 3 ***Multivariate Probit Selected Average Marginal Effects for the main model (full sample)***[9]

	(1) Exercise (+)	(2) Healthy Diet (+)	(3) No Smoke (+)	(4) Safe Alcohol (+)
north	-0.116***	0.0986*	0.241***	-0.0847**
	(-3.65)	(2.12)	(6.64)	(-2.71)
algarve	-0.0465	-0.357***	0.0814*	0.311***
	(-1.45)	(-6.81)	(2.29)	(9.83)
centro	-0.0680*	0.0112	0.281***	0.153***
	(-2.29)	(0.25)	(8.21)	(5.24)
alentejo	-0.233***	-0.258***	0.138***	0.452***
	(-7.34)	(-5.25)	(3.93)	(14.53)
raaçores	-0.333***	-0.455***	0.0428	0.476***
	(-9.76)	(-7.57)	(1.17)	(14.22)
ramadeira	0.0799*	-0.168**	0.224***	0.509***
	(2.46)	(-3.26)	(6.03)	(15.40)
basiceduc	0.162***	0.0295	-0.0418	0.0110
	(5.09)	(0.70)	(-1.01)	(0.39)
seceduc	0.305***	-0.182**	-0.0418	0.127***
	(7.95)	(-3.15)	(-0.90)	(3.53)
postsecsup	0.423***	-0.0501	0.00698	0.0871*
	(10.04)	(-0.81)	(0.14)	(2.16)
employed	-0.349***	-0.236***	0.260***	0.306***
	(-7.84)	(-3.38)	(6.15)	(7.35)
noworkforce	0.130***	-0.233***	0.582***	0.444***
	(4.17)	(-4.75)	(17.83)	(14.30)
income	0.0784***	0.0331**	-0.00330	-0.0460***
	(10.85)	(2.88)	(-0.40)	(-6.41)
healthsub	0.144***	0.0851*	-0.0621*	-0.0572**
	(6.69)	(2.43)	(-2.46)	(-2.61)
genpract	0.0929***	0.0203	0.135***	0.0780***
	(4.88)	(0.64)	(6.61)	(4.15)
Observations	30785	30785	30785	30785
Wald Chi2	43208.7	43208.7	43208.7	43208.7
Prob > chi2	0	0	0	0
Log lik.	-51611.9	-51611.9	-51611.9	-51611.9

t statistics in parentheses
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In Table 3, the main results from the multivariate probit for the main model are presented.

[8] The limitations of using such an instrument will be explored in Section 5.
[9] Full table in the Appendix (Table B.1).

According to the average marginal effects for the regression *exercise*, one can see that the probability of exercising at least once a week in the NUTS II regions "Norte", "Alentejo" and "Região Autónoma dos Açores" is significantly lower than in the region "Lisboa". Furthermore, having the secondary education level and post-secondary or higher education levels increases the probability of exercising at least once a week when compared to the individuals with no education level. It is important to highlight that as the education level of the individual increases, the probability of exercising at least once a week also increases. Nevertheless, an employed individual is significantly less likely to exercise at least once a week than an unemployed individual. Contrarily, being out of the labor force increases the probability when compared with the unemployment case. An individual with access to a healthcare subsystem is more likely to exercise at least once a week than one who doesn't.

Regarding the average marginal effects for the regression *healthydiet*, individuals living in the NUTS II regions "Algarve", "Alentejo", "Região Autónoma dos Açores" and "Região Autónoma da Madeira" display a significantly lower probability of having a healthy and balanced diet than those living in "Lisboa". It is important to mention that having the secondary education level, when compared to the case in which the individual did not attain any education level, is associated with a lower probability of having a healthy diet. Moreover, being employed or not being part of the labor force means that the likelihood of having a healthy diet, when compared to unemployed individuals, is lower.

Furthermore, individuals living in "Norte", "Centro", "Alentejo" and "Região Autónoma da Madeira" are more likely to be non-smokers than those living in "Lisboa". It is particularly important to highlight that an individual out of the labor force has a significantly higher probability of being a non-smoker than an unemployed individual. An employed individual is also more likely to be a non-smoker, but the result in this case is not as striking. On a different note, respondents who had an appointment with a General Practitioner in the twelve months

prior to the survey are also more likely to be non-smokers.

Finally, the region "Norte" is the only NUTS II region displaying a negative average marginal effect when it comes to safe alcohol consumption. This means that, with the exception of "Norte", individuals from all the NUTS II regions are more likely to have consumed alcohol prudently in the 12 months prior to the survey, when compared to the respondents from "Lisboa". Besides this, having completed secondary education increases the probability of consuming alcohol prudently, when compared to a case in which the individual did not complete any education level. As for being employed or out of the labor force, the probability of being a prudent alcohol consumer is higher than in the case of an unemployed individual. However, the likelihood is particularly higher for the individuals out of the labor force.

To better understand what the relationship of income with health behaviors is, the author computed the same regression as in the main model but specified the income quintiles separately. These results can be seen in Table 4. In general, the association of income with exercise is positive, and the results for the fourth and fifth income quintiles are particularly significant, when compared to individuals whose average income falls into the first income quintile. The same general positive impact of income on the probability of having a healthy diet can be seen (in Table 3). However, when analyzing the income quintiles separately, one concludes that individuals in the second income quintile display a lower probability of having a healthy diet than those in the first quintile. As for the results for the fifth income quintile, one can say that these are also positive and statistically significant in this case. Regarding non-smoking, even though the coefficient in the main regression (Table 3) is negative, when looking at the income quintiles, the only category displaying a negative value is the fifth income quintile. Finally, income displays a negative association with prudent alcohol consumption (*safealcohol*). The only income quintile presenting a positive coefficient is the second one, while the fourth and fifth income quintiles present significant negative average marginal effects.

Table 4 ***Multivariate Probit Average Marginal Effects for the income quintiles***

	(1) Exercise (+)	(2) Healthy Diet (+)	(3) No Smoke (+)	(4) Safe Alcohol (+)
income2	0.00914	-0.0486	0.0577*	0.00482
	(0.35)	(-1.22)	(1.97)	(0.19)
income3	0.0625*	0.0274	0.00958	-0.0787**
	(2.35)	(0.66)	(0.32)	(-3.05)
income4	0.171***	0.0753+	0.0355	-0.145***
	(6.13)	(1.70)	(1.14)	(-5.26)
income5	0.348***	0.218***	-0.0471	-0.235***
	(11.08)	(4.37)	(-1.32)	(-7.46)
Observations	30785	30785	30785	30785
Wald Chi2	43256.9	43256.9	43256.9	43256.9
Prob > chi2	0	0	0	0
Log lik.	-51554.2	-51554.2	-51554.2	-51554.2

t statistics in parentheses
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

From this analysis, one can conclude that, in general, the individuals who display a higher probability of exercising at least once a week possess a secondary or a post-secondary/higher education level. They also have access to a healthcare subsystem, are not part of the labor force and have an average income which places them in the fourth or fifth income quintiles. Individuals who present a higher probability of having a healthy diet are people whose income is included in the fifth quintile. Income seems to play an important role in this behavior, since a healthy diet is dependent on many factors not controlled by the individual such as food prices and food availability. It is important to note that the result for the secondary education level is significant and negative, which denotes the importance of diet education being administered in this education level. The individuals who possess a higher probability of not smoking are either employed or out of the labor force, had an appointment with a General Practitioner in the twelve months prior to the survey and are positioned in the second, third or fourth income quintiles. Individuals with the secondary education level, who are employed or out of the labor force have a higher probability of drinking alcohol prudently. Similar to the behavior observed in *nosmoke*, higher income has a relatively strong negative impact on prudent alcohol consumption.

4.2 Discussion

The differences observed between the regions can be understood in the light of the existing inequalities between them in the country, in terms of life conditions. For instance, rural

areas have significantly different conditions and lifestyles from urban areas. These variations can also impact the functioning of the regional healthcare services. In fact, Local Health Units located near the coast and in urban areas are more efficient (Moreira 2016). [10] This might explain why the region "Lisboa" presents a higher probability to exercise and have a healthy diet. These results point to a need to explore the differences in the functioning of the regional health administrations, to adopt similar effective policies in promoting health behaviors.

The results of this analysis are in line with those obtained by Balia and Jones (2008) and Contoyannis and Jones (2004). Education has an impact on the probability of having healthy habits, and, particularly, attaining the secondary education level has a positive impact on the probability of exercising and consuming alcohol prudently. The author conjectures that the results for this specific education level are related to the habits and beliefs developed and consolidated during the age at which secondary school is attained. Being employed is associated with two outcomes. On the one hand, there is an increase in the probability of not smoking and of consuming alcohol prudently. On the other hand, it displays a lower probability of exercising at least once a week and of having a healthy diet. These last results can be related to the levels of stress that the individuals are experiencing when employed and to the lack of time to exercise or eat three meals per day, for example. Income also displays two opposite effects. Contrarily to being employed, being on a higher income quintile is associated with an increase in the probability of exercising once a week and of having a healthy diet but it is associated with a decrease in the probability of drinking alcohol prudently and non-smoking. These results denote that there is an income effect which increases the probability of having unhealthy behaviors.

Even though these results are in line with many other studies, which demonstrate how belonging to a higher social class facilitates access to knowledge and resources, thus enhancing

[10] Local Health Units ("Unidades Locais de Saúde"-ULS) correspond to the vertical integration of healthcare which was carried out in Portugal to connect healthcare centers, hospitals and other entities responsible for the health of the local and regional population.

health and potentially leading to changes in health behaviors, they also illustrate the nuanced relationship between social class and health. People can "experience multiple social statuses simultaneously, and their effects on health behaviors are not simply additive" (Short and Mollborn 2015). This sheds some light on the duality of effects of some of the categories of social class analyzed, such as employment status and income.

On a different note, the results obtained for the individuals out of the labor force are very interesting, given the magnitude of the coefficient found and its statistical significance. These individuals have a higher probability of exercising at least once a week, and particularly of being non-smokers and prudent alcohol consumers. This can be explained by the sub-groups included in this category of employment status: retired people, students, people who do not work due to incapacity, domestic workers and people doing community service. However, a more detailed analysis would be needed to understand which sub-group is driving these results.

Another variable whose results are worth mentioning is the binary variable for whether the individual had an appointment with a General Practitioner in the twelve months prior to the survey. This has a particularly significant, and positive, impact on the probability of being a non-smoker. Even though it is consensual that General Practitioners can influence behaviors (Páscoa et al. 2021), it might be the case that doctors focus more on some behaviors than others. Nevertheless, one also needs to have in mind that behaviors can only be influenced if the patients are receptive. Hence, it might be the case that Portuguese people tend to take advice on smoking more seriously, since it is probably perceived as the behavior causing more health problems, due to all the anti-smoking campaigns performed through the years. On a final note, it is important to consider the lack of family doctors in the country, with approximately 14,7% of people without one (Mendes 2023). Given family doctors follow a given set of patients through their life, thus establishing a relationship of confidence with them, influencing behaviors becomes more effective. In fact, the scarcity of family doctors generates more

inequality regarding health behaviors. However, one should be careful in the interpretation of these results because these might not reflect a pure association between visiting a General Practitioner and having the health behaviors specified, since one can consider that people who worry more about their health will have appointments more often. Given the many factors interplaying in the analysis of this variable, the author found no need to correct for the possible endogeneity in this case, even if it is important to be aware of the drawback.

To conclude, these results point to the importance of some factors, like education, employment status and access to healthcare services, which can potentially be modified. An education component should be included in policies to stimulate healthy habits, but it should also be considered the role that medical professionals play in influencing these behaviors.

4.3 Robustness Checks

The main differences between the multivariate probit results for the main model and those for the univariate probit regression happen for healthy diet. For the three education levels the effects are positive and statistically significant, whereas before only the one for basic education was positive. This means that having any level of education increases the probability of having a healthy diet relative to having no education level, as initially expected. Furthermore, for the employment status categories, the impact on healthy diet becomes positive, but these results are not statistically significant, and the average marginal effects are very low. Moreover, for the regression for prudent alcohol consumption, all the education levels display a negative effect, which is statistically significant as well. All the other results are more or less similar.[11]

When looking at the results from the multivariate probit model when only the individuals with health insurance were considered, one can conclude that the main differences are observed in the education levels' impact. For the regression on exercise, the effect for the basic education level becomes negative. The same happens in the regression for healthy diet. All education

[11] Table B.3 in the Appendix contains all the results for the univariate probit model.

levels display a higher probability of being a non-smoker when compared with the case of having no education level, while in the main regression this only happens for the post-secondary/higher education levels. Finally, for prudent alcohol consumption, even though there is no change in the sign of the average marginal effects for the education levels, the results are stronger (the coefficients are higher and statistically significant).[12]

A univariate probit model with an IV was also considered.[13] While in the previous regressions the variable *feeldep* was used as a normal control, in this regression an instrumental variable (*lifeunsat*) was used to correct for possible endogeneity. For the regression regarding exercise, the average marginal effects for the regional dummies are the ones that display more differences when comparing to the main model, while all the other coefficients are very similar, both in terms of magnitude and sign. However, for *healthydiet*, major differences were found both in terms of the sign and magnitude of the average marginal effects. The impact of having a secondary education or a post-secondary/higher education level on having a healthy diet is now positive, which also happens for the categories considered for employment status. Both changes are in line with the differences verified for the univariate probit results for the main regression. Regarding the results for prudent alcohol consumption, as happens in the univariate probit model, the average marginal effects for all the education levels become negative.

To conclude, the results for the main model are similar across the different specifications considered, as the main differences worth highlighting are in the effect of the education levels.

5. Limitations

The main limitation of this analysis lies in the assumption that the error terms are normally distributed. If this crucial assumption is not verified, the maximum likelihood problem is not correctly defined, which renders the parameter estimates inconsistent. The assumption of the

[12] Table B.4 of the Appendix displays all the results for the multivariate probit with the sub-sample of individuals with health insurance.
[13] Table B.6 of the Appendix displays all the results for the univariate probit model with an IV.

normality of the error terms is a central hypothesis in all the estimation and in any statistical inference based on the parameter estimates. Moreover, guaranteeing the normality of the error term is fundamental in order to interpret the effects of the changes in the explanatory variables, since these correspond to changes in the probabilities of each of the outcomes (Johnson 1996).

Another major limitation of the analysis is present in the construction of the outcome variable *healthydiet*. Given it considered as criteria to eat three meals per day and the consumption of foods in the day before the survey was answered, it might be the case that it does not perfectly capture if in general the individual has a healthy and balanced diet, thus biasing the results downwards. Moreover, given the fact that most information, particularly the health behaviors, was self-reported, there is the probability that some bias could have been induced by this, since it is dependent on the ability of the individuals to remember facts.

Finally, the last major limitation of the analysis is only present in the univariate probit with an IV model. The instrument used (*lifeunsat*) displays a very low correlation with the instrumented variable (*feeldep*): only 0.3059. This means that, even though the instrument relevance condition is verified (the correlation between the two variables is not zero), the instrument is considered weak. In the case of a weak instrument, using the IV can actually lead to larger asymptotic biases than using the standard univariate probit.

6. Conclusion

In this analysis, the data collected through the National Health Surveys developed in 2014 and 2019 was used to identify the most relevant determinants of health prevention behaviors in Portugal. Hence, associations were established between the NUTS II regions, the education level, the employment status, the access to a healthcare subsystem and the appointments with a General Practitioner, and the following health behaviors: exercising at least once a week, having a healthy and balanced diet, not smoking and consuming alcohol prudently. A maximum likelihood estimation was performed through a multivariate probit model, to allow for the

correlation between the error terms of the health behaviors considered. The main limitation of this model is the normality assumption for the error terms.

The results found point to a strong association between social class and health behaviors. Secondary education has a positive impact on exercise and prudent alcohol consumption. This result denotes that health prevention education can have an important role in this education level. Being employed has a positive impact on non-smoking and on prudent alcohol consumption, but a negative one on exercising and on having a healthy diet. Furthermore, while higher income quintiles are associated with more exercise and a healthier diet, they are also associated with smoking and frequent alcohol consumption. This income effect on exercise, and particularly on diet, should be taken into account, since it highlights that financial constraints are an obstacle to having a healthy and balanced diet, going beyond the individual's choices. Some differences were also observed between the NUTS II regions, which possibly indicate that the health prevention strategies applied by the different regional health administrations should be analyzed in more detail so that they can be homogenized across the regions. The role of General Practitioners was more evident for non-smoking. However, the results for the other health behaviors also point to a positive influence of GPs, even if small, which denotes that the family doctors' role should be considered in policy making.

In further research, a more broaden set of health behaviors should be considered, such as sleep hygiene and stress management. It would also be important to study the underlying mechanisms through which education and the labor market influence health behaviors.

When analyzing the main determinants of health prevention behaviors, education comes as an important factor to have in consideration, as well as employment status. Income also plays a fundamental role in lifestyle decisions, while differences were found among the NUTS II regions. Finally, the role of family doctors should be given special consideration and prioritized when developing policies targeting health preventive behaviors.

References

Balia, Silvia, and Andrew M. Jones. 2008. "Mortality, Lifestyle and Socio-Economic Status." *Journal of Health Economics* 27, no. 1: 1–26. https://doi.org/10.1016/j.jhealeco.2007.03.001.

Contoyannis, Paul, and Andrew M. Jones. 2004. "Socio-Economic Status, Health and Lifestyle." *Journal of Health Economics* 23, no. 5: 965–995. https://doi.org/10.1016/j.jhealeco.2004.02.001.

Green, Edward C., Elaine M. Murphy, and Kristina Gryboski. 2020. "The Health Belief Model." *The Wiley Encyclopedia of Health Psychology*: 211–214. https://doi.org/10.1002/9781119057840.ch68.

Instituto Nacional de Estatística, IP. 2014. *Documento metodológico do Inquérito Nacional de Saúde 2014*. Lisboa: INE.

Instituto Nacional de Estatística, IP. 2019. *Documento metodológico do Inquérito Nacional de Saúde 2019*. Lisboa: INE.

Instituto Nacional de Estatística IP, Instituto Nacional de Saúde Doutor Ricardo Jorge IP. 2016. *Inquérito Nacional de Saúde 2014*. Lisboa: INE/INSA. https://www.ine.pt/ngt_server/attachfileu.jsp?look_parentBoui=263722386&att_display=n&att_download=y.

Instituto Nacional de Estatística IP, Instituto Nacional de Saúde Doutor Ricardo Jorge IP. 2021. *Estatísticas da Saúde 2019*. Lisboa: INE/INSA. https://www.ine.pt/ngt_server/attachfileu.jsp?look_parentBoui=492336067&att_display=n&att_download=y.

International Labour Organization. 2023. "The International Standard Classification of Occupations- ISCO-08." International Labour Organization. Accessed December 12, 2023. https://isco-ilo.netlify.app/en/isco-08/.

Johnson, Paul A. 1996. "A test of the normality assumption in the ordered probit model." *Metron* 54: 213-221. https://mpra.ub.uni-muenchen.de/10080/.

Jones, Andrew M., Nigel Rice, Teresa Bago d'Uva, and Silvia Balia. 2007. *Applied Health Economics*. London and New York: Routledge.

Kroenke, Kurt, Robert L. Spitzer, Janet B.W. Williams, and Bernd Löwe. 2009."An Ultra-Brief Screening Scale for Anxiety and Depression: The PHQ–4." *Psychosomatics* 50, no. 6: 613–621. https://doi.org/10.1016/s0033-3182(09)70864-3.

Mendes, Teresa Alves. 2023. "Retrato Dos Serviços de Saúde Em Portugal." Jornal de Negócios. Accessed December 12, 2023. https://www.jornaldenegocios.pt/negocios-iniciativas/portugal-health-summit/detalhe/retrato-dos-servicos-de-saude-em-portugal.

Ministério da Saúde. 2011. "Lei Orgânica." SNS. Accessed December 12, 2023. https://www.sns.gov.pt/institucional/ministerio-da-saude/lei-organica/.

Moreira, António André de Sousa. 2016. "Análise de Eficiência das Unidades Locais de Saúde: uma aplicação da Data Envelopment Analysis." Master's Thesis, Faculdade de Economia da Universidade do Porto. https://repositorio-aberto.up.pt/bitstream/10216/86534/2/163325.pdf.

National HIV Curriculum. 2023. "Patient Health Questionnaire-9 (PHQ-9)." National HIV Curriculum. Accessed December 12, 2023. https://www.hiv.uw.edu/page/mental-health-screening/phq-9.

OCDE/Observatório Europeu dos Sistemas e Políticas de Saúde. 2021. *Portugal: Perfil de Saúde do País 2021, Estado da Saúde na EU*. Paris: OCDE. Bruxelas: Observatório Europeu dos Sistemas e Políticas de Saúde. https://health.ec.europa.eu/system/files/2021-12/2021_chp_pt_portuguese.pdf.

Páscoa, Rosália, Andreia Teixeira, Micaela Gregório, Rosa Carvalho, and Carlos Martins. 2021. "Patients' Perspectives about Lifestyle Behaviors and Health in the Context of Family

Medicine: A Cross-Sectional Study in Portugal." *International Journal of Environmental Research and Public Health* 18, no. 6: 2981. https://doi.org/10.3390/ijerph18062981.

Pill, R., T. J. Peters, and M. R. Robling. 1995."Social Class and Preventive Health Behaviour: A British Example." *Journal of Epidemiology & Community Health* 49, no. 1: 28–32. https://doi.org/10.1136/jech.49.1.28.

Rafferty, Anthony, Pierre Walthery, and Sarah King-Hele. 2015. *Analysing change over time: repeated cross-sectional and longitudinal data.* UK: UK Data Service, University of Essex and University of Manchester. https://www.ukdataservice.ac.uk/media/455362/changeovertime.pdf.

Rodrigues, Carlos Farinha. 2023. "Portugal Desigual: Um Retrato Das Desigualdades de Rendimentos e Da Pobreza No País." Fundação Francisco Manuel dos Santos. Accessed December 12, 2023. https://www.ffms.pt/pt-pt/estudos/portugal-desigual-um-retrato-das-desigualdades-de-rendimentos-e-da-pobreza-no-pais.

Short, Susan E., and Stefanie Mollborn. 2015. "Social Determinants and Health Behaviors: Conceptual Frames and Empirical Advances." *Current Opinion in Psychology* 5: 78–84. https://doi.org/10.1016/j.copsyc.2015.05.002.

SNS, INSA. 2019. "Infográfico INSA ─ Doença Crónica." SNS. Accessed December 12, 2023. https://www.insa.min-saude.pt/infografico-insa-%E2%94%80-doenca-cronica/.

WHO. 2020. "Healthy Diet." World Health Organization. Accessed December 12, 2023. https://www.who.int/news-room/fact-sheets/detail/healthy-diet.

WHO. 2022. "Alcohol." World Health Organization. Accessed December 12, 2023. https://www.who.int/news-room/fact-sheets/detail/alcohol.

WHO. 2022. "Physical Activity." World Health Organization. Accessed December 12, 2023. https://www.who.int/news-room/fact-sheets/detail/physical-activity.

WHO. 2023. "No Level of Alcohol Consumption Is Safe for Our Health." World Health

Organization. Accessed December 12, 2023. https://www.who.int/europe/news/item/04-01-2023-no-level-of-alcohol-consumption-is-safe-for-our-health.

WHO. 2023. "Primary Health Care." World Health Organization. Accessed December 12, 2023. https://www.who.int/news-room/fact-sheets/detail/primary-health-care.

WHO. 2023. "Tobacco." World Health Organization. Accessed December 12, 2023. https://www.who.int/news-room/fact-sheets/detail/tobacco.

Wooldridge, Jeffrey M. 2013. *Introductory Econometrics: A Modern Approach*. Fifth Edition. USA: South-Western, Cengage Learning.

Appendix

Table A.1 ***Variable Definition and Descriptive Statistics***

Variables	Variable Definition	Mean	S.D.	Min	Max
region	NUTS II regions	17.85	5.33	11	30
north	Dummy variable (=1 if the individual lives in the region "Norte"; =0 otherwise)	0.15	0.35	0	1
algarve	Dummy variable (=1 if the individual lives in the region "Algarve"; =0 otherwise)	0.13	0.34	0	1
centro	Dummy variable (=1 if the individual lives in the region "Centro"; =0 otherwise)	0.21	0.41	0	1
lisbon	Dummy variable (=1 if the individual lives in the region "Lisboa"; =0 otherwise)	0.10	0.30	0	1
alentejo	Dummy variable (=1 if the individual lives in the region "Alentejo"; =0 otherwise)	0.16	0.37	0	1
raaçores	Dummy variable (=1 if the individual lives in the region "Região Autónoma dos Açores"; =0 otherwise)	0.12	0.33	0	1
ramadeira	Dummy variable (=1 if the individual lives in the region "Região Autónoma da Madeira"; =0 otherwise)	0.13	0.34	0	1
male	Dummy variable (=1 if the individual is male; =0 if the individual is female)	0.43	0.50	0	1
age	Age Groups (mid-point is specified)	54.37	18.76	17	87
agesq	The square of the age groups' mid-point	3308.27	2010.50	289	7569
national	Dummy variable (=1 if the individual has Portuguese nationality; =0 otherwise)	0.98	0.15	0	1
noeduclevel	Dummy variable (=1 if the individual did not complete any school level; =0 otherwise)	0.12	0.32	0	1
basiceduc	Dummy variable (=1 if the individual completed the basic education level; =0 otherwise)	0.57	0.50	0	1
seceduc	Dummy variable (=1 if the individual completed the secondary education level; =0 otherwise)	0.16	0.36	0	1
postsecsup	Dummy variable (=1 if the individual completed post-secondary/higher education level; =0 otherwise)	0.16	0.36	0	1
employed	Dummy variable (=1 if the individual is employed; =0 otherwise)	0.43	0.50	0	1

Table A.1 (continued)

Variables	Variable Definition	Mean	S.D.	Min	Max
unemployed	Dummy variable (=1 if the individual is unemployed; =0 otherwise)	0.10	0.29	0	1
noworkforce	Dummy variable (=1 if the individual is not part of the labor force; =0 otherwise)	0.47	0.50	0	1
skillevel1	Dummy variable (=1 if the individual's main job is classified in the skill level 1 according to the ISCO-08 structure; =0 otherwise)	0.06	0.23	0	1
skillevel2	Dummy variable (=1 if the individual's main job is classified in the skill level 2 according to the ISCO-08 structure; =0 otherwise)	0.22	0.41	0	1
skillevel3	Dummy variable (=1 if the individual's main job is classified in the skill level 3 according to the ISCO-08 structure; =0 otherwise)	0.05	0.21	0	1
skillevel4	Dummy variable (=1 if the individual's main job is classified in the skill level 4 according to the ISCO-08 structure; =0 otherwise)	0.11	0.31	0	1
income	Quintiles of the net monthly income per adult in the household	2.90	1.40	1	5
income1	Dummy variable (=1 if the net monthly income per adult in the household belongs to the first quintile; =0 otherwise)	0.21	0.41	0	1
income2	Dummy variable (=1 if the net monthly income per adult in the household belongs to the second quintile; =0 otherwise)	0.22	0.42	0	1
income3	Dummy variable (=1 if the net monthly income per adult in the household belongs to the third quintile; =0 otherwise)	0.20	0.40	0	1
income4	Dummy variable (=1 if the net monthly income per adult in the household belongs to the fourth quintile; =0 otherwise)	0.18	0.39	0	1
income5	Dummy variable (=1 if the net monthly income per adult in the household belongs to the fifth quintile; =0 otherwise)	0.18	0.38	0	1
healthsub	Dummy variable (=1 if the individual has access to a healthcare subsystem; =0 otherwise)	0.18	0.39	0	1
healthins	Dummy variable (=1 if the individual has health insurance; =0 otherwise)	0.18	0.39	0	1
longhealthprob	Dummy variable (=1 if the individual has a long-term health problem; =0 otherwise)	0.60	0.49	0	1
actlim	Dummy variable (=1 if the individual has a physical limitation with impact on daily life; =0 otherwise)	0.34	0.47	0	1
chronicdis	Dummy variable (=1 if the individual suffers or has suffered in the 12 months prior to the survey with a chronic disease[14]; =0 otherwise)	0.65	0.48	0	1
litint	Dummy variable (=1 if the individual has felt little interest in doing the usual things more than half of the days or almost every day in the 2 weeks prior to the survey; =0 otherwise)	0.09	0.29	0	1
feeldep	Dummy variable (=1 if the individual felt depressed more than half of the days or almost every day in the 2 weeks prior to the survey; =0 otherwise)	0.09	0.29	0	1

[14] The chronic diseases considered were: hypertension, arthrosis, back and neck chronic pain, diabetes, allergies and depression.

Table A.1 (continued)

Variables	Variable Definition	Mean	S.D.	Min	Max
distsleep	Dummy variable (=1 if the individual experienced changes in sleeping pattern half of the days or almost every day in the 2 weeks prior to the survey; =0 otherwise)	0.17	0.37	0	1
apetdist	Dummy variable (=1 if the individual experienced changes in appetite half of the days or almost every day in the 2 weeks prior to the survey; =0 otherwise)	0.05	0.21	0	1
badfeel	Dummy variable (=1 if the individual felt useless or guilty half of the days or almost every day in the 2 weeks prior to the survey; =0 otherwise)	0.06	0.23	0	1
slow	Dummy variable (=1 if the individual has moved or talked slower than normal half of the days or almost every day in the 2 weeks prior to the survey; =0 otherwise)	0.03	0.17	0	1
lifesat	Dummy variable (=1 if the individual felt satisfied with life; =0 otherwise)	0.73	0.44	0	1
lifeunsat	Dummy variable (=1 if the individual felt unsatisfied with life; =0 otherwise)	0.19	0.39	0	1
genpract	Dummy variable (=1 if the individual had an appointment with a General Practitioner in the 12 months prior to the survey; =0 otherwise)	0.74	0.44	0	1
preg	Dummy variable (=1 if the individual is pregnant; =0 otherwise)	0.00	0.06	0	1
lighteff	Dummy variable (=1 if the activities the individual performs require light physical effort; =0 otherwise)	0.50	0.50	0	1
modeff	Dummy variable (=1 if the activities the individual performs require moderate physical effort; =0 otherwise)	0.37	0.48	0	1
demandeff	Dummy variable (=1 if the individual performs physically demanding activities; =0 otherwise)	0.08	0.28	0	1
none	Dummy variable (=1 if the individual does not perform an activity that requires physical effort; =0 otherwise)	0.05	0.22	0	1
physicexdays	Number of days in a week in which the individual does physical exercise	1.06	1.95	0	7
exercise	Dummy variable (=1 if the individual does physical exercise at least once a week; =0 otherwise)	0.30	0.46	0	1
healthydiet	Dummy variable (=1 if the individual has balanced and healthy eating habits; =0 otherwise)	0.05	0.21	0	1
nosmoke	Dummy variable (=1 if the individual doesn't smoke; =0 otherwise)	0.82	0.39	0	1
safealcohol	Dummy variable (=1 if the individual didn't consume alcoholic beverages at least once a week in the 12 months prior to the survey; =0 otherwise)	0.61	0.49	0	1
lifecond	Self-reported satisfaction with life conditions	4.41	1.61	1	7
pastsmoker	Dummy variable (=1 if the individual has smoked in the past but wasn't a smoker at the time of the survey; =0 otherwise)	0.22	0.41	0	1
t2020	Dummy variable (=1 if the data was collected in 2020; =0 otherwise)	0.01	0.09	0	1
t2019	Dummy variable (=1 if the data was collected in 2019; =0 otherwise)	0.42	0.49	0	1
t2014	Dummy variable (=1 if the data was collected in 2014; =0 otherwise)	0.57	0.50	0	1
N = 30785					

Table A.2 *Descriptive Statistics by Survey*

	Survey 2014		Survey 2019	
	N=17470		N=13315	
	Mean	S.D.	Mean	S.D.
region	17.76	5.33	17.97	5.32
north	0.15	0.36	0.14	0.35
algarve	0.14	0.35	0.12	0.32
centro	0.21	0.40	0.21	0.41
lisbon	0.11	0.31	0.09	0.29
alentejo	0.15	0.36	0.17	0.38
raaçores	0.11	0.32	0.14	0.34
ramadeira	0.13	0.33	0.13	0.34
male	0.44	0.50	0.43	0.50
age	53.16	18.78	55.96	18.61
agesq	3178.84	1998.83	3478.10	2013.21
national	0.98	0.15	0.98	0.15
noeduclevel	0.13	0.33	0.10	0.30
basiceduc	0.57	0.49	0.57	0.50
seceduc	0.15	0.36	0.16	0.37
postsecsup	0.15	0.36	0.17	0.37
employed	0.42	0.49	0.44	0.50
unemployed	0.12	0.32	0.07	0.25
noworkforce	0.46	0.50	0.49	0.50
skillevel1	0.06	0.23	0.06	0.23
skillevel2	0.21	0.41	0.23	0.42
skillevel3	0.05	0.21	0.04	0.20
skillevel4	0.11	0.31	0.11	0.31
income	2.90	1.42	2.90	1.37
income1	0.22	0.42	0.19	0.39
income2	0.21	0.41	0.25	0.43
income3	0.20	0.40	0.21	0.41
income4	0.19	0.39	0.18	0.38
income5	0.18	0.39	0.18	0.38
healthsub	0.18	0.38	0.19	0.39
healthins	0.16	0.37	0.21	0.41
longhealthprob	0.62	0.49	0.58	0.49
actlim	0.36	0.48	0.31	0.46
chronicdis	0.62	0.49	0.68	0.47
litint	0.09	0.29	0.08	0.28
feeldep	0.10	0.30	0.08	0.28
distsleep	0.17	0.38	0.16	0.36
apetdist	0.05	0.22	0.05	0.21
badfeel	0.06	0.24	0.05	0.22
slow	0.03	0.18	0.03	0.16
lifesat	0.72	0.45	0.75	0.44
lifeunsat	0.20	0.40	0.18	0.38
genpract	0.74	0.44	0.75	0.43
preg	0.00	0.06	0.00	0.07
lighteff	0.47	0.50	0.53	0.50
modeff	0.39	0.49	0.35	0.48
demandeff	0.09	0.29	0.07	0.26
none	0.05	0.21	0.06	0.23
physicexdays	1.13	2.01	0.97	1.87
exercise	0.32	0.47	0.27	0.45
healthydiet	0.05	0.22	0.04	0.19
nosmoke	0.81	0.40	0.84	0.37
safealcohol	0.60	0.49	0.62	0.49
lifecond	4.34	1.64	4.52	1.56
pastsmoker	0.22	0.41	0.21	0.41

Table A.3 *Descriptive Statistics for the Sub-Samples*

	Sub-sample of individuals with Health Insurance		**Sub-sample of individuals without Health Insurance**	
	N= 5621		**N = 25164**	
	Mean	**S.D.**	**Mean**	**S.D**
region	17.55	4.78	17.92	5.44
north	0.12	0.32	0.15	0.36
algarve	0.17	0.38	0.12	0.33
centro	0.17	0.38	0.22	0.41
lisbon	0.17	0.38	0.08	0.28
alentejo	0.16	0.37	0.16	0.37
raaçores	0.10	0.31	0.13	0.33
ramadeira	0.10	0.30	0.14	0.34
male	0.48	0.50	0.42	0.49
age	46.78	15.46	56.07	19.01
agesq	2427.53	1517.83	3505.01	2053.91
national	0.97	0.18	0.98	0.15
noeduclevel	0.02	0.12	0.14	0.35
basiceduc	0.41	0.49	0.61	0.49
seceduc	0.25	0.43	0.13	0.34
postsecsup	0.32	0.47	0.12	0.33
employed	0.70	0.46	0.37	0.48
unemployed	0.06	0.23	0.10	0.31
noworkforce	0.25	0.43	0.52	0.50
skillevel1	0.05	0.22	0.06	0.24
skillevel2	0.31	0.46	0.20	0.40
skillevel3	0.09	0.29	0.03	0.18
skillevel4	0.24	0.43	0.08	0.27
income	3.67	1.34	2.73	1.35
income1	0.10	0.30	0.23	0.42
income2	0.11	0.32	0.25	0.43
income3	0.17	0.38	0.21	0.41
income4	0.24	0.43	0.17	0.38
income5	0.37	0.48	0.14	0.34
healthsub	0.18	0.38	0.18	0.39
longhealthprob	0.46	0.50	0.63	0.48
actlim	0.19	0.40	0.37	0.48
chronicdis	0.54	0.50	0.67	0.47
litint	0.05	0.22	0.10	0.30
feeldep	0.04	0.21	0.10	0.30
distsleep	0.12	0.33	0.18	0.38
apetdist	0.03	0.17	0.05	0.22
badfeel	0.03	0.16	0.06	0.24
slow	0.01	0.11	0.03	0.18
lifesat	0.84	0.37	0.71	0.45
lifeunsat	0.11	0.31	0.21	0.41
genpract	0.70	0.46	0.76	0.43
preg	0.01	0.09	0.00	0.06
lighteff	0.50	0.50	0.49	0.50
modeff	0.38	0.49	0.37	0.48
demandeff	0.10	0.29	0.08	0.27
none	0.02	0.15	0.06	0.23
physicexdays	1.45	2.05	0.97	1.92
exercise	0.43	0.50	0.27	0.44
healthydiet	0.04	0.20	0.05	0.21
nosmoke	0.78	0.42	0.83	0.38
safealcohol	0.55	0.50	0.62	0.49
lifecond	4.88	1.43	4.31	1.63
pastsmoker	0.25	0.43	0.21	0.41
t2020	0.01	0.11	0.01	0.09
t2019	0.48	0.50	0.41	0.49
t2014	0.51	0.50	0.58	0.49

Table A.4 *Sample Characterization (full sample)*

	Absolute Frequency	Relative Frequency (%)
Region (NUTS II)		
Norte	4539	14.74
Algarve	4037	13.11
Centro	6418	20.85
Lisboa	3069	9.97
Alentejo	4970	16.14
Região Autónoma dos Açores	3783	12.29
Região Autónoma da Madeira	3969	12.89
Gender		
Male	13354	43.38
Female	17431	56.62
Age		
15-19	1173	3.81
20-24	1156	3.76
25-29	1143	3.71
30-34	1621	5.27
35-39	2221	7.21
40-44	2612	8.48
45-49	2438	7.92
50-54	2539	8.25
55-59	2677	8.70
60-64	2798	9.09
65-69	2771	9.00
70-74	2525	8.20
75-79	2227	7.23
80-84	1728	5.61
85+	1156	3.76
Education Level		
No Education Level	3607	11.72
Basic Education	17557	57.03
Secondary Education	4782	15.53
Post-secondary/Higher Education	4839	15.72
Employment Status		
Employed	13279	43.13
Unemployed	2936	9.54
Out of the Labor Force	14570	47.33
Skill Level of Main Job (ISCO-08)[15]		
Skill Level 1	1762	5.72
Skill Level 2	6791	22.06
Skill Level 3	1410	4.58
Skill Level 4	3316	10.77
Income		
Quintile 1	6443	20.93
Quintile 2	6913	22.46
Quintile 3	6228	20.23
Quintile 4	5662	18.39
Quintile 5	5539	17.99
Healthcare Subsystem	5652	18.36
Health Insurance	5621	18.26
Chronic Disease	19866	64.53
Year		
2014	17470	56.75
2019	13051	42.39
2020[16]	264	0.86

[15] 17506 individuals answered "not applicable" when asked about the main job.
[16] The data collected in 2020 still refers to the 2019 NHS.

Table A.5 *Sample Characterization by survey*

	Survey 2014 (n=17470)		Survey 2019 (n=13315)	
	Absolute Frequency	**Relative Frequency (%)**	**Absolute Frequency**	**Relative Frequency (%)**
Region (NUTS II)				
Norte	2641	15.12	1898	14.25
Algarve	2484	14.22	1553	11.66
Centro	3592	20.56	2826	21.22
Lisboa	1872	10.72	1197	8.99
Alentejo	2690	15.40	2280	17.12
Região Autónoma dos Açores	1954	11.18	1829	13.74
Região Autónoma da Madeira	2237	12.80	1732	13.01
Gender				
Male	7609	43.55	5745	43.15
Female	9861	56.45	7570	56.85
Age				
15-19	706	4.04	467	3.51
20-24	674	3.86	482	3.62
25-29	692	3.96	451	3.39
30-34	1071	6.13	550	4.13
35-39	1438	8.23	783	5.88
40-44	1595	9.13	1017	7.64
45-49	1413	8.09	1025	7.70
50-54	1460	8.36	1079	8.10
55-59	1429	8.18	1248	9.37
60-64	1496	8.56	1302	9.78
65-69	1492	8.54	1279	9.61
70-74	1284	7.35	1241	9.32
75-79	1217	6.97	1010	7.59
80-84	932	5.33	796	5.98
85+	571	3.27	585	4.39
Education Level				
No Education Level	2244	12.84	1363	10.24
Basic Education	9978	57.12	7579	56.92
Secondary Education	2624	15.02	2158	16.21
Post-secondary/Higher Education	2624	15.02	2215	16.64
Employment Status				
Employed	7416	42.45	5863	44.03
Unemployed	2012	11.52	924	6.94
Out of the Labor Force	8042	46.03	6528	49.03
Skill Level of Main Job (ISCO-08)[17]				
Skill Level 1	985	5.64	777	5.84
Skill Level 2	3710	21.24	3081	23.14
Skill Level 3	841	4.81	569	4.27
Skill Level 4	1880	10.76	1436	10.78
Income				
Quintile 1	3881	22.22	2562	19.24
Quintile 2	3635	20.81	3278	24.62
Quintile 3	3467	19.85	2761	20.74
Quintile 4	3286	18.81	2376	17.84
Quintile 5	3201	18.32	2338	17.56
Healthcare Subsystem	3125	17.89	2527	18.98
Health Insurance	2865	16.40	2756	20.70
Chronic Disease	10.783	61.72	9083	68.22

[17] 17506 individuals in the full sample answered “not applicable” when asked about the main job.

Table A.6 *Sample Characterization (sub-samples)*

	Sub-sample of Individuals with Health Insurance (n=5621)		Sub-sample of Individuals without Health Insurance (n=25164)	
	Absolute Frequency	**Relative Frequency (%)**	**Absolute Frequency**	**Relative Frequency (%)**
Region (NUTS II)				
Norte	667	11.87	3872	15.39
Algarve	957	17.03	3080	12.24
Centro	978	17.40	5440	21.62
Lisboa	959	17.06	2110	8.38
Alentejo	906	16.12	4064	16.15
Região Autónoma dos Açores	590	10.50	3193	12.69
Região Autónoma da Madeira	564	10.03	3405	13.53
Gender				
Male	2670	47.50	10684	42.46
Female	2951	52.50	14480	57.54
Age				
15-19	213	3.79	960	3.81
20-24	219	3.90	937	3.72
25-29	291	5.18	852	3.39
30-34	497	8.84	1124	4.47
35-39	681	12.12	1540	6.12
40-44	802	14.27	1810	7.19
45-49	626	11.14	1812	7.20
50-54	515	9.16	2024	8.04
55-59	525	9.34	2152	8.55
60-64	454	8.08	2344	9.31
65-69	328	5.84	2443	9.71
70-74	233	4.15	2292	9.11
75-79	148	2.63	2079	8.26
80-84	57	1.01	1671	6.64
85+	32	0.57	1124	4.47
Education Level				
No Education Level	87	1.55	3520	13.99
Basic Education	2309	41.08	15248	60.59
Secondary Education	1413	25.14	3369	13.39
Post-secondary/Higher Education	1812	32.24	3027	12.03
Employment Status				
Employed	3907	69.51	9372	37.24
Unemployed	323	5.75	2613	10.38
Out of the Labor Force	1391	24.75	13179	52.37
Skill Level of Main Job (ISCO-08)[18]				
Skill Level 1	279	4.96	1483	5.89
Skill Level 2	1736	30.88	5055	20.09
Skill Level 3	532	9.46	878	3.49
Skill Level 4	1360	24.19	1956	7.77
Income				
Quintile 1	566	10.07	5877	23.35
Quintile 2	641	11.40	6272	24.92
Quintile 3	958	17.04	5270	20.94
Quintile 4	1369	24.36	4293	17.06
Quintile 5	2087	37.13	3452	13.72
Healthcare Subsystem	1012	18.00	4640	18.44
Chronic Disease	3022	53.76	16844	66.94
Year				
2014	2865	50.97	14605	58.04
2019	2684	47.75	10367	41.20
2020[19]	72	1.28	192	0.76

[18] 17506 individuals in the full sample answered "not applicable" when asked about the main job.
[19] The data collected in 2020 still refers to the 2019 NHS.

Table A.7 *Health Prevention Behaviors Characterization by Survey*

	Survey 2014 (n=17470)		Survey 2019 (n=13315)	
	Absolute Frequency	**Relative Frequency (%)**	**Absolute Frequency**	**Relative Frequency (%)**
Frequency of physical exercise in a week (number of days)				
Zero	11922	68.24	9682	72.71
One	798	4.57	405	3.04
Two	1455	8.33	992	7.45
Three	1033	5.91	777	5.84
Four	478	2.74	341	2.56
Five	631	3.61	440	3.30
Six	205	1.17	160	1.20
Seven	948	5.43	518	3.89
Healthy Diet	904	5.17	513	3.85
Smoking	3397	19.44	2184	16.40
Alcohol consumption (weekly basis)	6970	39.90	5108	38.36

Table A.8 *Health Prevention Behaviors Characterization (sub-samples)*

	Sub-sample of Individuals with Health Insurance (n=5621)		Sub-sample of Individuals without Health Insurance (n=25164)	
	Absolute Frequency	**Relative Frequency (%)**	**Absolute Frequency**	**Relative Frequency (%)**
Frequency of physical exercise in a week (number of days)				
Zero	3183	56.63	18421	73.20
One	316	5.62	887	3.52
Two	704	12.52	1743	6.93
Three	544	9.68	1266	5.03
Four	241	4.29	578	2.30
Five	267	4.75	804	3.20
Six	75	1.33	290	1.15
Seven	291	5.18	1175	4.67
Healthy Diet	238	4.23	1179	4.69
Smoking	1259	22.40	4322	17.18
Alcohol consumption (weekly basis)	2502	44.51	9576	38.05

Table A.9 *Health Prevention Behaviors by Education Level*

	Exercise (%)	**Healthy Diet (%)**	**No Smoking (%)**	**Safe Alcohol (%)**
Education Level				
Basic Education	24.83	4.95	81.60	58.63
Secondary Education	40.88	3.22	74.74	61.59
Post-Secondary/Higher Education	50.82	4.94	80.45	58.67

Table B.1 ***Multivariate Probit Average Marginal Effects for the main model (full sample)***

	(1) Exercise (+)	(2) Healthy Diet (+)	(3) No Smoke (+)	(4) Safe Alcohol (+)
north	-0.116***	0.0986*	0.241***	-0.0847**
	(-3.65)	(2.12)	(6.64)	(-2.71)
algarve	-0.0465	-0.357***	0.0814*	0.311***
	(-1.45)	(-6.81)	(2.29)	(9.83)
centro	-0.0680*	0.0112	0.281***	0.153***
	(-2.29)	(0.25)	(8.21)	(5.24)
alentejo	-0.233***	-0.258***	0.138***	0.452***
	(-7.34)	(-5.25)	(3.93)	(14.53)
raaçores	-0.333***	-0.455***	0.0428	0.476***
	(-9.76)	(-7.57)	(1.17)	(14.22)
ramadeira	0.0799*	-0.168**	0.224***	0.509***
	(2.46)	(-3.26)	(6.03)	(15.40)
basiceduc	0.162***	0.0295	-0.0418	0.0110
	(5.09)	(0.70)	(-1.01)	(0.39)
seceduc	0.305***	-0.182**	-0.0418	0.127***
	(7.95)	(-3.15)	(-0.90)	(3.53)
postsecsup	0.423***	-0.0501	0.00698	0.0871*
	(10.04)	(-0.81)	(0.14)	(2.16)
employed	-0.349***	-0.236***	0.260***	0.306***
	(-7.84)	(-3.38)	(6.15)	(7.35)
noworkforce	0.130***	-0.233***	0.582***	0.444***
	(4.17)	(-4.75)	(17.83)	(14.30)
income	0.0784***	0.0331**	-0.00330	-0.0460***
	(10.85)	(2.88)	(-0.40)	(-6.41)
healthsub	0.144***	0.0851*	-0.0621*	-0.0572**
	(6.69)	(2.43)	(-2.46)	(-2.61)
genpract	0.0929***	0.0203	0.135***	0.0780***
	(4.88)	(0.64)	(6.61)	(4.15)
male	0.102***	-0.103***	-0.512***	-1.048***
	(5.89)	(-3.84)	(-26.86)	(-64.00)
age	-0.0296***	-0.0290***	-0.0249***	-0.0202***
	(-13.06)	(-9.34)	(-9.58)	(-9.80)
agesq	0.000128***	0.000331***	0.000419***	0.0000966***
	(5.40)	(10.39)	(14.38)	(4.52)
national	-0.163***	-0.757***	0.195***	0.901***
	(-3.66)	(-12.45)	(4.14)	(20.51)
skillevel2	0.199***	-0.00566	-0.0136	-0.174***
	(4.99)	(-0.09)	(-0.35)	(-4.65)
skillevel3	0.416***	0.0904	0.0961+	-0.211***
	(8.08)	(1.04)	(1.84)	(-4.18)
skillevel4	0.459***	0.0608	0.109*	-0.214***
	(9.72)	(0.79)	(2.26)	(-4.67)
longhealthprob	-0.0208	-0.0351	0.0118	0.150***
	(-0.86)	(-0.90)	(0.44)	(6.36)
actlim	-0.177***			
	(-8.42)			
chronicdis	-0.00739	0.0352	0.118***	-0.0920***
	(-0.31)	(0.88)	(4.41)	(-3.84)
litint	-0.0691+	-0.118+	-0.0478	0.0835*
	(-1.77)	(-1.94)	(-1.09)	(2.26)
distsleep	0.0717**	-0.0635+	-0.0184	0.0529*
	(2.89)	(-1.67)	(-0.63)	(2.21)
apetdist	0.00988	-0.127+	-0.183***	0.0634
	(0.22)	(-1.78)	(-3.72)	(1.46)
badfeel	-0.0773	-0.0196	-0.0263	0.00716
	(-1.58)	(-0.27)	(-0.49)	(0.16)
slow	-0.0239	0.00291	0.0635	0.156**
	(-0.40)	(0.03)	(0.98)	(2.79)
feeldep	-0.0564	-0.141*	-0.0249	0.222***
	(-1.34)	(-2.22)	(-0.52)	(5.59)
preg	-0.562***	-0.266	0.462**	0.567***
	(-4.57)	(-1.09)	(3.08)	(3.53)
lighteff	0.147***			
	(3.69)			

Table B.1 (continued)

	(1) Exercise (+)	(2) Healthy Diet (+)	(3) No Smoke (+)	(4) Safe Alcohol (+)
modeff	0.212***			
	(5.15)			
demandeff	-0.0268			
	(-0.54)			
lifecond	0.0529***	-0.0208*	0.0906***	0.0162**
	(9.92)	(-2.56)	(15.50)	(3.17)
pastsmoker	0.0711***			
	(3.46)			
t2020	0.106	-0.148	0.169+	0.0295
	(1.28)	(-1.04)	(1.65)	(0.35)
t2019	-0.135***	-0.194***	0.0315+	0.0835***
	(-8.06)	(-7.28)	(1.66)	(5.13)
Observations	30785	30785	30785	30785
Wald Chi2	43208.7	43208.7	43208.7	43208.7
Prob > chi2	0	0	0	0
Log lik.	-51611.9	-51611.9	-51611.9	-51611.9

t statistics in parentheses
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.2 *Multivariate Probit Average Marginal Effects for the main model (extended)*

	(1) Exercise (+)	(2) Healthy Diet (+)	(3) No Smoke (+)	(4) Safe Alcohol (+)
north	-0.102**	0.117*	0.233***	-0.0981**
	(-3.19)	(2.49)	(6.39)	(-3.13)
algarve	-0.0283	-0.338***	0.0719*	0.297***
	(-0.88)	(-6.41)	(2.02)	(9.35)
centro	-0.0540+	0.0276	0.273***	0.141***
	(-1.81)	(0.62)	(7.95)	(4.80)
alentejo	-0.216***	-0.238***	0.129***	0.438***
	(-6.77)	(-4.81)	(3.66)	(14.02)
raaçores	-0.327***	-0.444***	0.0387	0.466***
	(-9.56)	(-7.37)	(1.05)	(13.89)
ramadeira	0.0917**	-0.149**	0.216***	0.494***
	(2.82)	(-2.89)	(5.82)	(14.93)
basiceduc	0.171***	0.0255	-0.0447	0.0129
	(5.36)	(0.61)	(-1.08)	(0.46)
seceduc	0.310***	-0.200***	-0.0412	0.136***
	(8.11)	(-3.46)	(-0.89)	(3.80)
postsecsup	0.402***	-0.102	0.0233	0.114**
	(9.53)	(-1.63)	(0.45)	(2.82)
employed	-0.327***	-0.223**	0.253***	0.301***
	(-7.33)	(-3.17)	(5.96)	(7.19)
noworkforce	0.143***	-0.226***	0.580***	0.439***
	(4.62)	(-4.60)	(17.76)	(14.15)
income2	0.00914	-0.0486	0.0577*	0.00482
	(0.35)	(-1.22)	(1.97)	(0.19)
income3	0.0625*	0.0274	0.00958	-0.0787**
	(2.35)	(0.66)	(0.32)	(-3.05)
income4	0.171***	0.0753+	0.0355	-0.145***
	(6.13)	(1.70)	(1.14)	(-5.26)
income5	0.348***	0.218***	-0.0471	-0.235***
	(11.08)	(4.37)	(-1.32)	(-7.46)
healthsub	0.129***	0.0587+	-0.0533*	-0.0406+
	(5.94)	(1.66)	(-2.10)	(-1.84)
genpract	0.0972***	0.0253	0.133***	0.0752***
	(5.11)	(0.80)	(6.50)	(3.99)
male	0.103***	-0.110***	-0.509***	-1.044***
	(5.92)	(-4.11)	(-26.74)	(-63.73)
age	-0.0284***	-0.0282***	-0.0250***	-0.0210***
	(-12.57)	(-9.09)	(-9.65)	(-10.22)

Table B.2 (continued)

	(1) Exercise (+)	(2) Healthy Diet (+)	(3) No Smoke (+)	(4) Safe Alcohol (+)
agesq	0.000118***	0.000323***	0.000420***	0.000104***
	(4.97)	(10.15)	(14.40)	(4.87)
national	-0.128**	-0.731***	0.185***	0.874***
	(-2.87)	(-11.92)	(3.91)	(19.82)
skillevel2	0.196***	-0.0134	-0.0134	-0.168***
	(4.93)	(-0.21)	(-0.35)	(-4.48)
skillevel3	0.405***	0.0686	0.103*	-0.197***
	(7.87)	(0.78)	(1.98)	(-3.90)
skillevel4	0.438***	0.0311	0.123*	-0.196***
	(9.24)	(0.40)	(2.54)	(-4.26)
longhealthprob	-0.0212	-0.0350	0.0121	0.150***
	(-0.88)	(-0.90)	(0.45)	(6.35)
actlim	-0.175***			
	(-8.32)			
chronicdis	-0.00634	0.0356	0.117***	-0.0922***
	(-0.26)	(0.89)	(4.37)	(-3.84)
litint	-0.0666+	-0.116+	-0.0479	0.0821*
	(-1.71)	(-1.92)	(-1.09)	(2.22)
distsleep	0.0722**	-0.0630+	-0.0192	0.0522*
	(2.90)	(-1.66)	(-0.66)	(2.18)
apetdist	0.0112	-0.125+	-0.184***	0.0620
	(0.25)	(-1.76)	(-3.74)	(1.43)
badfeel	-0.0758	-0.0178	-0.0275	0.00634
	(-1.55)	(-0.24)	(-0.51)	(0.14)
slow	-0.0230	0.00395	0.0629	0.154**
	(-0.38)	(0.04)	(0.97)	(2.76)
feeldep	-0.0561	-0.142*	-0.0239	0.222***
	(-1.34)	(-2.23)	(-0.50)	(5.58)
preg	-0.562***	-0.255	0.465**	0.563***
	(-4.56)	(-1.06)	(3.10)	(3.51)
lighteff	0.165***			
	(4.12)			
modeff	0.233***			
	(5.63)			
demandeff	-0.00338			
	(-0.07)			
lifecond	0.0539***	-0.0215**	0.0908***	0.0163**
	(10.13)	(-2.65)	(15.57)	(3.21)
pastsmoker	0.0685***			
	(3.33)			
t2020	0.113	-0.139	0.166	0.0255
	(1.36)	(-0.97)	(1.62)	(0.30)
t2019	-0.132***	-0.186***	0.0310	0.0799***
	(-7.88)	(-6.97)	(1.63)	(4.90)
Observations	30785	30785	30785	30785
Wald Chi2	43256.9	43256.9	43256.9	43256.9
Prob > chi2	0	0	0	0
Log lik.	-51554.2	-51554.2	-51554.2	-51554.2

t statistics in parentheses
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.3 ***Univariate Probit Average Marginal Effects for the main model***

	(1) Exercise (+)	(2) Healthy Diet (+)	(3) No Smoke (+)	(4) Safe Alcohol (+)
north	-0.0793*	0.288***	0.199***	-0.228***
	(-2.48)	(5.63)	(5.31)	(-7.01)
algarve	0.00824	-0.0842	0.00613	0.111***
	(0.25)	(-1.46)	(0.16)	(3.35)
centro	-0.0272	0.211***	0.228***	0.00701
	(-0.90)	(4.29)	(6.44)	(0.23)

Table B.3 (continued)

	(1) Exercise (+)	(2) Healthy Diet (+)	(3) No Smoke (+)	(4) Safe Alcohol (+)
alentejo	-0.188***	-0.0263	0.0679+	0.285***
	(-5.85)	(-0.49)	(1.86)	(8.94)
raaçores	-0.289***	-0.250***	-0.0127	0.329***
	(-8.39)	(-3.87)	(-0.34)	(9.37)
ramadeira	0.124***	0.0482	0.162***	0.358***
	(3.77)	(0.86)	(4.19)	(10.44)
basiceduc	0.216***	0.205***	-0.161***	-0.136***
	(6.56)	(4.53)	(-3.74)	(-4.76)
seceduc	0.391***	0.198**	-0.197***	-0.155***
	(9.72)	(3.15)	(-3.91)	(-4.07)
postsecsup	0.506***	0.301***	-0.145**	-0.181***
	(11.58)	(4.48)	(-2.59)	(-4.28)
employed	-0.308***	0.0116	0.215***	0.128**
	(-6.80)	(0.16)	(5.13)	(3.02)
noworkforce	0.202***	0.0814	0.481***	0.199***
	(6.23)	(1.45)	(13.88)	(6.03)
income	0.0777***	0.0337**	-0.00355	-0.0454***
	(10.64)	(2.80)	(-0.43)	(-6.28)
healthsub	0.122***	-0.0368	-0.0301	0.0292
	(5.60)	(-1.01)	(-1.17)	(1.29)
genpract	0.102***	0.0765*	0.122***	0.0439*
	(5.35)	(2.32)	(5.91)	(2.27)
male	0.111***	-0.0583*	-0.528***	-1.107***
	(6.35)	(-2.14)	(-27.67)	(-65.99)
age	-0.0179***	0.0383***	-0.0454***	-0.0675***
	(-6.68)	(8.04)	(-12.33)	(-23.81)
agesq	0.0000191	-0.000254***	0.000620***	0.000521***
	(0.71)	(-5.83)	(15.66)	(19.15)
national	0.0402	0.197+	-0.0480	0.177**
	(0.79)	(1.81)	(-0.84)	(3.24)
skillevel2	0.186***	-0.0598	0.00556	-0.119**
	(4.62)	(-0.92)	(0.15)	(-3.19)
skillevel3	0.391***	-0.0207	0.134*	-0.111*
	(7.53)	(-0.24)	(2.53)	(-2.17)
skillevel4	0.432***	-0.0813	0.146**	-0.101*
	(9.04)	(-1.05)	(2.97)	(-2.17)
longhealthprob	-0.0230	-0.0358	0.0120	0.153***
	(-0.95)	(-0.93)	(0.44)	(6.33)
actlim	-0.169***			
	(-7.93)			
chronicdis	-0.00987	0.0208	0.116***	-0.0891***
	(-0.41)	(0.52)	(4.22)	(-3.60)
litint	-0.0609	-0.0874	-0.0563	0.0642+
	(-1.53)	(-1.49)	(-1.28)	(1.71)
distsleep	0.0730**	-0.0547	-0.0261	0.0399+
	(2.89)	(-1.45)	(-0.90)	(1.65)
apetdist	0.0210	-0.0515	-0.201***	0.0121
	(0.46)	(-0.71)	(-4.11)	(0.28)
badfeel	-0.0677	0.0234	-0.0346	-0.0160
	(-1.35)	(0.33)	(-0.65)	(-0.35)
slow	-0.0132	0.0409	0.0611	0.157**
	(-0.22)	(0.46)	(0.94)	(2.78)
feeldep	-0.0473	-0.104+	-0.0340	0.194***
	(-1.11)	(-1.72)	(-0.71)	(4.89)
preg	-0.534***	-0.0640	0.414**	0.432*
	(-4.29)	(-0.25)	(2.74)	(2.54)
lighteff	0.267***			
	(6.23)			
modeff	0.331***			
	(7.52)			
demandeff	0.0981+			
	(1.89)			
lifecond	0.0653***	0.0372***	0.0753***	-0.0261***
	(11.65)	(4.22)	(12.20)	(-4.80)

Table B.3 (continued)

	(1) Exercise (+)	(2) Healthy Diet (+)	(3) No Smoke (+)	(4) Safe Alcohol (+)
pastsmoker	0.107***			
	(5.38)			
t2020	0.118	-0.106	0.159	-0.00404
	(1.41)	(-0.75)	(1.64)	(-0.05)
t2019	-0.132***	-0.197***	0.0307	0.0858***
	(-7.94)	(-7.29)	(1.61)	(5.21)
Constant	-0.808***	-3.639***	1.015***	2.596***
	(-8.03)	(-19.01)	(8.89)	(25.90)
Observations	30785	30785	30785	30785
Wald Chi2	3960.3	523.2	3582.3	6423.5
Prob > chi2	0	1.52e-89	0	0
Log lik.	-16545.9	-5454.7	-12347.2	-16757.7

t statistics in parentheses
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.4 ***Multivariate Probit Average Marginal Effects for the main model (sub-sample of individuals with health insurance)***

	(1) Exercise (+)	(2) Healthy Diet (+)	(3) No Smoke (+)	(4) Safe Alcohol (+)
north	-0.0185	0.166	0.0857	-0.0500
	(-0.28)	(1.58)	(1.17)	(-0.73)
algarve	-0.0798	-0.277*	0.0139	0.200**
	(-1.32)	(-2.53)	(0.21)	(3.23)
centro	-0.117+	0.00274	0.0978	0.0512
	(-1.95)	(0.03)	(1.47)	(0.83)
alentejo	-0.183**	-0.0654	-0.00683	0.332***
	(-2.99)	(-0.62)	(-0.10)	(5.21)
raaçores	-0.292***	-0.391**	0.00378	0.242***
	(-4.18)	(-2.70)	(0.05)	(3.38)
ramadeira	0.0599	0.0848	0.158*	0.319***
	(0.86)	(0.73)	(2.01)	(4.39)
basiceduc	-0.0521	-0.676***	0.336*	0.483***
	(-0.37)	(-3.95)	(2.10)	(3.68)
seceduc	0.0362	-0.870***	0.332*	0.532***
	(0.25)	(-4.75)	(2.04)	(3.94)
postsecsup	0.181	-0.664***	0.466**	0.452**
	(1.23)	(-3.55)	(2.78)	(3.26)
employed	-0.251*	-0.0763	0.204+	0.387***
	(-2.27)	(-0.40)	(1.83)	(3.49)
noworkforce	0.113	-0.0798	0.478***	0.353***
	(1.34)	(-0.57)	(5.27)	(4.08)
income	0.0967***	0.0143	-0.0179	-0.0329*
	(5.98)	(0.51)	(-1.01)	(-1.98)
healthsub	0.132**	-0.0111	-0.0536	-0.0813
	(2.75)	(-0.13)	(-0.99)	(-1.64)
genpract	0.0512	0.0196	0.102*	0.0596
	(1.31)	(0.28)	(2.40)	(1.48)
male	0.142***	-0.0972	-0.300***	-0.991***
	(3.80)	(-1.51)	(-7.41)	(-26.27)
age	-0.0116+	-0.0227*	-0.0239**	-0.0305***
	(-1.76)	(-2.38)	(-3.08)	(-4.76)
agesq	-0.0000420	0.000293**	0.000382***	0.000157*
	(-0.57)	(2.82)	(4.29)	(2.22)
national	-0.144	-0.419**	0.0247	0.705***
	(-1.54)	(-2.98)	(0.23)	(7.41)
skillevel2	0.107	-0.0250	-0.0519	-0.258**
	(1.21)	(-0.16)	(-0.58)	(-2.88)
skillevel3	0.285**	-0.0337	0.0593	-0.227*
	(2.76)	(-0.18)	(0.56)	(-2.18)
skillevel4	0.258**	0.00418	0.0792	-0.255**
	(2.64)	(0.02)	(0.79)	(-2.59)

Table B.4 (continued)

	(1) Exercise (+)	(2) Healthy Diet (+)	(3) No Smoke (+)	(4) Safe Alcohol (+)
longhealthprob	0.0332	-0.00418	0.104+	0.0260
	(0.68)	(-0.05)	(1.94)	(0.52)
actlim	-0.192***			
	(-3.77)			
chronicdis	-0.0841+	-0.0364	0.00661	0.0272
	(-1.74)	(-0.43)	(0.12)	(0.54)
litint	-0.0933	0.122	-0.146	0.212*
	(-0.93)	(0.70)	(-1.39)	(2.06)
distsleep	0.0599	-0.143	-0.0601	0.0823
	(1.01)	(-1.36)	(-0.92)	(1.35)
apetdist	-0.0258	-0.114	-0.244*	-0.000956
	(-0.23)	(-0.54)	(-2.07)	(-0.01)
badfeel	-0.0196	-0.510+	0.0235	0.196
	(-0.14)	(-1.66)	(0.16)	(1.37)
slow	0.0279	0.382	0.246	-0.189
	(0.15)	(1.31)	(1.25)	(-1.02)
feeldep	-0.150	-0.105	-0.0216	0.221+
	(-1.28)	(-0.52)	(-0.18)	(1.86)
preg	-0.649**	-0.226	0.334	0.991**
	(-2.98)	(-0.52)	(1.37)	(2.90)
lighteff	0.00934			
	(0.08)			
modeff	0.0697			
	(0.62)			
demandeff	-0.143			
	(-1.14)			
lifecond	0.0522***	-0.0199	0.0751***	0.0350**
	(4.09)	(-0.94)	(5.54)	(2.73)
pastsmoker	0.0740+			
	(1.71)			
t2020	0.204	-0.512	0.273	-0.324+
	(1.33)	(-1.31)	(1.45)	(-1.96)
t2019	-0.0629+	-0.0350	0.0767+	0.0631+
	(-1.72)	(-0.55)	(1.92)	(1.68)
Observations	5621	5621	5621	5621
Wald Chi2	6695.1	6695.1	6695.1	6695.1
Prob > chi2	0	0	0	0
Log lik.	-10547.3	-10547.3	-10547.3	-10547.3

t statistics in parentheses
+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.5 ***Multivariate Probit Average Marginal Effects for the main model (sub-sample of individuals without health insurance)***

	(1) Exercise (+)	(2) Healthy Diet (+)	(3) No Smoke (+)	(4) Safe Alcohol (+)
north	-0.101**	0.101+	0.285***	-0.0896*
	(-2.68)	(1.89)	(6.65)	(-2.48)
algarve	-0.0190	-0.366***	0.108*	0.343***
	(-0.50)	(-6.04)	(2.53)	(9.20)
centro	-0.0279	0.0206	0.332***	0.174***
	(-0.79)	(0.40)	(8.15)	(5.11)
alentejo	-0.229***	-0.288***	0.183***	0.481***
	(-6.06)	(-5.09)	(4.36)	(13.29)
raaçores	-0.312***	-0.456***	0.0571	0.533***
	(-7.79)	(-6.76)	(1.33)	(13.78)
ramadeira	0.116**	-0.196***	0.246***	0.548***
	(3.07)	(-3.33)	(5.69)	(14.41)
basiceduc	0.179***	0.0605	-0.0633	-0.00216
	(5.43)	(1.38)	(-1.46)	(-0.07)
seceduc	0.324***	-0.145*	-0.0596	0.134***
	(7.81)	(-2.28)	(-1.19)	(3.42)

Table B.5 (continued)

	(1) Exercise (+)	(2) Healthy Diet (+)	(3) No Smoke (+)	(4) Safe Alcohol (+)
postsecsup	0.424***	-0.0430	-0.0588	0.122**
	(9.01)	(-0.61)	(-1.02)	(2.67)
employed	-0.378***	-0.254***	0.249***	0.284***
	(-7.68)	(-3.35)	(5.41)	(6.28)
noworkforce	0.139***	-0.234***	0.579***	0.439***
	(4.13)	(-4.41)	(16.35)	(13.09)
income	0.0643***	0.0352**	0.00256	-0.0454***
	(7.82)	(2.75)	(0.27)	(-5.61)
healthsub	0.180***	0.103**	-0.0572*	-0.0442+
	(7.23)	(2.59)	(-1.96)	(-1.77)
genpract	0.107***	0.0268	0.141***	0.0845***
	(4.91)	(0.75)	(6.05)	(3.96)
male	0.0959***	-0.101***	-0.574***	-1.070***
	(4.85)	(-3.41)	(-26.44)	(-58.65)
age	-0.0324***	-0.0283***	-0.0262***	-0.0197***
	(-13.14)	(-8.41)	(-9.29)	(-8.79)
agesq	0.000156***	0.000325***	0.000435***	0.0000954***
	(6.10)	(9.56)	(13.84)	(4.18)
national	-0.156**	-0.819***	0.219***	0.899***
	(-3.02)	(-11.79)	(4.10)	(17.88)
skillevel2	0.203***	-0.00717	0.00662	-0.154***
	(4.51)	(-0.10)	(0.16)	(-3.72)
skillevel3	0.406***	0.148	0.114+	-0.232***
	(6.63)	(1.46)	(1.83)	(-3.87)
skillevel4	0.496***	0.0686	0.127*	-0.233***
	(8.91)	(0.75)	(2.22)	(-4.31)
longhealthprob	-0.0331	-0.0459	-0.0187	0.188***
	(-1.19)	(-1.04)	(-0.61)	(7.00)
actlim	-0.174***			
	(-7.50)			
chronicdis	0.0174	0.0589	0.154***	-0.128***
	(0.63)	(1.29)	(4.97)	(-4.68)
litint	-0.0653	-0.146*	-0.0238	0.0663+
	(-1.54)	(-2.25)	(-0.49)	(1.67)
distsleep	0.0708**	-0.0600	-0.0106	0.0529*
	(2.58)	(-1.47)	(-0.32)	(2.02)
apetdist	0.0136	-0.127+	-0.167**	0.0681
	(0.28)	(-1.67)	(-3.07)	(1.45)
badfeel	-0.0868+	0.0243	-0.0473	-0.0165
	(-1.65)	(0.32)	(-0.82)	(-0.34)
slow	-0.0262	-0.0206	0.0351	0.195***
	(-0.41)	(-0.22)	(0.51)	(3.31)
feeldep	-0.0446	-0.148*	-0.0350	0.219***
	(-0.99)	(-2.20)	(-0.67)	(5.18)
preg	-0.537***	-0.289	0.576**	0.388*
	(-3.59)	(-0.97)	(2.95)	(2.12)
lighteff	0.177***			
	(4.09)			
modeff	0.247***			
	(5.52)			
demandeff	0.00200			
	(0.04)			
lifecond	0.0501***	-0.0186*	0.0928***	0.0110+
	(8.47)	(-2.09)	(14.16)	(1.96)
pastsmoker	0.0645**			
	(2.74)			
t2020	0.0604	-0.0673	0.114	0.157
	(0.61)	(-0.43)	(0.94)	(1.55)
t2019	-0.162***	-0.238***	0.0234	0.0943***
	(-8.56)	(-7.99)	(1.08)	(5.18)

Table B.5 (continued)

	(1) Exercise (+)	(2) Healthy Diet (+)	(3) No Smoke (+)	(4) Safe Alcohol (+)
Observations	25164	25164	25164	25164
Wald Chi2	36437.0	36437.0	36437.0	36437.0
Prob > chi2	0	0	0	0
Log lik.	-40844.0	-40844.0	-40844.0	-40844.0

t statistics in parentheses
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.6 *Univariate Probit Average Marginal Effects for the IV model*

	(1) Exercise (+)	(2) Healthy Diet (+)	(3) No Smoke (+)	(4) Safe Alcohol (+)
feeldep	-1.193**	-1.567*	-1.770***	2.046***
	(-2.65)	(-2.50)	(-4.17)	(5.55)
north	-0.0611+	0.296***	0.213***	-0.238***
	(-1.87)	(6.00)	(5.90)	(-7.56)
algarve	0.00993	-0.0784	0.00762	0.0995**
	(0.31)	(-1.41)	(0.21)	(3.09)
centro	-0.0148	0.216***	0.232***	-0.0135
	(-0.49)	(4.58)	(6.78)	(-0.45)
alentejo	-0.182***	-0.0227	0.0660+	0.260***
	(-5.69)	(-0.44)	(1.88)	(8.09)
raaçores	-0.266***	-0.218***	0.0134	0.277***
	(-7.37)	(-3.32)	(0.36)	(7.41)
ramadeira	0.116***	0.0395	0.144***	0.338***
	(3.53)	(0.73)	(3.83)	(9.86)
basiceduc	0.203***	0.183***	-0.166***	-0.110***
	(6.05)	(4.01)	(-4.03)	(-3.74)
seceduc	0.366***	0.168**	-0.210***	-0.115**
	(8.67)	(2.68)	(-4.37)	(-3.01)
postsecsup	0.474***	0.262***	-0.164**	-0.136**
	(10.15)	(3.86)	(-3.07)	(-3.20)
employed	-0.301***	0.00512	0.193***	0.127**
	(-6.69)	(0.07)	(4.61)	(3.09)
noworkforce	0.190***	0.0741	0.444***	0.189***
	(5.83)	(1.36)	(11.78)	(5.81)
income	0.0750***	0.0311**	-0.00464	-0.0407***
	(10.05)	(2.65)	(-0.59)	(-5.65)
healthsub	0.114***	-0.0410	-0.0353	0.0344
	(5.21)	(-1.17)	(-1.43)	(1.60)
genpract	0.0941***	0.0676*	0.108***	0.0475**
	(4.90)	(2.10)	(5.24)	(2.58)
male	0.0919***	-0.0786**	-0.520***	-0.993***
	(4.77)	(-2.86)	(-24.36)	(-22.59)
age	-0.0167***	0.0369***	-0.0419***	-0.0628***
	(-6.14)	(7.64)	(-10.85)	(-18.31)
agesq	0.0000156	-0.000240***	0.000581***	0.000479***
	(0.58)	(-5.50)	(13.57)	(15.15)
national	0.0429	0.193+	-0.0377	0.155**
	(0.86)	(1.85)	(-0.69)	(2.98)
skillevel2	0.174***	-0.0636	-0.00317	-0.101**
	(4.34)	(-1.02)	(-0.09)	(-2.80)
skillevel3	0.369***	-0.0273	0.115*	-0.0923+
	(7.04)	(-0.33)	(2.26)	(-1.89)
skillevel4	0.416***	-0.0764	0.136**	-0.0936*
	(8.64)	(-1.03)	(2.87)	(-2.10)
longhealthprob	-0.0172	-0.0223	0.0261	0.126***
	(-0.72)	(-0.59)	(0.99)	(5.13)
actlim	-0.145***			
	(-6.09)			
chronicdis	-0.00522	0.0272	0.117***	-0.0912***
	(-0.22)	(0.70)	(4.45)	(-3.88)

Table B.6 (continued)

	(1) Exercise (+)	(2) Healthy Diet (+)	(3) No Smoke (+)	(4) Safe Alcohol (+)
litint	0.410*	0.522+	0.664***	-0.710***
	(2.16)	(1.92)	(3.63)	(-4.39)
distsleep	0.156***	0.0586	0.107*	-0.104**
	(3.86)	(0.96)	(2.46)	(-2.72)
apetdist	0.0988+	0.0521	-0.0676	-0.118*
	(1.81)	(0.61)	(-1.09)	(-2.25)
badfeel	0.296+	0.488*	0.520***	-0.607***
	(1.95)	(2.33)	(3.56)	(-4.72)
slow	0.112	0.205+	0.251**	-0.0638
	(1.42)	(1.85)	(3.10)	(-0.84)
preg	-0.565***	-0.116	0.319*	0.472**
	(-4.63)	(-0.48)	(2.18)	(3.00)
lighteff	0.246***			
	(5.57)			
modeff	0.296***			
	(6.27)			
demandeff	0.0726			
	(1.37)			
lifecond	0.0532***	0.0211+	0.0528***	-0.00532
	(6.80)	(1.85)	(5.79)	(-0.77)
pastsmoker	0.0989***			
	(4.94)			
t2020	0.121	-0.0949	0.156+	-0.0116
	(1.48)	(-0.69)	(1.69)	(-0.14)
t2019	-0.132***	-0.193***	0.0229	0.0855***
	(-8.02)	(-7.22)	(1.24)	(5.36)
Constant	-0.718***	-3.390***	1.043***	2.295***
	(-6.56)	(-12.91)	(9.51)	(16.24)
First-stage regressions (feeldep)				
north	0.0140**	0.0149**	0.0149**	0.0149**
	(2.82)	(2.98)	(2.98)	(2.98)
algarve	0.00245	0.00206	0.00206	0.00206
	(0.51)	(0.43)	(0.43)	(0.43)
centro	0.0104*	0.0108*	0.0108*	0.0108*
	(2.28)	(2.38)	(2.38)	(2.38)
alentejo	0.000140	0.000657	0.000657	0.000657
	(0.03)	(0.14)	(0.14)	(0.14)
raaçores	0.0126*	0.0142**	0.0142**	0.0142**
	(2.54)	(2.86)	(2.86)	(2.86)
ramadeira	-0.00299	-0.00273	-0.00273	-0.00273
	(-0.64)	(-0.59)	(-0.59)	(-0.59)
basiceduc	-0.00609	-0.00823	-0.00823	-0.00823
	(-1.13)	(-1.53)	(-1.53)	(-1.53)
seceduc	-0.0119*	-0.0137*	-0.0137*	-0.0137*
	(-1.97)	(-2.27)	(-2.27)	(-2.27)
postsecsup	-0.0158*	-0.0170**	-0.0170**	-0.0170**
	(-2.45)	(-2.65)	(-2.65)	(-2.65)
employed	0.00221	-0.000771	-0.000771	-0.000771
	(0.35)	(-0.12)	(-0.12)	(-0.12)
noworkforce	-0.00200	-0.000248	-0.000248	-0.000248
	(-0.38)	(-0.05)	(-0.05)	(-0.05)
income	-0.000332	-0.000434	-0.000434	-0.000434
	(-0.31)	(-0.41)	(-0.41)	(-0.41)
healthsub	-0.00375	-0.00390	-0.00390	-0.00390
	(-1.28)	(-1.33)	(-1.33)	(-1.33)
genpract	-0.00420+	-0.00373	-0.00373	-0.00373
	(-1.68)	(-1.50)	(-1.50)	(-1.50)
male	-0.0140***	-0.0154***	-0.0154***	-0.0154***
	(-5.90)	(-6.61)	(-6.61)	(-6.61)
age	0.000541	0.000282	0.000282	0.000282
	(1.31)	(0.69)	(0.69)	(0.69)
agesq	-0.00000222	0.000000928	0.000000928	0.000000928
	(-0.50)	(0.21)	(0.21)	(0.21)

Table B.6 (continued)

First-stage regressions (feeldep)				
national	0.00532	0.00625	0.00625	0.00625
	(0.90)	(1.06)	(1.06)	(1.06)
skillevel2	-0.00507	-0.00368	-0.00368	-0.00368
	(-1.00)	(-0.73)	(-0.73)	(-0.73)
skillevel3	-0.00771	-0.00437	-0.00437	-0.00437
	(-1.29)	(-0.74)	(-0.74)	(-0.74)
skillevel4	-0.00241	0.000953	0.000953	0.000953
	(-0.41)	(0.17)	(0.17)	(0.17)
longhealthprob	0.00377	0.00720*	0.00720*	0.00720*
	(1.22)	(2.36)	(2.36)	(2.36)
actlim	0.0144***			
	(4.33)			
chronicdis	0.00389	0.00472	0.00472	0.00472
	(1.34)	(1.63)	(1.63)	(1.63)
litint	0.403***	0.406***	0.406***	0.406***
	(37.26)	(37.61)	(37.61)	(37.61)
distsleep	0.0719***	0.0731***	0.0731***	0.0731***
	(13.19)	(13.44)	(13.44)	(13.44)
apetdist	0.0672***	0.0683***	0.0683***	0.0683***
	(5.94)	(6.05)	(6.05)	(6.05)
badfeel	0.306***	0.307***	0.307***	0.307***
	(22.97)	(23.07)	(23.07)	(23.07)
slow	0.109***	0.111***	0.111***	0.111***
	(7.16)	(7.31)	(7.31)	(7.31)
preg	-0.0380*	-0.0372*	-0.0372*	-0.0372*
	(-2.48)	(-2.43)	(-2.43)	(-2.43)
lighteff	-0.00979			
	(-1.35)			
modeff	-0.0190**			
	(-2.61)			
demandeff	-0.0175*			
	(-2.17)			
lifecond	-0.00286**	-0.00323**	-0.00323**	-0.00323**
	(-2.82)	(-3.20)	(-3.20)	(-3.20)
pastsmoker	-0.00359			
	(-1.36)			
t2020	0.00498	0.00419	0.00419	0.00419
	(0.56)	(0.46)	(0.46)	(0.46)
t2019	-0.00337	-0.00346	-0.00346	-0.00346
	(-1.38)	(-1.43)	(-1.43)	(-1.43)
lifeunsat	0.0554***	0.0573***	0.0573***	0.0573***
	(11.08)	(11.46)	(11.46)	(11.46)
Constant	0.0156	0.00862	0.00862	0.00862
	(1.02)	(0.62)	(0.62)	(0.62)
/				
athrho2_1	0.241*	0.313*	0.375***	-0.406***
	(2.46)	(2.20)	(3.75)	(-4.51)
lnsigma2	-1.588***	-1.587***	-1.587***	-1.587***
	(-163.06)	(-162.62)	(-162.62)	(-162.62)
Observations	30785	30785	30785	30785
Wald Chi2	4182.8	630.7	4341.0	7702.8
Prob > chi2	0	1.19e-111	0	0
Log lik.	-11352.1	-284.5	-7172.1	-11579.0

t statistics in parentheses
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$