UNIVERSITY OF
BIRMINGHAM

Understanding the Determinants of Health-compromising Behaviours in China

by

LIN FU

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

Department of Economics

Birmingham Business School

College of Social Sciences

University of Birmingham

September 2019

Abstract

Using of the 2011, 2013, and 2015 waves of the China Health and Retirement Longitudinal Study, we identify the role of cognitive abilities on three unhealthy behaviours among the middle-aged and elderly people in China. We find that individuals with higher cognitive ability scores are less likely to be current smokers, heavy alcohol drinkers, and physically inactive.

Based on the same dataset, we investigate the influence of the time to the first cigarette (TTFC) after waking up on smoking cessation among the middle-aged and elderly people in China. We find that the TTFC contributes to smoking cessation. Also, the positive association between the TTFC and the probability of smoking cessation is higher among light smokers than among heavy smokers.

Using of the 2014 wave of the China Family Panel Studies, we investigate the extent to which time preferences may be associated with the probability of being overweight or obese and Body Mass Index (BMI). We find that time preferences are positively associated with the probability of being overweight or obese and BMI in the Chinese context. In addition, the association between time preferences and the probability of being overweight is higher than the corresponding relationship with obesity.

Declaration

I confirm that the thesis is my own work. I have not submitted this thesis, in whole or in part, for any other degree or qualification at this or any other universities.

To my dearest parents

(Xichen Fu & Cailing Hao)

Acknowledgements

It is my great pleasure to express my sincere thanks to supervisors, parents, friends and other PhD colleagues, because this PhD thesis could not be completed without the support and guidance of them. It has been an excellent experience working on this research project over the last four years.

I would like to deliver my deepest and sincere gratitude to my primary supervisor, Professor Alessandra Guariglia. She gave me enormous support through all stages of my PhD study. She helped me narrow down my research topics, helped me solve my problems when I encounter difficulties in running regression models, encouraged me to continue working hard when I lost my confidence, and trained me to be a professional class teacher. Also, her outstanding research results in the area of health economics inspired my research. Moreover, she is a great life mentor who has guided me to be an academic researcher and has encouraged me to deliver the research results at international academic conferences, such as the Chinese Economic Association annual conference. I thank her for her knowledge, wisdom, selflessness, support and encouragement.

I would also like to thank my secondary supervisor, Mr Nicholas Horsewood, for all the knowledge and suggestions he has imparted to me. He gave me insightful and valuable advice on my research. His ideas and inspiration on my work encouraged me to think deeper about my results, which helped me to be more confident when presenting my research results to different audiences. I benefited from every meeting with him.

I am grateful to Professor Sasha Talavera, for his advice in regard to econometric doubts I had to face. I will give my special appreciation to Dr. Danying Li, Dr. Yanan Zhang. Lin Tian, Yu Pei, Wei Jin, Jingmin Zhu, and Chenyang Wang. I am very appreciative for their endless support and comments on my work.

I would like to take this opportunity to acknowledge all the training and support that I have received from the Department of Economics at the University of Birmingham. In particular I would like to thank Marleen Vanstockem, Fiona Barron, and Anne Brookes, for their support and care.

Last but not least, I will give all my gratitude to my parents. I could not finish my studies without their devotion, support and sacrifice. My deepest gratitude to my mother, Cailing Hao, and my father, Xichen Fu, who have given me their support.

Table of Contents

List of Abbreviations

BMI	Body Mass Index
CDC	Center for Disease Control and Prevention
CESD	Center for Epidemiological Studies Depression Scale
CFPS	China Family Panel Studies
CHARLS	China Health and Retirement Longitudinal Study
CHNS	China Health and Nutrition Survey
CPD	Cigarettes Smoked per Day
ELSA	English Longitudinal Study of Ageing
GATS	Global Adult Tobacco Survey
HRS	Health and Retirement Study
IPAQ	International Physical Activity Questionnaire
IQ	Intelligence Quotient
IV	Instrumental Variable
K6	6-item Kessler psychological distress scale
LPM	Linear Probability Models
MET	Metabolic Equivalent of Task
MetS	Metabolic Syndrome
NHANES	National Health and Nutrition Examination Surveys
NHIS	National Health Interview Surveys
NIAAA	National Institute on Alcohol Abuse and Alcoholism
NLSY	National Longitudinal Survey of Youth

OLS	Ordinary Least Squares
SES	Socio-economics Status
TTFC	Time to the First Cigarette
WGOC	Working Group on Obesity in China
WHO	World Health Organisation

List of Tables

Chapter Four

Chapter One: Introduction

1.1 Research background

Healthy behaviours play an important role in improving health and reducing medical costs, while unhealthy behaviours make people at risk of diseases[1]. Smoking, physical inactivity and obesity are three top factors that lead to global mortality as they are associated with higher risks of developing chronic diseases (World Health Organisation, 2009). Additionally, Cook and Moore (2002) show that heavy drinking is significantly associated with organ damage, and results in disability and early death.

Recently, unhealthy behaviours, such as smoking, drinking, physical inactivity, and obesity, have become growing concerns in China. Focusing on smoking, in 2010, 28.1% of adults in China are current smokers. However, the rate of successful smoking cessation is very low, and only 16.9% of ever smokers have quit smoking successfully (Li *et al.*, 2011a). Comparatively, there has been a large drop of the smoking rate in developed countries. For instance, in the United States, the prevalence of current smoking has fallen from 40% in the 1970s to 25% in 2000 (Cutler *et al.*, 2009). It has also been shown that 59.1% of American adults who have ever smoked have quit by 2015 (Babb *et al.*, 2017). Culture differences may be the reason for the differences in the prevalence of smoking between China and other developed countries. Therefore, lack of smoking-related legislations may be responsible for the low quitting rate in China (Lim, *et al.*, 2013).

[1] This part draws on my first year proposal (Fu, 2016)

According to data from the 2007 China Chronic Disease and Risk Factor Surveillance, 35.7% of the respondents are current drinkers. Male and female account for 55.6% and 15.0% of the prevalence of drinking (Li *et al*., 2011b). WHO (2014) shows that 38.3% of the global population aged 15 and older are current drinkers, and 16% of them are acknowledged as heavy drinkers. The drinking rate in China is as high as the average drinking rate in other countries.

Based on data from the 2010 International Collaborative Study of Cardiovascular Disease in Asia, 33.7% of Chinese adults aged 35 to 74 are physically inactive (Muntner *et al*., 2005). Hallal *et al*. (2012) show that approximately one third of adults were physically inactive around the world in 2011. The probability of physical inactivity in Southeast Asia accounts for 17%, while the corresponding probability in America and the eastern Mediterranean reaches 43%. The prevalence in China is same as the average global one, but is higher than the corresponding probability in Southeast Asia.

At the same time, the prevalence of obesity in China has increased significantly over the past decades. Based on the China Health and Nutrition Survey, Mi *et al*. (2015) show that the prevalence of obesity increased from 2.88% to 11.8% in males from 1991 to 2011, and from 4.55% to 11.0% in females during the same period. Similarly, the global prevalence of overweight in 2016 was 39%, and that of obesity was 13% (WHO, 2018).

It is worth mentioning that China is facing an issue of population ageing, because of the declining fertility rate and increasing life expectancy. The proportion of elderly people in

China is expected to be 30% of the total population in 2050, compared with 10% in 2000 (Kinsella and He, 2009; Mai *et al*., 2013). An ageing population leads to a higher economic burden for the government, because an older population often faces more health issues than the younger generation (Chatterji *et al*., 2008; Yang *et al*., 2016; Strauss *et al*., 2010). Therefore, it is necessary to undertake further studies that focus on older people's behaviours in China, in order to improve their health status, as well as decrease burdens on families and society.

1.2 Thesis outline and proposed research questions

The remainder of this thesis is organised as follows. Chapter Two examines the relationship between cognitive abilities and three types of unhealthy behaviours, namely smoking, heavy drinking, and physical inactivity. This study centres on participants who are aged over 45 in China. Chapter Three investigates the influence of the time to the first cigarette after waking up on smoking cessation among middle-aged and elderly people in China. Chapter Four investigates the extent to which time preferences may be related to the probability of being overweight, being obese and to Body Mass Index (BMI). All three Chapters use nationally representative datasets in China. Finally, Chapter five concludes the thesis.

In Chapter Two, we identify the determinants of smoking, drinking, and physical inactivity, focusing on cognitive ability. Cognitive ability, which represents the brain-based ability to understand and respond to new information, is recognised as an important contributor to unhealthy behaviours. Anstey *et al*. (2009) and Weiser *et al*. (2010) show that individuals with higher cognitive ability scores are less likely to be current smokers in Australia and Israel, because of a higher health literacy level and better self-control ability. Based on the same reason, Anstey *et al*. (2009) also show that people with higher cognitive scores are less likely to engage in physical inactivity. To the best of my knowledge, no previous study has examined the association between cognitive ability and heavy drinking in older people. For the first time in the literature, we examine the association between cognitive ability and unhealthy behaviours in the Chinese context. We use the 2011, 2013, and 2015 waves of the China Health and Retirement Longitudinal Study (CHARLS), a nationally representative and longitudinal survey of Chinese residents aged 45 and older and their spouses. In our dataset, on average, 27.6% of the middle-aged and elderly respondents are current smokers. Approximately 1 out of 10 respondents drinks heavily, and 1 out of five respondents is physically inactive. By estimating simultaneous Probit models, and controlling for demographics, health and financial variables, we find that individuals with higher cognitive ability scores are less likely to be current smokers and heavy alcohol drinkers, and less prone to be physically inactive. The results are robust to using an alternative model, which is the multivariate Probit model. However, cognition is not significantly associated with the number of cigarettes smoked and the unit of alcohol drunk. Moreover, this study proves the spousal concordance of unhealthy behaviours among the middle-aged and elderly married couples. Given the context of fast ageing in China, our results provide a better understanding of unhealthy behaviours among the middle-aged and elderly people in China, and suggest that improving cognitive abilities may be effective in reducing the prevalence of unhealthy behaviours, especially among middle-aged and elderly people.

In Chapter Three, we contribute to the literature by investigating the influence of the time to the first cigarette (TTFC) after waking up on smoking cessation among middle-aged and elderly people in China. Many studies have identified the TTFC as a determinant of smoking cessation among adults in other countries (Baker *et al*., 2007; Fagerstrom *et al*., 2003; Hyland *et al*., 2004; Mercincavage *et al*., 2013; West *et al*. 2001). However, little research on this topic has been conducted in the Chinese context. The study fills this gap in the literature. Using the 2011, 2013, and 2015 waves of the CHARLS, we find that the TTFC contributes to smoking cessation for middle-aged and elderly people in China. Also, the positive association is more prevalent among light smokers than among heavy smokers. Our results suggest that more health education should be provided to middle-aged and elderly people in China to increase their awareness of the benefits of smoking cessation. Also, policy makers should advise people to delay the time to their first cigarette after waking up, which could then make smoking cessation easier.

In Chapter Four, we investigate the extent to which time preferences may be related to the probability of being overweight or obese and BMI. To the best of our knowledge, the role of time preferences in explaining obesity in China has not been investigated with national representative datasets. We contribute to the literature by filling this gap. Following the standard from the Working Group on Obesity in China (Zhou, 2002), obesity is present if the BMI is 28 kg/m^2 or greater, and overweight manifests itself if the BMI is 24 kg/m^2 or greater. Making use of the 2014 wave of the China Family Panel Studies, we use a set of linear probability models and Probit models, and find that time preferences are positively associated with the probability of being overweight or obese in the Chinese context. In addition, the association between time preferences and the probability of being overweight is higher than the corresponding relationship with obesity. Based on the results of ordinary least square estimations, time preferences are also positively and significantly associated with BMI. These results are robust to using alternative measures of time preferences and different estimation methods, such as Quantile regressions and Ordered Probit models. Our results suggest that guiding people to value the future could contribute to reducing the prevalence of overweight and obesity in China.

Chapter Two: Understanding the Determinants of Unhealthy Behaviours in China: Do Cognitive Abilities Play a Role?

2.1 Introduction

Unhealthy behaviours in China have recently become an important health concern for researchers, as these behaviours are seriously damaging people's health status (Ezzati and Lopez, 2003; Jin *et al*., 2011; Ng *et al*., 2010; Strauss *et al*., 2010; Yuan *et al*., 1997). It is worth mentioning that the proportion of ageing population is growing rapidly in recent years, and this proportion is expected to be 30% in 2050 (Kinsella and He, 2009; Mai *et al*., 2013). However, a few studies have examined unhealthy behaviours among middle-aged and elderly people in China. It is necessary to examine risky behaviours among aged population, as there are a large number of people and aged people have worse health status than younger generation. Without controlling for the higher prevalence of risky behaviours among olders, China will face a serious economic burden in the future (Chatterji *et al*., 2008; Yang *et al*., 2016; Strauss *et al*., 2010).

Several studies examine the determinants of unhealthy behaviours in China, such as smoking, drinking, and physical inactivity (Chan and Leung, 2015; Chou, 2008; Cui *et al*., 2012; Gao *et al*., 2013; Strauss *et al*., 2010; Wang *et al*., 2017). They find that age, gender, education, income, employment, social capital, and stress are important factors influenced on unhealthy behaviours.

Other studies have identified cognitive ability as a critical determinant of these unhealthy behaviours among adults in countries, such as the Australia and Israel (Anstey *et al*., 2009; Weiser *et al*., 2010). Both Anstey *et al*. (2009) and Weiser *et al*. (2010) find that high cognitive abilities contribute to healthy behaviours. However, to the best of our knowledge, there is no empirical study in the literature examining the relationship between cognitive abilities and unhealthy behaviours in the Chinese context, especially among middle-aged and elderly people. We will fill this gap in the literature by looking at the relationship between cognitive abilities and three types of unhealthy behaviours, namely smoking, drinking, and physical inactivity, among people who are over 45 in China[2].

To this end, we use the 2011, 2013, and 2015 waves of the China Health and Retirement Longitudinal Study (CHARLS). Controlling for demographic variables, health variables and financial variables, we find that individuals with higher cognitive ability scores are less likely to be current smokers and heavy alcohol drinkers, and less prone to be physically inactive. Specifically, people with higher mental intactness have a lower change of smoking, and people with higher memory have lower probabilities of heavy drinking and physical inactivity. However, there is no significant association between cognition and the number of cigarettes smoked and the unit of alcohol drunk. Additionally, we find that the probabilities of these three unhealthy behaviours are positively influenced by the corresponding behaviours of their spouses.

[2] This part draws on my first year proposal (Fu, 2016)

China is facing an issue of population ageing, and a growing number of studies on health behaviours among aged population are needed. Our results provide a better understanding of health behaviours among the middle-aged and elderly people in China, and suggest that more education should be provided targeting middle-aged and elderly people in China to improve their cognitive abilities, which can be useful for reducing the prevalence of unhealthy behaviours.

The remainder of this chapter is laid out as follows. Section 2.2 summarises the existing related studies. Section 2.3 presents our hypotheses about the relationship between cognitive abilities and risky behaviours. Section 2.4 describes our specifications and estimation methodology. Section 2.5 describes the dataset and illustrates some descriptive statistics. Section 2.6 presents the empirical results, and section 2.7 concludes.

2.2 Literature review

2.2.1 What is meant by unhealthy behaviours and why are they important?

Unhealthy behaviours have recently become important concerns for researchers, because these behaviours affect people's health status and death rates. Smoking, drinking alcohol, drug use, unprotected sex, poor diets and sedentary lifestyles are different kinds of unhealthy behaviours (Cawley and Ruhm, 2011). World Health Organisation (WHO, 2009) reports that tobacco use, physical inactivity and obesity are three out of top five global leading risk factors for mortality worldwide, as they enhance the risk of chronic diseases. WHO (2009) also shows that unhealthy behaviours affect countries across all income groups, including high, middle and low groups, which means that unhealthy behaviours may affect every individual in every household no matter how rich or poor they are.

2.2.1.1 Smoking

Cigarette smoking is the biggest avoidable issue in the world. In 2004, around 9% of global deaths were due to smoking (WHO, 2009). According to the data from the WHO's Global Burden of Disease database, 1874 million people in developing countries and 795 million people in developed countries aged 30 years and older are interviewed. Smoking causes an estimated 2.43 million deaths of adults in developed countries, which represents 19% of total adult mortality. Meanwhile, smoking causes an estimated 2.41 million deaths of adults in developing countries in 2000, accounting for 9% of total adult mortality (Ezzati and Lopez, 2003). Moreover, Wang (2015) uses the 1991 to 2011 waves of the China Health and Nutrition Survey (CHNS) from China, and indicates that smoking decreases the likelihood of overweight and obesity.

In China, smoking is really popular. For example, a study based on the Global Adult Tobacco Survey (GATS) China, covering 13,354 participants in China, reports that in 2010, 28.1% of adults were current smokers. 52.9% of men and 2.4% of women were smokers (Li *et al*., 2011a). Strauss *et al*. (2010) further confirm this result, using the 2008 CHARLS pilot with 2,238 respondents who are 45 and older. According to their data, 52% of men report that they are currently smoking.

2.2.1.2 Drinking

There are three types of alcoholic beverages worldwide, including liquor, beer, and wine. Liquor includes all types of beverages with high alcohol content, such as white liquor, whisky, vodka, gin, rum and others. Beer and wine have low alcohol content. In China, rice wine, which is a kind of wine, is also popular, because it has low alcohol content.

Drinking is harmful to personal health and public security. For example, Cook and Moore (2002) summarise the literature on drinking and its consequences. They report that, from the perspective of personal health, chronic heavy drinking may cause organ damage, which may result in disability and early death. Other possible consequences of excessive alcohol consumption include traumatic injuries, cognitive impairment, addiction, reduced productivity, neglect of family responsibilities, and birth defects. Cook and Moore (2002) also show that, from the perspective of public security, the possible effects of alcohol abuse are property damage from accidents, criminal victimization, domestic violence, unwanted sexual encounters and venereal diseases. Additionally, Halme *et al*. (2010) estimate the prevalence of alcohol consumption levels in Finland among people aged above 65, and also investigate the association between heavy drinking and all-cause mortality. Following the combined data from the Health 2000 Study and the official Cause-of-Death Register which covers 1,669 respondents, Halme *et al*. (2010) report the prevalence of heavy drinking (defined as drinking more than 8 standard drinks per week) is 20.3% in males and 1.2% in females. Also, their study shows that a man drinks more than 15 standard drinks per week has a two-fold risk of death compared with an abstinent man. The alcohol consumption is too low for females to analyse the association between drinking and mortality.

Many papers examine the consequences of alcohol consumption among middle-aged and elderly people in China. Yuan *et al*. (1997) estimate the relationship between drinking and mortality among 18,224 men aged 45-64 years in Shanghai. They find that those who consume 1-14 drinks a week show a significant 19% reduction in overall mortality relative to lifelong non-drinkers. Also, heavy drinking, which is defined as consuming 29 or more drinks weekly, is significantly associated with increased risks of death from cancer of the upper aerodigestive tract, hepatic cirrhosis, and stroke. Additionally, based on 15,540 Chinese adults aged 35 to 74 from a cross-sectional survey in 2000-2001, Reynolds *et al*. (2007) find that alcohol consumption is an important risk factor of overweight and obesity in China. Moreover, Jin *et al*., (2011) investigates the relationship between alcohol consumption and metabolic syndrome (MetS) among 19,215 middle-aged and elderly Chinese during the period June 2008-April 2009, which finds that excessive alcohol consumption (in excess of 50 grams of ethanol per day) is associated with higher prevalence of MetS in Chinese men.

Comparatively, other studies suggest low-to-moderate alcohol consumption is beneficial to health. Making use of a cohort of 67,706 middle-aged and elderly respondents in Sweden, Bellavia *et al*. (2014) conduct a dose-response analysis to find the association between alcohol consumption and survival time. They conclude that low alcohol consumption is associated with improved survival of up to 1.5 years for women and up to 1.3 years for men.

Li *et al*. (2011b) analyse the prevalence of drinking among 49,527 Chinese residents over 31 provinces in China. Their data, taken from the 2007 China Chronic Disease and Risk Factor Surveillance, shows that 35.7% of the respondents are current drinkers. Male and female account for 55.6% and 15.0% of the prevalence of drinking.

2.2.1.3 Physical inactivity

According to WHO (2010), there are a variety of physical activities for adults, including physical activities for entertainments, transportations, occupations, and household works. WHO (2010) also shows that the purpose of these physical activities is to improve people's health conditions from cardiorespiratory and muscular aspects, and to reduce the risks of noncommunicable diseases, depression and cognitive declines.

Using the WHO global health observatory data in 2011, Hallal *et al*. (2012) describe worldwide physical inactivity levels in adults aged 15 or older from 122 countries. They conclude that approximately one third of adults have low physically activity, with proportions ranging from 17.0% in southeast Asia to about 43% in the Americas and the eastern Mediterranean[3]. Moreover, according to data from the International Collaborative Study of Cardiovascular Disease in Asia during the period 2000-2001, 33.7% of the Chinese adults aged 35 to 74 are physically inactive (Muntner *et al*., 2005).

---

[3] These values represent the weighted average of the proportion of physical inactivity in the 122 countries, taking into account population sizes.

Based on data from 1.5 million health plan members of the Blue Cross Blue Shield of Minnesota in the US, who are 18 years old and over, matched with the 2000 Behavioural Risk Factor Surveillance System, Garrett *et al*. (2004) estimate the physical inactivity-related medical claim amount among these members in 2000. They find that around one third of colon cancer, heart disease, osteoporosis, and stroke cases in the US are attributable to physical inactivity. Garrett *et al*. (2004) also show that physical inactivity costs $86 million in expenditures for medical treatment within the health plan population in 2000.

Furthermore, based on the cross-sectional dataset from the China Kadoorie Biobank between 2004 and 2008, 466,605 participants aged 30 to 79 join in this survey. Du *et al*. (2003) report that low level of physical activity and excess sedentary leisure time are independently and jointly associated with greater obesity, which has been shown to be at the origin of several cardiovascular diseases, type 2 diabetes, and certain cancers (WHO, 2018). The positive relationship between physical inactivity and obesity is further confirmed by Ng *et al*. (2010), who estimates the effects of physical inactivity on weight in male adults. Ng *et al*. (2010) make use of the 1991, 1993, 1997, 2000, 2004, and 2006 waves of CHNS, which include 10,935 male adults aged from 18 to 55. The results show that 5.4% of weight gain among men in China within the six waves is due to declines in physical activity.

2.2.1.4 Other unhealthy behaviours

Apart from smoking, drinking, and physical inactivity, there exist other types of unhealthy behaviours, such as diet quality and drug use.

According to a follow-up cohort study with 3,982 French community dwellers aged 65 or over during 2001-2012, Artaud *et al*. (2013) consider eating fruit or vegetables less than once a day to be an unhealthy behaviour. They also include the other three unhealthy behaviours in their study, namely physical inactivity, smoking, and drinking. Artaud *et al*. (2013) show that smoking, physical inactivity, and bad diet quality with less fruit and vegetables intake are independently associated with greater hazard of incident disability, while drinking is not significantly associated with the hazard. This study also finds that this hazard increases gradually with the number of unhealthy behaviours. For instance, participants who have three unhealthy behaviours concurrently have a 2.53-fold increased hazard of disability compared with those with no unhealthy behaviours.

Stenbacka *et al*. (2010) estimate the effect of illicit drugs on users' mortality rate in Stockholm, Sweden, based on a cohort of 1,705 identified substance abusers. Participants with substance abuse are identified in 1967, and then they are followed over 37 years until 2003. The authors find that more than half of the respondents die at an average age of 47, which is 25 to 30 years earlier than the general population. The cause of deaths is various by age groups. Stenbacka *et al*. (2010) show that accidents and suicide are the major reasons of death among the youngest respondents, who are between 15 and 24 years old. By contrast, cardiovascular diseases and tumours are the main cause of death for people aged above 55.

2.2.2 General determinants of unhealthy behaviours

2.2.2.1 Studies on countries other than China

2.2.2.1.1 Demographic determinants

Schneider and Schneider (2006) test the effect of education, income, and labour force participation on health behaviours, such as smoking, exercising, and nutrition. Using the German Socio-Economic Panel, which is a representative longitudinal study of private households in Germany, they study 9,282 respondents with average age at 46.4. The results of simultaneous equation model show that both education and income are negatively associated with risky behaviours, namely smoking, physical inactivity, and having bad diet habits. These findings can be explained considering that educated people may have better health knowledge, which is important in understanding the consequences of risky behaviours. Also, richer individuals can afford health-conscious foods and medical services which are not fully covered by the insurance system, thus, they are more likely to have a healthy diet, quit smoking successfully, and do exercise frequently. In addition, Schneider and Schneider (2006) find that employed individuals are more likely to have healthier behaviours than unemployed individuals, such as non-smoking, doing exercise, and following a health-conscious diet, because they have higher opportunity costs for illness. For instance, if an employee is ill, he/she reduces current and future earnings, while an unemployed individual does not have such concerns.

Using a variety of US data sets, Cutler and Lleras-Muney (2010) examine the relationship between education and unhealthy behaviours among individuals aged 25 or above. They focus on broad explanations for unhealthy behaviours rather than explanations specific to any particular behaviour. The 1990, 1991, 1994 and 2000 waves of the National Health Interview Survey (NHIS) are used, and the sample size is up to approximately 23,000. Cutler and Lleras-Muney (2010) find that controlling for age, gender, and parental background, a high educational level is negatively associated with the prevalence of smoking, heavy drinking, being obese, but is positively associated with the prevalence of driving safely, living in a safe house, and using preventive care. The authors repeat their analysis using data in the US from the 3rd wave of Health and Retirement Study (HRS) for middle-aged and elderly people, the 1979 National Longitudinal Survey of Youth, the National Survey of Midlife Development in the United States during the period 1955-1996, the Survey on Smoking, and find similar results to the NHIS data. Cutler and Lleras-Muney (2010) also use the 6th wave of the National Child Development Study in the UK, which consists of nearly 6500 people aged 41 to 42, and find, once again, similar results. The negative relationship between education and unhealthy behaviours can be explained considering that education raises health knowledge, which helps people recognise the harms of risky behaviours and ultimately improves healthier behaviours. Another reason is the difference of individuals' tastes, which means that educated people are more likely to value the future, and avoid unhealthy behaviours. This finding is consistent with Schneider and Schneider's (2006) conclusion.

Iparraguirre (2015) looks at the socioeconomic determinants of alcohol drinking among middle-aged and elderly people in England. Using a sample of 9,251 respondents aged 50 or over from the 2010 wave of the English Longitudinal Survey of Ageing (ELSA), the results show that good health status and higher education are positively associated with the probability of being a heavy drinker for both males and females. The positive relationship between education and heavy drinking is inconsistent with Cutler and Lleras-Muney (2010). This inconsistence may be caused by generational trait. Iparraguirre (2015) focuses on middle-aged and elderly people, while the other two studies focus adults. Moreover, single, separated or divorced respondents show a higher likelihood of heavy drinking in males. It is possible that heavy drinkers within marriage always receive drinking cessation advice from their spouse.

Palipudi *et al*. (2012) investigate the determinants of current smoking in several low-and-middle income countries. Nationally representative data from the GATS is conducted over the period 2008–2010 in 13 low-and-middle income countries, which covers 209,027 respondents aged 15 and older. The results show that the smoking rate is higher for males across all selected countries and for urban areas in 8 of the 13 countries. Consistent with the findings in developed countries (Cutler and Lleras-Muney, 2010; Schneider and Schneider, 2006), education and wealth are significantly inversely associated with the prevalence of smoking in many low-and-middle income countries, excluding China.

2.2.2.1.2 Other determinants

Unemployment has a strong association with unhealthy behaviours. Ruhm (2005) examines the changes of lifestyle when unemployment rate increases. He uses data from the 1987 to 2000 waves of the Behavioural Risk Factor Surveillance System for American adults, where more than 50,000 participations are surveyed annually, and show that lifestyles become healthier during temporary economic downturns. Ruhm (2005) uses the state unemployment rate as an alternative proxy of macroeconomic conditions, and reports that a one-point drop in the percentage of employed people is associated with a reduction in the prevalence of smoking, obesity, severe obesity, physical inactivity and multiple health risks by 0.6, 0.4, 0.8, 0.7 and 1.1 percentage points, respectively. These findings may be caused by the fact that working hours decline in bad times. Therefore, people have more time to undertake exercises. Also, falling working hours may reduce the job-related stress, and people's demand for tobacco reduces accordingly.

However, there are opposite point of views in terms of the association between unemployment and unhealthy behaviours. Using data from the 2002 to 2010 National Survey on Drug Use and Health in the US, which consists of 405,000 adult participants aged 18 or older. Compton *et al*. (2014) compare the health-risk factors among unemployed and employed people in the periods 2002–2004, 2005–2007, 2008, and 2009–2010, which cover the periods before and after the 2008 economic crisis. Controlling for demographic characteristics, mental health, and local county unemployment rates, their results reveal that individual unemployment is associated with higher rates of smoking, heavy alcohol drinking, illicit drug use. These positive associations persist during the economic distress. The possible reason for the opposite conclusion with Ruhm (2005) is stress-reactivity or impulsivity, which mean that unemployed people are more likely to undertake risky behaviours to alleviate their stress from unemployment, and they are less concerned about the damage to their body caused by these risky behaviours. Additionally, Latif (2014) confirms the positive associations between unemployment rate and risky behaviours, namely alcohol drinking and smoking. Using longitudinal data from the Canadian National Population Health Survey in the period 1994-2009, which includes 31,171 observations aged 16 to 64 in each wave, Latif (2014) investigates the impact of provincial unemployment rate on individual alcohol drinking and smoking behaviours in the Canada. Controlling for individual characteristics, such as age, marital status, education, home ownership, income, and mental health status, he finds that the unemployment rate has a significantly positive association with the probability of drinking, especially among males. Although he finds no evidence for the association between unemployment and the probability of being a smoker, he finds a significant positive association between the unemployment and the number of cigarettes smoked daily among the daily smokers, especially for males. Latif (2014) interprets his findings considering that the leisure or stress effect of unemployment dominates the income effect, leading to an increase in drinking and cigarettes smoked particularly for the male group.

Social factors are known as crucial determinants for unhealthy behaivours. Using data from the 2004 wave of the English Longitudinal Study of Ageing (ELSA), which consists of 8,688 individuals aged 50 and older in England. Shankar *et al*. (2011) investigate the impact of social isolation and loneliness on health-related behaviours, namely smoking and physical inactivity. Loneliness represents a subjective feeling of social activities, whilst social isolation indicates objective social activities, including marital status, frequency of contacting with friends and relatives, and participation in social activities. Controlling for age, gender, whether the respondent is ill or not, depression, and wealth, Shankar *et al*. (2011) find that both social isolation and loneliness are independently associated with a higher likelihood of smoking, physical inactivity, and both these two health-risk behaviours. This may be caused by the fact that isolated people may lack supports from others, thus, they have less motivations for adopting and maintaining healthier behaviours.

Several studies point out the spousal concordance of unhealthy behaviours. For example, Li *et al*. (2013) estimate the role of the spouse in shaping physical activity behaviour among 5,074 married couples aged 50 and over in the US. Using the 2004 to 2010 waves of the HRS, they find that physical inactivity is interrelated between couple members. It implies that the respondent's physical activity level is positively predicted by that of his/her spouse's. The possible reason is because of the shared resources, such as the physical environment, social networks, and financial resources. Also, assortative mating is another reason for this phenomenon, as people are attracted to marry those who share similar characteristics. Additionally, Stimpson *et al*. (2006) examine the spousal effect of the unhealthy behaviours among older husbands and wives. Three types of behaviours are considered, including body fatness, smoking, and alcohol consumption. A cohort of 553 Mexican American couples, who were 65 or over, participated in the Hispanic Established Populations for the Epidemiologic Studies of the Elderly in 1993. Stimpson *et al*. (2006) run the multivariate regression model, and prove the significant spousal effect among the older Mexican American couples. In detail, the individual's BMI is positively related to the corresponding value of his/her spouse. The probability of smoking and drinking by the respondent would be higher if his/her spouse had smoked cigarettes or drunk alcohol ever. The possible reason for the findings is that couple members share the same household context.

2.2.2.1.3 Determinants of multiple unhealthy behaviours

It is common that people with one unhealthy behaviour tends to have more unhealthy behaviours concurrently. This co-occurrence is normally called multiple unhealthy behaviours (Chou, 2008; Laaksonen *et al*., 2003; Schuit *et al*., 2002). Using data from the MORGEN study between 1994 and 1997, which covers 16,789 respondents aged 20–59 years, Schuit *et al*. (2002) explore the determinants of multiple lifestyle factors in the Netherlands. Four lifestyle risk factors are smoking, drinking, physical inactivity, and low vegetable and fruit consumption. The results uncover that around 20% of the respondents adopt three or four unhealthy behaivours simultaneously, and only 10% of the respondents do not have any lifestyle risk factor. Schuit *et al*. (2002) find that the higher the number of unhealthy behaviours clustering together, the lower the respondents' educational level and self-rated health. Also, they show that unemployed respondents who have experienced a deterioration in health in the past year, and males living without partner tend to have multiple lifestyle risk factors. The inverse association between health status and clustering can be explained considering that respondents' attitudes to self-reported health status reflect the respondents' awareness of potential harmful effects of unhealthy behaviours. Namely, those respondents who adopt several unhealthy behaviours simultaneously are aware of the possible future negative effects of the behaviours.

Laaksonen *et al*. (2003) explore the determinants of multiple unhealthy behaviours among Finnish adults in the 1990s. The data is collected from annual cross-sectional surveys within the project for Monitoring Health Behaviour, which consists of 12,192 men and 13,822 women from 1991 to 1998. The study uses a combination of the four behaviours, including smoking, heavy alcohol drinking, physical inactivity and unhealthy diet. Laaksonen *et al*. (2003) find that the probability of engaging in three or four unhealthy behaviours is lower among the respondents who are older, married, and have higher educational level in both males and females, but it is higher in urban areas than rural areas in females. The authors explain that marriage prevents male and female adults from unhealthy behaviours, because married people are more likely to receive smoking cessation advice and supports from their spouses. Also, due to the different levels of urbanization, the social resources in urban and rural areas are different.

2.2.2.2 Studies on China

Strauss *et al*. (2010) examine the relationships between socio-economic status (SES) and smoking among Chinese men. The data contains 1,111 respondents aged 45 or older from the 2008 CHARLS pilot data fielded in the Zhejiang and Gansu provinces. Two indicators of SES, education and income, are used in their analysis. The results show that current smoking has no relationship with education, but has a significantly positive relationship with income. Strauss *et al*.'s (2010) results demonstrate the association between SES and smoking in China, which is different from most previous studies in western developed countries, such as the US, the UK, and Germany (Cutler and Lleras-Muney, 2010; Schneider and Schneider, 2006). The possible explanation is that the level of economic development in China is different from the corresponding levels in developed countries.

Also, Chan and Leung (2015) address the association between sociodemographic factors and unhealthy behaviours among Hong Kong people. Making use of a sample of 459 observations with average age at 54.8, the study finds that a specific group of people, who are males, with younger age, have higher educational levels, and employed, are more likely to have risky behaviours. Specifically, they are more likely to smoke cigarettes and drink alcohol, but are less likely to exercise daily. Chan and Leung (2015) explain that Chinese culture encourages smoking and drinking in males, but not in females. Additionally, they explain that sedentary lifestyles are common in western countries, and westernization in Hong Kong tends to affect young men with higher socioeconomic status, because these group of people are considered as elites and are more likely to be exposed to Western lifestyles. By contrast, older women in Hong Kong with lower educational level are more likely to do exercises, because a majority of them are homemakers, which provide them with enough time engaging in regular physical activities.

This finding is confirmed by Wang *et al*. (2017), who emphasise gender differences in several aspects, such as mental health, smoking, drinking, and chronic diseases, in older adults in China. The authors use the Jilin Provincial Chronic Disease Survey in 2012, which includes 4,115 people aged 60 to 79. The results find that women, compared with men, are less likely to be current smokers and frequent drinkers. This can be explained considering the strong cultural and social pressure against women smoking and drinking.

Stress is also considered as an independent variable for smoking behavior in China. Cui *et al*. (2012) examine the association between stress and smoking among rural–urban migrants by involving 1,595 respondents aged 18 or older in the research. Both perceived work stress and perceived life stress are examined in their analysis. The regression results show that compared with their counterparts with low stress, those who have high perceived life stress are 1.45 times more likely to smoke, while perceived work stress has no association with the prevalence of smoking. The possible reason is that life stress includes work stress, and life stress is a better indicator of stress for individuals. Also, Cui *et al*. (2012) explain that rural–urban migrants lack stress-management skills and awareness of tobacco control, thus, they are more likely to smoke to reduce their anxiety. Additionally, they find that the respondents who work in construction have a higher prevalence of smoking than for those works in other occupations. This can be explained considering that smoking is prevalent in construction industry, and people are influenced by colleagues.

Gao *et al*. (2013) examine the relationship between social capital and the likelihood of smoking among Chinese male workers. Using data from a survey in Shanghai, 1,603 male employees from 35 workplaces are interviewed in 2012. Social capital is a multidimensional concept, and includes the shared values, attitudes, norms of trust and reciprocity, and practices of collective action in the workplace. Controlling for variables, including age, education, marital status, occupation, and work stress, Gao *et al*. (2013)

observe that the prevalence of smoking is lower among employees with high social capital than their counterparts with low social capital. This finding is consistent with Shankar *et al*. (2011). Three reasons are used for explanation. Firstly, smoking is always used to deal with stress, but high social capital helps people improve their coping abilities, which reduces workers' motivations for cigarettes. Secondly, high social capital means they are more likely to receive more instrumental and emotional supports from their colleagues. The third possible reason for the negative relationship is that people with high social capital in workplace have good self-control on smoking, because they have a strong belief that smoking is harmful to their health.

Chou (2008) investigates the determinants of multiple risk behaviours in Hong Kong. Using data from a survey which is conducted in Hong Kong in 2004 to assess the health status of older adults, 4,812 Chinese women and men over 60 are interviewed. Chou (2008) uses four common lifestyle risk factors in the analysis, including cigarette smoking, heavy alcohol drinking, physical inactivity, and lack of vegetables and fruits consumption. He/she finds that higher educated men belonging to the older age groups tend to show a higher degree of clustering, and employed people tend to display a lower degree of clustering. The possible reason for the positive association between age and unhealthy behaviours is that older adults may not know the health benefits of reducing their unhealthy behaviours, or they feel no need to improve their health.

Consistent with the previous findings in the US (Li *et al*., 2013; Stimpson *et al*., 2006), Jurj *et al*. (2006) estimate the extent to which a wife's smoking, drinking, and physical activity behaviours are influenced by her husband's corresponding behaviours. Based on a survey with 66,130 married women aged 40 to 70, who participated in the baseline survey of the Shanghai Women's Health Study between 1997 and 2000, the authors show that the wives would be more than twice as likely to be smokers, regularly drinkers, and to be physically active, if their husbands had the same habits. The shared marital environment may contribute to the spousal concordance of the health behaviours between husbands and wives.

2.2.3 The relationship between cognitive abilities and unhealthy behaviours

2.2.3.1 Definition of cognitive ability

There is no specific definition for cognitive ability. In some previous studies, term intelligence is used to replace cognitive ability. Hunter (1986) explains that intelligence focuses mainly on the potential development of the respondent's ability that is due to genes, while cognitive abilities measure the current abilities. Therefore, intelligence is more suitable for children, and cognitive abilities apply to mature people. Specifically, gene comes with the children from birth, and is determined by the genes from their parents (Trzaskowski *et al*., 2014). For example, some children start to speak earlier, learn new information faster than their peers, and we expect they will be smarter than their peer in the future as well, controlling for all the other factors. This is intelligence. When the children grow up, all their experiences, such as wealth level, education, location, friends, and so on, determine their mental development, and their cognitive ability (Hunter, 1986). Therefore, we can define cognition as the way an individual understands and responds to new information. Cognitive abilities are brain-based skills that focus mainly on the mechanisms of learning, memorising, problem-solving and concentrating (Lei *et al*., 2012; Lei and Liu, 2018; McArdle *et al*., 2007).

Cognitive abilities could be determined by a bunch of factors, such as education, environment, genes, gender, and so on[4]. Using a number of data sets from the UK and the U.S, Cutler and Lleras-Muney (2010) claim that cognitive abilities are positively associated with education. This finding is confirmed by Banks *et al*. (2010) among the middle-aged and elderly people within the ELSA 2002 survey. Banks *et al*. (2010) find that numerical ability, which is the most representative cognitive ability, is positively correlated with education status. Also, they find that the numerical ability of males exceeds that of females in all education groups, and the numerical ability of younger individuals is greater than that of their older counterparts. Additionally, a study of 839 twin pairs who completed the National Merit Scholastic Qualifying Test in 1962 shows that external variables such as environment impact on the cognitive abilities for adolescents from poorer backgrounds, while internal variables such as genes account for a higher cognitive ability from affluent homes (Harden *et al*., 2007).

2.2.3.2 The relationship between cognitive abilities and unhealthy behaviours

2.2.3.2.1 The significantly negative relationship

Anstey *et al*. (2009) investigate the association between cognitive function in adulthood and health behaviours in Australia. Their sample comes from the first and second waves

[4] This part draws on my first year proposal (Fu, 2016)

of the PATH Through Life Project over the period 1999-2006, which is a large community survey with information on health and wellbeing of people aged in their 20s, 40s, and 60s. Two measurements of cognitive function are used in their study, including verbal ability and processing speed. With control variables, such as demographic variables and self-rated health status, Anstey *et al*. (2009) show that for the 40s and 60s, higher level of verbal ability and processing speed are associated reduced likelihood of current smoking and lower level of physical activity, but are associated with higher probability of light or moderate alcohol consumption. They find that cognitive function has negative relationships with smoking and physical inactivity. The possible reason for these relationships is that the respondent with a higher cognitive function is more likely to have an improved health literacy, which ultimately contributes to engaging in healthier behaviours. By contrast, the authors find that cognitive function is positively associated with light to moderate drinking. This may be caused by the fact that many people consider a small amount of alcohol drinking as beneficial for health.

Weiser *et al*. (2010) test the differences in cognitive functioning between the smokers and non-smokers. Their sample covers 20,221 male subjects from the Israeli military recruits in 1988. The intellectual assessments are tested at ages 16-17 before recruited to military service, and the smoking status is analysed when they were recruited to the Israeli Defence Forces at 18. Weiser *et al*. (2010) find that smokers have significantly lower cognitive test scores compared to the non-smokers, and this negative association remains after controlling for SES. They run an additional test among 70 sibling pairs, and show that the smoking brother has poorer cognitive test scores than the non-smoking brother. This result means that in addition to genetic and environmental factors, cognitive ability is an important determinant of smoking. This can be explained considering that intelligent individuals are more likely to control their curiosity about tobacco, when informed that smoking is harmful to their health. This finding is consistent with Anstey *et al*. (2009).

Moreover, Kubička *et al*. (2001) aim to identify the childhood personality predictors of drinking and smoking behaviours in adults. Based on a 24-year follow-up study in Prague in the Czech Republic, 440 subjects, who are born in 1960s, participate in five examinations at ages 9-10, 21-23, 28-31, and 32-35. The results show that a low intelligence quotient (IQ) observed in childhood is a risk factor of being a smoker in adulthood, but childhood IQ has no relation to drinking in adulthood. The authors speculate that the effect of intelligence on smoking is due to education, because subjects with low IQ rarely attain a higher level of education. Therefore, people with low IQ are less likely to be exposed to information on health risks associated with smoking.

2.2.3.2.2 The insignificant relationship

Several studies show that childhood cognitive ability is not significantly associated with unhealthy behaviours in adults. For instance, Taylor *et al*. (2005) examine the association between childhood IQ and smoking behaviour in Scottish adults. Their data includes 938 respondents from the Midspan prospective studies in the 1970s, who are matched with their cognitive ability test results on the Scottish Mental Survey 1932. Taylor *et al*. (2005) show that intelligence at age 11 is not directly associated with smoking, although it has indirect effects on smoking consumption via deprivation. Specifically, a lower childhood

IQ at age 11 is associated with greater deprivation, and individuals living in more deprived areas are more likely to smoke.

Mortensen *et al*. (2005) examine the relationship between intelligence in young adults and drinking habits later in life among Danish men. A total of 1,799 men participates in the prospective cohort study in the area of Zealand in Denmark. The intelligence test is completed on individuals at 18 years old between 1956-1977. Two fellow-ups, which collect information on alcohol consumption, are conducted in 1981-1983 and 1992-1994, separately. The authors define heavy drinking as drinking more than 21 drinks per week, and a drink in Denmark contains approximately 12g of ethanol. The results find that IQ is not associated with heavy drinking, which is similar to the study published by Kubička *et al*. (2001).

Also, Batty *et al*. (2006) investigate the relationship between childhood IQ and alcohol induced hangovers in middle aged Scottish men and women. A cohort of 12,150 people, who were born between 1950 and 1956, participated in the IQ test in 1962. Between 2000 and 2003, 7,184 people were re-surveyed with drinking behaviour. Self-reported hangover is defined as drinking two or more occasion per month, and it is a proxy for binge drinking. Batty *et al*. (2006) report that higher IQ scores in childhood are negatively associated with the prevalence of hangovers in middle age. Moreover, they find that the test of IQ at 11 is more strongly associated with hangovers than those administered at either 7 or 9 years, because the later IQ test results are more representative for the mental ability of adult life. However, the association is considerably weakened after controlling for adult indicators of socioeconomic position, such as education.

A similar analysis is conducted by Batty *et al*. (2007a) with data from the Aberdeen Children of the 1950s study. In 1962, 12,150 children participated in this school-based survey at age 11, when the IQ scores were assessed. Forty years later, 7183 participants were re-surveyed, and their information on risk factors for premature mortality were reported. The aim of their study is to investigate the association between IQ in childhood and risk factors in middle-aged people, namely, smoking, heavy alcohol drinking, obesity, and so on. Batty *et al*. (2007a) find that childhood IQ is negatively associated with the prevalence of smoking and heavy alcohol consumption. After statistically controlling for socioeconomic position, particularly in adulthood, the association between IQ and any unhealthy behaviour disappears. This is consistent with the findings in Kubička *et al.* (2001), and can be explained considering that IQ and education are strongly related.

Referring to physical activity, Batty *et al*. (2007b) investigate the relationship between childhood mental ability and physical activity in adulthood, using the 1970 British cohort study. This longitudinal study starts the initial survey in 1970, and five follow-up surveys are conducted in 1975, 1980-1981, 1986, 1999-2000, respectively. The cognitive ability is tested in 1980-1981 when 8,282 participants are 10 year old. The physical activity levels are collected in 1999-2000 when the respondents are 30. The study finds that mental ability is positively related to the prevalence of physical activity in both genders, in particular, intense activity, which means that people with higher childhood mental ability are more likely to do exercises. However, the positive relationship is attenuated after adjusting for socioeconomics position, such as education.

2.2.4 Contribution

This study investigates the relationship between cognitive abilities and three major types of unhealthy behaviours, namely smoking, heavy drinking, and physical inactivity, among people who are 45 and older in China. To this end, the 2011, 2013, and 2015 waves of the CHARLS datasets are used in the analysis. To the best of our knowledge, a few studies on this relationship have concentrated on older people, and no empirical study has been conducted investigating this association in the Chinese context. Our results provide a better understanding of unhealthy behaviours among the middle-aged and elderly people in China.

2.3 Hypotheses

This section discusses the hypotheses on the associations between cognitive abilities and risky behaviours that we will test.

2.3.1 Hypothesis on smoking

Smoking is the most damaging health behaviour, with tobacco use accounting for around 9% of mortality around the world (WHO, 2009). Anstey *et al*. (2009) and Weiser *et al*. (2010) provide compelling evidence in Australia and Israel, and show that individuals with higher cognitive scores have reduced likelihood of current smoking. This can be explained considering that people with higher cognitive function tend to have adequate health literacy, which is important in understanding the negative effects of cigarette smoking. Also, people with higher cognitive ability may have higher level of self-control ability, because they value the future than their counterparts with lower cognitive ability.

Additionally, Kubička *et al*. (2001) find that intelligence in childhood is negatively associated with smoking in adulthood. It is common that cognitive ability changes over time, and some extent of decline in cognitive ability occurs when people get older. However, many studies show that intelligence in childhood contributes to a large extent to cognitive ability in older age (Deary *et al*., 2000; Deary, 2014). This means that people with higher IQ in childhood are more likely to have a higher level of cognitive ability in older age. Therefore, it is reasonable to expect that middle-aged and elderly people with higher cognitive abilities are less likely to smoke cigarettes, compared with their counterparts with lower cognitive abilities. Also, those with higher cognitive abilities are less likely to smoke a large number of cigarettes per day. Our first hypothesis takes therefore the following form:

H1. Higher cognitive abilities are associated with a lower probability of smoking, as well as a lower number of cigarettes consumed per day, among the middle-aged and elderly in China.

2.3.2 Hypothesis on heavy drinking

The negative effects of drinking, especially heavy drinking, pose significant economic burdens for countries (Cook and Moore, 2002). Anstey *et al*. (2009) report that higher cognitive abilities are associated with a higher prevalence of light or moderate drinking in middle-aged and elderly people in Australia. To the best of our knowledge, no previous study has examined the association between cognitive abilities and heavy drinking in older people[5]. We will focus on investigating this topic in China. We expect that people with high cognitive abilities tend to behave healthier than their counterparts with low cognitive abilities, because they have a better understanding on the consequences of heavy drinking, and tend to avoid the damage of excessive alcohol use. Therefore, we propose the following hypothesis:

H2. Higher cognitive abilities are associated with a lower probability of heavy drinking, as well as a lower number of alcohol drinks consumed per week, among the middle-aged and elderly in China.

2.3.3 Hypothesis on physical inactivity

Physical inactivity causes a high rate of diseases, and increases national medical pressures (Du *et al*., 2003; Garrett *et al*., 2004; Ng *et al*., 2010). Anstey *et al*. (2009) report that in comparison with their lower scoring counterparts, people with higher cognitive ability scores are less likely to engage in physical inactivity. This can be explained considering the improved health literacy and good self-control. In line with their findings, our third hypothesis takes the following form:

H3. Higher cognitive abilities are associated with a lower probability of undertaking physical inactivity among the middle-aged and elderly in China.

[5] This part draws on my first year proposal (Fu, 2016)

2.3.4 Hypothesis on spousal effect

The older spouses have a concordance in their health behaviours, such as smoking, alcohol drinking, and physical inactivity (Li *et al*., 2013; Stimpson *et al*., 2006; Jurj *et al*., 2006). In China, Jurj *et al*. (2006) also find that people's unhealthy behaviours are more likely to be influenced by the spousal unhealthy behaviours, because the couple members share the same environment. Moreover, Li *et al*. (2013) provide a further explanation, and show that the assortative mating matters, because people are attracted to marry those who share similar characteristics. In line with their findings, we expect that the respondent would be more likely to smoke, drink heavily, and be physically inactive, if his/her spouse had the same habits. Therefore, we propose the fourth hypothesis:

H4. The spousal unhealthy behaviours are positively associated with the respondent's corresponding behaviours among the middle-aged and elderly in China.

2.4 Specifications and estimation methodology

This section includes the specifications used in this chapter.

2.4.1 Baseline specifications

To document the effect of cognition on health behaviours among the middle-aged and elderly people in China, we apply a simultaneous Probit model. This can be seen as a system of three seemingly unrelated Probit models. We run the regressions for the probability of smoking, heavy drinking, and physical inactivity concurrently. Our simultaneous Probit model takes the following form:

$$\begin{cases} Smoking_{i,t} = \beta_0 + \beta_1 M_{i,t-1} + \beta_2 MI_{i,t-1} + \beta_3 D_{i,t} + \beta_4 X_{i,t-1} + a_i + \varepsilon_{i,t} + v_t \\ Drinking_{i,t} = \beta_0 + \beta_1 M_{i,t-1} + \beta_2 MI_{i,t-1} + \beta_3 D_{i,t} + \beta_4 X_{i,t-1} + a_i + u_{i,t} + v_t \\ Inactivity_{i,t} = \beta_0 + \beta_1 M_{i,t-1} + \beta_2 MI_{i,t-1} + \beta_3 D_{i,t} + \beta_4 X_{i,t-1} + a_i + z_{i,t} + v_t \end{cases} \quad (1)$$

where the subscript *i* represents individuals, and *t*, time. On the left hand side, the dependent variables $Smoking_{i,t}$, $Drinking_{i,t}, and\ Inactivity_{i,t}$ are dummy variables, which equal to one if individual *i* respectively smokes currently, drinks heavily, and is physically inactive, and zero otherwise. We consider two aspects of cognitive abilities, namely memory (*M*) and mental intactness (*MI*). $D_{i,t}$ is a vector of control variables at time *t* that affect the respondent's propensity to adopt risky behaviours, including gender, age, age squared, and a dummy indicating whether the respondent lives in a rural area. $X_{i,t-1}$ is a vector of control variables at time *t-1.* It includes the education level, income, marital status, number of chronic diseases, self-rated health status, social interactions, depression level, work type, wealth, and number of hospital visits in the past year. These lagged explanatory variables take into account the fact that it may take some time for each independent variable to affect the dependent variable. The definitions of all these variables are shown in Appendix 2.1. Provincial dummies are also included in the model to control for provincial-level effects. The error term is made up of three components in each equation. a_i takes into account the unobserved individual heterogeneity. v_t is the error term for a time-specific effect. $\varepsilon_{i,t}, u_{i,t}, and\ z_{i,t}$ are the time-varying error terms, and are assumed to share a multivariate normal distribution with a mean of 0 and a variance of 1.

For our Hypotheses 1, 2, and 3 to be satisfied, both 'β_1' and 'β_2' in the three components of Eq. 1 are expected to be significant and negative, implying that cognitive abilities in the previous wave are negatively associated with unhealthy behaviours, namely smoking, heavy drinking, and physical inactivity.

Males are expected to be positively related to smoking, drinking, and physical inactivity, because of Chinese culture (Chan and Leung, 2015; Chou, 2008; Wang *et al*., 2017). Also, we expect age have a positive association with these three unhealthy behaviours, meaning that the young generation behaves healthier than old generation, because the young people have open minds, and are more likely to adopt advices from physicians. Thus, the rate of risky behaviours is lower among young people (Chou, 2008; Cui *et al*., 2012). Following the previous literature in China (Chan and Leung, 2015; Chou, 2008; Strauss *et al*., 2010), SES is assumed to have a positive relationship with smoking, drinking, and physical inactivity, respectively. The reason is that people with high SES have better financial abilities, which make them affordable for cigarettes and alcohol consumptions. They can also afford more expensive transportation, such as driving a car, which highly reduce the chance of physical activity. Therefore, variables *income, above, primary,* and *wealth* are expected to show positively associated with risky behaviours, respectively. Marriage shows a protective effect on risky behaviours (Iparraguirre, 2015; Laaksonen *et al*., 2003). Compared with single people, we expect that variable *married* has a negatively relationship with all three unhealthy behaviours. However, *divorced* and *widowed* have positive relationships with these behaviours. It is possible that divorced people are in the pain of marriage failures and widowed people are in the pain of losing their spouses, which contribute to the urges for risky behaviours. Additionally, we expect that living in rural area is positively associated with risky behaviours, because urban people are more likely to experience the health promotion campaigns, while rural people have relatively lower social resources on these campaigns (Laaksonen *et al*., 2003; Palipudi *et al*., 2012).

Employment is expected to be positively associated with unhealthy behaviours (Chan and Leung, 2015; Ruhm, 2005), because employed people have long working hours which make people have no time to undertake exercises. Also, the employed people have more pressures from work, which increases their dependence on cigarettes and alcohol. Furthermore, poor health status is assumed to have a negative relationship with unhealthy behaviours (Iparraguirre, 2015), because people are more likely to accept the health-promoting advice from physicians and family members when their health status are not good. Another possible reason for the negative relationship between poor health status and risky behaviours is that the middle-aged and elderly people with poor health conditions tend to behave well in order to alleviate their existing disease conditions. Thus, we expect the variables, such as *slfhp*, *chronics*, and *hsptim1y*, have negative relationships with unhealthy behaviours, while *slfhvg* and *slfhg* show positive relationships with unhealthy behaviours. According to Cui *et al*. (2012), *CESD* is expected to be positively related to the three unhealthy behaviours. This can be explained by the fact that depressed people lack emotional management skills, and are more likely to undertake risky behaviours to ease their depression. Additionally, the number of social activates is assumed to have a negative association with three unhealthy behaviours (Gao *et al*., 2013;

Shankar *et al.*, 2011). The reason is that social people are more likely to receive the instrumental and emotional supports from their families and friends.

2.4.2 Fixed effects models

Given the availability of the number of cigarettes and units of alcohol, we test the associations between cognition and the number of cigarettes smoked per day, as well as the unit of alcohol drunk per week. The following equations are therefore estimated using a fixed effects model:

$$Smokenum_{i,t} = \beta_0 + \beta_1 M_{i,t-1} + \beta_2 MI_{i,t-1} + \beta_3 D_{i,t} + \beta_4 X_{i,t-1} + a_i + \varepsilon_{i,t} + v_t \qquad (2)$$

$$Drinknum_{i,t} = \beta_0 + \beta_1 M_{i,t-1} + \beta_2 MI_{i,t-1} + \beta_3 D_{i,t} + \beta_4 X_{i,t-1} + a_i + u_{i,t} + v_t \qquad (3)$$

The specification of the fixed effects model is similar to Eq. 1, except the dependent variables. In Eq. 2, the dependent variable $Smokenum_{i,t}$ indicates the number of cigarettes smoked per day. Also, the dependent variable $Drinknum_{i,t}$ in Eq. 3 represents the unit of alcohol consumed per week. In light of Hypotheses 1 and 2, we expect 'β_1' and 'β_2' in Eq. 2 and 3 to be significant and negative. In other words, we expect cognitive abilities in the previous wave to be negatively related with the number of cigarettes smoked and the number of units of alcohol drunk.

2.4.3 Simultaneous Probit models and fixed effects models among married respondents

In order to test the spousal effect, we add the spousal unhealthy behaviour status as an independent variable in the models. The specifications of Eq. 1 become Eq. 4. Similarly, Eq. 2 and 3 become Eq. 5 and 6, respectively. Only the subsample of respondents who are married is considered in this section.

$$\begin{cases} Smoking_{i,t} = \beta_0 + \beta_1 M_{i,t-1} + \beta_2 MI_{i,t-1} + \beta_3 D_{i,t} + \beta_4 X_{i,t-1} + \beta_5 SS_{i,t-1} + a_i + \varepsilon_{i,t} + v_t \\ Drinking_{i,t} = \beta_0 + \beta_1 M_{i,t-1} + \beta_2 MI_{i,t-1} + \beta_3 D_{i,t} + \beta_4 X_{i,t-1} + \beta_5 SD_{i,t-1} + a_i + u_{i,t} + v_t \\ Inactivity_{i,t} = \beta_0 + \beta_1 M_{i,t-1} + \beta_2 MI_{i,t-1} + \beta_3 D_{i,t} + \beta_4 X_{i,t-1} + \beta_5 SI_{i,t-1} + a_i + z_{i,t} + v_t \end{cases} \tag{4}$$

$$Smokenum_{i,t} = \beta_0 + \beta_1 M_{i,t-1} + \beta_2 MI_{i,t-1} + \beta_3 D_{i,t} + \beta_4 X_{i,t-1} + \beta_5 SS_{i,t-1} + a_i + \varepsilon_{i,t} + v_t \tag{5}$$

$$Drinknum_{i,t} = \beta_0 + \beta_1 M_{i,t-1} + \beta_2 MI_{i,t-1} + \beta_3 D_{i,t} + \beta_4 X_{i,t-1} + \beta_5 SD_{i,t-1} + a_i + u_{i,t} + v_t \tag{6}$$

In Eq. 4, 5 and 6, the variables $SS_{i,t-1}$, $SD_{i,t-1}$ and $SI_{i,t-1}$ denote whether the respondent's spouse is a current smoker, a heavy drinker, or physically inactive, respectively. Also, the vector of control variables $X_{i,t-1}$ no longer contains marital status in all of the three equations, because only married respondents are included in the sample. According to Hypothesis 4, we expect 'β_5' to be significant and positive in the three equations above, meaning that spousal risky behaviours affect the respondents' risky behaviours positively.

2.4.4 Estimation methodology

In Eq. 1 and Eq. 4, we use the simultaneous Probit model in the estimations, because our three left hand side variables are dichotomous (Roodman, 2011). This study does not use the independent models for the three unhealthy behaviours, considering that the estimated marginal effects may be inefficient when the correlations between the error terms in the three equations are neglected. However, the simultaneous equation model allows us to estimate the three unhealthy behaviours simultaneously, accounting for the potential correlations between smoking, heavy drinking, and physical inactivity. This model controls for the potential bias that may arise from correlations.

In the previous studies, Schneider and Schneider (2006) use the simultaneous Probit model to examine the determinants of health behaviours in Germany, namely smoking, physical inactivity, and having bad diet habits. We follow Schneider and Schneider's (2006) approach, and carry out the estimation by using Stata's Conditional Mixed-process (*cmp*) model. This model estimates a simulated maximum-likelihood three-equation Probit model using the Geweke-Hajivassiliour-Keane (GHK) algorithm (Roodman, 2011). Additionally, we use the cluster option at the personal level to control for the unobserved individual heterogeneity in the three equations.

Also, we use the multivariate Probit model for the robustness test. With the Stata program *mvprobit*, this model simulates maximum likelihood using the GHK simulator, which has no restriction on the number of model equations (Cappellari and Jenkins, 2003). The cluster option at the personal level is included. Moreover, the number of replications is set at 96 draws, which is slightly larger than the square root of the sample size. In this way, the simulated maximum likelihood estimator can provide a good estimate of the correlations between the error terms. For all the models mentioned above, we report and discuss the marginal effects in Tables 2.2, Table 2.4 and Table A in Appendix 2.2.

Additionally, the fixed effects Probit model is analysed, because the number of cigarettes smoked per day and the unit of alcohol drunk per week are available, and our two left hand side variables are continuous (Greene, 2011). In this study, the physical activity cannot be counted by a specific number, because physical exercises in the CHARLS include three categories at different intensities and the exercise time is in a range instead of the actual time spent. Therefore, we ignore the association between cognition and the time of physical activity. After running a Hausman test, the results show that the fixed effects model is preferred compared with the random effects model. The fixed effects model removes the effects of time-invariant characteristics, so the estimate is not biased because of the omitted time-invariant characteristics, such as the culture. As we use the panel data for the analyses, this fixed effects model accounts for the unobserved respondent-specific heterogeneity. We report and discuss coefficients in Tables 2.3 and Table 2.5.

2.5 Data and summary statistics

This section describes the dataset used in this chapter, and presents summary statistics.

2.5.1 The China Health and Retirement Longitudinal Survey

Our analysis uses data from the China Health and Retirement Longitudinal Survey (CHARLS). The CHARLS is a national representative, longitudinal survey of Chinese residents and their spouses aged 45 and older. It is based on the HRS in the United States, and relates to similar surveys such as the ELSA and the Survey of Health, Ageing and Retirement in Europe. This survey includes assessment of the demographic backgrounds, health circumstances, job conditions as well as financial circumstances of respondents (Zhao *et al*., 2012).

The baseline national wave of CHARLS was being fielded in 2011, and it randomly selects households from maps and listings within each rural or urban community. There are 34 provincial-level administrative units in China, but 6 provinces level units are excluded in the CHARLS, namely Hainan, Taiwan, Tibet, Ningxia, Hong Kong, and Macau. The baseline dataset contains 10,199 households and 17,705 individuals, covering 150 counties and 450 villages in 28 provinces in China. Follow-up interviews were conducted biannually. The second wave was conducted on 2013, and it contains

10,803 households and 18,605 individuals. Among these interviewees, 15,770 are follow-up interviewees, and 2,834 are new interviewees. 431 individuals from the baseline survey pass away during the period from 2011 to 2013. Finally, the third wave of the survey was being fielded in 2015, which added 574 new interviewees based on the 2013 wave of the CHARLS. More details on the CHARLS are available in Zhao *et al.* (2012).

In this study, we focus on the individuals who are interviewed in all three waves, because information on lifestyle and health behaviours are included in all three waves. The questions about smoking and heavy drinking are asked to all the respondents, but the questions about physical inactivity are only presented to a random subsample of households, which accounts for around half of the total respondents. In order to control for the potential effect of extreme values and faulty data, we drop observations below the 1st percentile and above the 99th percentile for each continuous main regression variable. This leaves us with a sample of 9,088 observations in the main estimation. The sample has a balanced structure over the three waves, which accounts for unobserved heterogeneity by each individual and reduces attrition bias by missing values.

2.5.2 Cognitive abilities and unhealthy behaviours

2.5.2.1 Cognitive abilities

Following Lei *et al*. (2012) and Lei and Liu (2018), we include two measures of cognitive abilities in this paper, namely memory and mental intactness[6]. The first domain estimates

---

[6] This part draws on my first year proposal (Fu, 2016)

episodic memory, which includes both immediate memory and delayed memory. In the CHARLS, a list consisting of 10 words is read by interviewers. Immediate memory is measured by the number of words that respondents can remember immediately. After being asked several other questions, they are asked to recall the words that the interviewer read five minutes earlier. The number of words the respondent recalls is referred to as delayed memory. Episodic memory is calculated as the average of immediate and delayed recalled results (McArdle *et al*., 2007).

The second domain estimates mental intactness, where 11 questions are asked to identify the respondent's mental status, including successively subtracting 7 from 100 (up to five times), awareness of the date, month, year (using either solar or lunar calendar), the day of the week and season of the year, and the ability to redraw a picture of two overlapping pentagons. We measure mental intactness through counting each correct answer. This yields a score ranging from 0 to 11.

2.5.2.2 Smoking

Smoking status is derived from the questions on current smoking habits that yields a dummy variable *smoke* coded 1 if the respondent currently smokes, and 0 otherwise. Also, the variable *smokenum* indicates the number of cigarettes smoked per day.

2.5.2.3 Heavy drinking

Heavy drinking is determined based on recommendations from the United States. Two institutes, namely the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and the Center for Disease Control and Prevention (CDC), publish the same drinking standard. For men who are 65 years or younger, heavy drinking is typically defined as consuming 15 drinks or more per week. For women who are 65 years or younger, it is defined as consuming 8 drinks or more per week. For people older than 65 years, heavy drinking is defined as consuming 8 drinks or more per week for both men and women. Each standard drink contains roughly 14 grams of pure alcohol. We are following this standard because it is suitable for the middle-aged and elderly people, which can be explained by the fact that the influence of alcohol consumption on health differs between those aged 65 or less and those older than 65 (Lin *et al*., 2011; Shaw and Agahi, 2012). Also, this standard is widely used in the current studies on alcohol consumption among middle-aged and elderly people (Li *et al*., 2019; Molander *et al*., 2010; Kirchner *et al*., 2007).

The CHARLS questionnaire asks respondents: 'Did you drink any alcoholic beverages, such as beer, wine, or liquor in the past year? (If 'yes') How often?' For those who answer 'Drink but less than once a month' or 'None of these', their questions end in terms of drinking section, and they are classified as non-heavy drinkers. For those who answer 'Drink more than once a month', two more questions are asked repeatedly based on three different types of alcoholic beverages, such as liquor, wine and beer. The first question is: 'How often did you drink this type of alcoholic beverage per month in the last year?' and the second question is: 'The last time you drank this type of alcoholic beverage last year, how many liang or bottle did you drink?' Liang is a Chinese unit of alcohol, and 1 liang equals to approximately 50ml of liquor. It is worth noting that the average amount of alcohol drinks in the last year is not available in this dataset. So, we use the amount of drinks in the last time to represent the average amount. Using this information, we firstly translate the unit of liang or bottle to drink. Then, we are able to calculate the number of drinks for every respondent per week through multiplying the weekly frequency by the number of alcohol drinks in the last time. Using this measure, if a middle-aged man drinks 15 or more drinks per week, he is defined as 'heavy drinking', and if a woman drinks 8 or more drinks per week, she is defined as 'heavy drinking'. Additionally, the variable *drinknum* represents the unit of alcohol consumed per week.

2.5.2.4 Physical inactivity

Physical inactivity is defined using the International Physical Activity Questionnaire (IPAQ), Which is widely used, especially in the large population studies (Fogelholm *et al*., 2016; Guthold *et al*., 2008; Haseli-Mashhadi *et al*., 2009). Although CHARLS has its own questions on physical activity, all the information needed to define physical inactivity based on the IPAQ standard is available in the survey. Therefore, IPAQ is compatible with CHARLS in terms of physical inactivity.

Specifically, CHARLS divides all kinds of physical activity into three groups, Group A, B and C. Group A includes vigorous activities, which refer to hard or high intensity physical effort, which make you breathe much harder than normal (such as heavy lifting, aerobics, fast bicycling, and cycling with a heavy load). Group B includes moderate-intensity activities, which refer to moderate physical effort that you do in the usual week, and make you breathe somewhat harder than normal (such as carrying light loads, bicycling at a regular pace, and mopping the floor). Group C focuses on a specific activity, which is walking. It includes walking at work, at home, walking to travel from place to place, and any other waking that you might do solely for recreation, sport, exercise, or leisure. For every group, respondents are asked the question 'during a usual week, did you do any activity (Group A or B or C) for at least 10 minutes continuously?' Respondents are not given further question of this group if their answers are 'No' to this question, alternatively, the respondents are asked to answer other questions from the other two groups. If the answer is 'yes', follow-up questions are: 'During a usual week, on how many days did you do activity (Group A or B or C) for at least 10 minutes?' and 'How much time did you usually spend doing activity (Group A or B or C) on one of those days?' Therefore, we know the activity days in every week and time spent exercising each time in each group.

In the classification of IPAQ, there are three categories of physical activity based on total energy consumed every week. The first category is high level. The second category is moderate level. If the respondent does not meet criteria for either high or moderate level, he/she belongs to the low level, which is the third category. Therefore, the dummy variable *inact* equals to 1 if the respondent belongs to the low level of physical activity group, and 0 if the respondent belongs to the moderate or high level of physical activity groups.

IPAQ combines the information in Group A, B and C, and lists the criteria for moderate and high levels of physical activity. For example, if the respondent has three or more days of vigorous activity of at least 20 minutes per day, he/she is considered in the moderate level. If the respondent has 5 or more days of moderate-intensity activity or walking of at least 30 minutes each day, he/she belongs to the moderate level. Additionally, IPAQ uses the metabolic equivalent of task (MET) in their criteria, which represents the energy expenditure during an activity. IPAQ also gives the researchers the MET levels for different groups, which are 8.0 MET for Group A, 4.0 MET for Group B, and 3.3 MET for Group C, separately. We use the continuous score that is calculated by multiplying MET levels, minutes of activity, and frequency per week, which is an indicator for energy expenditure during an activity. If the respondent has three or more days of vigorous activity per week and achieves 1500 MET-minutes per week, he/she is defined in the high level.

2.5.3 Summary statistics

Table 2.1 presents the descriptive statistics for main sample used in the simultaneous Probit model. This table presents the number of observations, mean value, standard deviation, minimum value, median value, and maximum value for each variable.[7] 27.6 percent of respondents are current smokers. Around one out of ten respondents are heavy alcohol drinkers according to the NIAAA and CDC standard. Moreover, around one out of five respondents are classified as showing a low level of physical activity. The average number of cigarettes consumed per day is 3.707, and the average alcohol drinks consumed

[7] See Appendix 2.1 for detailed definitions of all variables used in this study.

per week is 4.541 per week. Only 8,199 and 8,986 observations are reported for the variables *smokenum* and *drinknum*, respectively, because of a small amount of missing values on the observed numbers of cigarettes and alcohol drinks.

Education is divided into three categories, including illiteracy, primary school level, and secondary school level or above. Firstly, 25.4 percent of the respondents are illiterate. And around 40 percent of the respondents have a primary school-equivalent level of education. Thirdly, 33.9 percent of the respondents obtain at least a secondary school level of education. This prevalence of education is in line with data on the rates of education from other studies on China (Fend *et al*., 2013; Ning *et al*., 2016). Also, the average annual income per capita is 10,679 YUAN, and the average wealth per capita is 50,534 YUAN. 47.4 percent of the respondents are male, and the average age is 61.4. 88.8 percent of the respondents are married, while the probability of single, divorced, and widowed respondents are 0.682, 0.781, and 9.38 percent. 63.4 percent of the respondents live in a rural area. The average number of chronic diseases is 1.362. We include four categories of self-rated health status. Only around 10 percent of respondents say their health is very good, 16.2 percent say it is good, around half of the respondents rate it as fair, and 19.4 percent of respondents say their health condition is poor. The average number of social activities is 0.551.

CHARLS includes a short version of the Center for Epidemiological Studies Depression Scale (CESD), and the mean value is 8.028. There are four categories of work status. 28.5 percent of the respondents do not work, 52.5 percent of the respondents are engaging in agricultural work, 11.9 percent of the respondents are engaging in employed work, and only 7.03 percent of the respondents are engaging in self-employed work. In the sample, the average number of times the respondents received inpatient care during the past year is 0.146. Additionally, the mean value of memory is 3.617, and the mean value of mental intactness is 7.65.

2.6 Empirical findings

This section presents the empirical results aimed at identifying the relationship between cognition and unhealthy behaviours among the middle-aged and elderly respondents. As we use the simultaneous Probit models in the main analysis, marginal effects are presented and discussed in Tables 2.2, Table 2.4 and Table A in Appendix 2.2. The fixed effects Probit model is analysed in Tables 2.3 and Table 2.5, and corresponding coefficients are presented.

2.6.1 Baseline specifications

Table 2.2 presents the results of the simultaneous Probit model in Eq. 1 for the probability of smoking (column 1), heavy drinking (column 2), and physical inactivity (column 3). The estimates were obtained with the Stata *cmp* command. Next, this section discusses the results in detail.

Focusing on column 1, mental intactness is negatively related to the probability of smoking. The marginal effect suggests that, an increase in mental intactness by one unit, is associated with a 0.5 percentage point lower probability of smoking, keeping all other control variables fixed. However, memory is not significantly associated with the prevalence of smoking. This can be explained considering that memory and mental intactness represent different aspects of cognitive abilities, and only one aspect of them is significantly associated with smoking. Respondents with higher mental intactness are receptive to new information and may benefit from health promotion campaigns. Therefore, they are more likely to have adequate health literacy, and are less likely to smoke. The negative relationship between cognitive abilities and smoking is consistent with our first hypothesis and with the literature (Anstey *et al*., 2009; Deary *et al*., 2000; Deary, 2014; Weiser *et al*., 2010; Kubička *et al*., 2001). The results suggest that individuals with higher cognitive abilities, and particularly with a higher mental intactness score, are less likely to be current smokers.

In column 2 of Table 2.2, memory shows a significantly negative association with the probability of drinking, while mental intactness shows an insignificant association with it. The marginal effect suggests that a one-unit higher memory score is related to a 0.5 percentage point lower chance of drinking heavily, keeping all other variables fixed. Memory is shown to have a significantly negative association with drinking, and the possible reason is that respondents with good memory are more likely to remember the harmful effects of excessive alcohol use, and therefore have a good self-control on heavy drinking. Consistent with our second hypothesis, individuals with higher cognitive abilities, particularly a higher memory score, are less prone to be heavy drinker.

In column 3, the association between cognition and physical inactivity is similar to that in column 2. Specifically, memory is significantly negatively associated with the probability of physical inactivity, but mental intactness is insignificantly associated with it. A one-unit higher memory score is associated with a 0.8 percentage point lower probability of being physically inactive, keeping other control variables constant. This result is consistent with the third hypothesis and with previous literature (Anstey *et al*., 2009), which proves that individuals with higher cognitive abilities, especially a higher memory ability, are less likely to be physically inactive.

The magnitudes of cognitive abilities in columns 1, 2 and 3 seem quite small. The corresponding magnitudes in the previous study are between 0.009 and 0.033, which focus on the 40s and 60s in Australia (Anstey *et al*., 2009). The magnitudes of cognitive abilities in our study are a little smaller than that in Anstey *et al*. (2009). One possible reason is that our study covers a great number of elderly people who are over 70, and these people have relatively lower cognitive abilities than the middle-aged people. Specifically, 1,703 out of 9,088 respondents in the sample are 70 years or over. The mean values of memory and mental intactness among the elderly people who are 70 or over are 2.970 and 6.726, respectively. By contrast, the corresponding values among the middle-aged people are higher at 3.767 and 7.863, respectively. The relatively lower cognitive abilities of the elderly people in our sample may contribute to the smaller magnitudes in our results.

Consistent with our expectations and the previous studies (Wang *et al*., 2017; Iparraguirre, 2015; Chan and Leung, 2015; Cui *et al*., 2012; Shankar *et al*., 2011; Chou, 2008; Laaksonen *et al*., 2003), the results show that, in column 1, older widowed males with higher CESD scores, and living in rural areas, are more likely to be smokers. By contrast, respondents with more chronic diseases and more hospital visits in the last year are less likely to be smokers. Similarly, in column 2, older males, who are currently working in agriculture, living in rural areas, and having good or excellent self-rated health status, are more likely to be heavy drinkers. By contrast, respondents with more chronic diseases, poorer self-reported health status, and more hospital visits, are less likely to be heavy drinkers. In column 3, the respondents who living in rural areas and having fewer social activities are more likely to be physically inactive.

However, several findings are inconsistent with our expectations. For example, in column 1, the number of social activities shows a positive relationship with smoking. This can be explained considering that as a part of Chinese culture, social networking requests people to smoke in most circumstances. Also, *wealth* at time *t-1* is significantly negatively associated with smoking. This may be caused by the fact that affluent people can afford smoking cessation interventions, and quit smoking successfully, while poor people are less likely to pay money for these services and, thus, have a relatively low rate of successful smoking cessation (Palipudi *et al*., 2012; Schneider and Schneider, 2006). Additionally, we could not find significant relationships between smoking and several determinants, including education, income, marital status, self-rated health status, and employment. The corresponding significant associations in previous studies (Chan and Leung, 2015; Strauss *et al*., 2010; Chou, 2008; Ruhm, 2005; Laaksonen *et al*., 2003) may be caused by not including cognitive abilities in the estimations (omitted variable bias), as well as by using different populations, because this study is looking at the middle-aged and elderly people while the previous studies are looking at all the adults. For the same reason, in column 2, determinants, such as education, income, wealth, marital status, social activity, and CESD, have no significant association with the probability of heavy drinking.

Additionally, in column 3, respondents with at least a secondary school level of education are less likely to be physically inactive. This finding can be explained considering that people with higher educational level have better health knowledge, which improves their understanding in the negative impacts of physical inactivity (Schneider and Schneider, 2006). Therefore, highly educated respondents are less likely to be physically inactive. Also, both *income* and *wealth* are negatively associated with the prevalence of physical inactivity. Considering that pollution prevents people in China from exercising outdoors, compared with poor people, rich people can afford the extra costs of exercises, such as gym membership fees, badminton court fees and so on. The variable *age* has a significantly negative association with inactivity. This result can be explained considering that the older respondents are more flexible with time management, because the majority of them are retired after 60. Therefore, older people are more likely to devote their time to physical exercises and less likely to be physically inactive. Moreover, poor self-reported health status and chronic diseases contribute to physical inactivity, because people with these conditions may not be able to do moderate or higher intensity exercises. Compared with the respondents who do no work, working respondents are less likely to be physically inactive. One possible reason is that they have higher opportunity costs for illness, which motivate them to exercise (Schneider and Schneider, 2006). Another is that their jobs may require lots of physical activities.

At the bottom of the Table 2.2, the */atanhrho_ij* statistics are shown, indicating the correlation between the error terms of equation *i* and equation *j* (Roodman, 2011). */atanhrho_ij* is the transformed version of rho, because rho is a bounded scale. The transformed rho changes the value into an unbounded scale by using the arc-hyperbolic tangent function. The coefficients show that the correlations between the error terms of smoking and heavy drinking, as well as between heavy drinking and physical inactivity, are statistically significant. These results show interdependence across the equations. Without controlling for these possible correlations, the results reported may be biased and inconsistent.

The previous results were estimated using the Stata command *cmp*. To confirm the robustness of the main results, an alternative model, which is the multivariate Probit model obtained using the Stata command *mvprobit*, is estimated. The specification of this model is also Eq. 1. The results are presented in Table A in Appendix 2.2. They are consistent with those reported in Table 2.2. The only difference is from the age squared variable. *age2* is significantly negative in column 1 and is significantly positive in column 3 of Table A in Appendix 2.2. However, *age2* is zero in all the columns in Table 2.2. Consistent with the existing studies (Lawrence *et al*., 2014; García *et al*., 2011; Pabayo *et al*., 2011; Skowron *et al*., 2008; Miles, 2006), age has non-linear relationships with smoking and physical inactivity. Specifically, there is an inverted U-shaped relationship between age and smoking. The marginal effects in column 1 of Table A in Appendix 2.2 suggest that the distribution of the probability of smoking peaks at the age of 43. The current study only focuses on people who are 45 or older. Therefore, there is a negative relationship between age and the probability of smoking in this sample. This finding can be explained considering that older respondents are likely to stop smoking because of health problems, thus, the probability of smoking tends to reduce with age (Lee *et al*., 2016; Choi and DiNitto, 2014; Tsai *et al*., 2012). Similarly, the marginal effects in column 3 of Table A in Appendix 2.2 suggest that there is a negative relationship between age and the probability of physical inactivity before the age of 49.5, and a positive relationship thereafter. This finding is possible because people' health status deteriorates with age, and older people with poor health condition may not be able to do moderate or high-intensity exercises.

2.6.2 Fixed effects models

Table 2.3 presents the estimates of the fixed effects models aimed to assess whether cognition is associated with the number of cigarettes smoked (column 1) and the number of units of alcohol consumed (column 2) among the middle-aged and elderly respondents in China. This table corresponds to the estimates of Eq. 2 and 3, respectively.

Contrary to the results in Table 2.2, the coefficients associated with both memory and mental intactness at time *t-1* are insignificant in columns 1 and 2. The results suggest that, keeping all the other control variables fixed, cognitive abilities do not affect the number of cigarettes smoked nor the units of alcohol consumed. In other words, cognition only affects the decision of undertaking unhealthy behaviours, such as smoking and heavy drinking, but it does not affect the number of cigarettes and units of alcohol consumed. This can be explained considering that people with adequate health literacy and a higher level of self-control can protect themselves from the attraction of cigarettes and alcohol.

Yet, for those who have already started smoking and drinking heavily, the impact of cognition is limited. These findings about cognitive abilities in Table 2.3 are inconsistent with our Hypotheses 1 and 2.

Consistent with the results in Table 2.2, the divorced, widowed status, and the number of social activities are significantly positively associated with the number of cigarettes smoked (column 1). However, the significant effect of all the other control variables in column 1 disappears. Moreover, in column 2, all the independent variables show no significant associations with the number of alcohol units consumed per week. These findings are different from our expectations. Additionally, the effects of gender and whether the respondent lives in a rural area are not reported in the Table 2.3 because these variables are time-invariant.

2.6.3 Models among married respondents

This section includes two tables aimed at investigating the spousal effects on the unhealthy behaviours. Table 2.4 shows the estimates of the simultaneous Probit models among married respondents, where spousal smoking, drinking, and physical inactivity status are included (Eq. 4). Column 1 presents the results for smoking; column 2, for heavy drinking; and column 3 for physical inactivity. Table 2.5 presents the fixed effects estimates of Eq. 5 and 6 for the number of cigarettes smoked (column 1) and the number of units of alcohol drunks (column 2).

Considering the spousal effects in Table 2.4, we can see that the marginal effects associated with the spousal smoking, heavy drinking, and physical inactivity are positive and significant in all three columns. The results suggest that if the respondent's spouse was a smoker in the previous waves, the respondent has a 6.1 percentage points higher probability of smoking in the current wave. Similarly, the spousal heavy drinking status in the previous wave is associated with a 4.2 percentage points higher probability of heavy drinking, and the spousal physical inactivity status in the previous wave is associated with a 3.6 percentage points higher probability of physical inactivity by the respondent. These findings are consistent with Hypothesis 4 and the literature (Li *et al*., 2013; Stimpson *et al*., 2006; Jurj *et al*., 2006).

Turning to the cognitive abilities, in these new models, lagged mental intactness is significantly negatively related to smoking, and lagged memory is significantly negatively related to heavy drinking. These results are generally consistent with those reported in Table 2.2. However, memory has an insignificant association with physical inactivity, which is different from the results in Table 2.2. A possible reason is that the inclusion of spousal physical inactivity in the model attenuates the relationship between cognition and the probability of physical inactivity.

In column 3, the coefficient associated with the *primary* dummy is significantly positive, which is consistent with our expectation and the literature (Chan and Leung, 2015; Chou, 2008). Different with our expectation, *male* shows a negative association with physical inactivity. This can be explained considering that married men take the majority of the physical labour, compared with their wives (Li *et al*., 2013). This is because men are stronger and can undertake harder and higher intensity physical works. Another possible reason is that in the Chinese culture, males typically do physical works in a family, such as heavy lifting and gardening.

As for the other differences between Table 2.2 and Table 2.4, the significance of the marginal effects associated with some of the control variables disappears. Examples of these variables are *rural* in column 1, *age* and *rural* in column 2, and *above*, *chronics*, *social*, *workemp*, *workself*, and *wealth* in column 3. A possible explanation is that married respondents are a subsample of the population, and the influences of these variables are attenuated by spousal effects.

After including the spousal smoking and drinking status in the columns 1 and 2 of Table 2.5, respectively, the results uncover a significantly positive relationship between the spouse's smoking status and the number of cigarettes consumed by the respondent. Specifically, a respondent whose spouse was a smoker in the previous wave is likely to smoke 0.47 more cigarette per day than his/her counterpart whose spouse was not a smoker. This finding supports our Hypothesis 4 and is in line with the results we obtained for the probability of smoking among married respondents in Table 2.4. However, no significant association is found between spousal heavy drinking status and the number of drinks consumed. This can be explained considering that heavy drinking always happens between friends and colleagues, when the spouse is not present, therefore, the spousal effect is limited. Consistent with the results in Table 2.3, the coefficient associated with the variable *social* is significantly positively associated with the number of cigarettes smoked (column 1).

2.7 Conclusion

Unhealthy behaviours among the Chinese ageing population have recently become important health concerns for researchers. However, to the best of our knowledge, the relationship between cognitive abilities and risky behaviours has not been investigated in the Chinese context[8]. This chapter fills this gap in the literature by examining the association between cognitive abilities and risky behaviours, namely smoking, drinking, and physical inactivity, among the middle-aged and elderly people in China. We use the 2011, 2013, and 2015 waves of the CHARLS dataset, and find that cognitive abilities are generally negatively associated with the probability of smoking, drinking, and physical inactivity, respectively. Specifically, higher mental intactness is associated with a lower change of smoking, and higher memory is associated with a lower prevalence of heavy drinking and physical inactivity. Two methods of estimation are used, including the simultaneous Probit model and multivariate Probit model. However, from the results of the fixed effects models, we do not find the significant associations between cognition and the number of cigarettes smoked per day, as well as the number of alcohol drinks consumed per week. Among the married respondents, the probabilities of smoking, heavy drinking, and physical inactivity are significantly positively influenced by that of their spouses'. Also, the respondents are likely to smoke more cigarettes per day if their spouses were smokers in the previous wave, but the corresponding spousal effect has not been found for the number of alcohol drunk.

[8] This part draws on my first year proposal (Fu, 2016)

A limitation to this study is that diet quality and drug use are not included in the analysis. This is caused by the fact that these information are not available in the CHARLS dataset. Also, when calculating the amount of alcohol drinks per week, we use the amount of drinks in the last time to represent the average amount. This may cause a bias with the real values.

Moreover, there are potential reverse causality issues from unhealthy behaviours to cognitive abilities. For example, smoking may be associated with cognitive impairment in midlife. Smokers who survive into later life may be at risk of clinically significant cognitive declines, because smoking increases lung damage, atherosclerosis, and oxidative stress, which may lead to related diseases, such as chronic obstructive pulmonary disease, stroke, brain atrophy, and dementia. As a result, smoking is associated with cognitive impairment for older adults (Sabia *et al*., 2012; Collins *et al*., 2009; Huang *et al*., 2009). The current study tries to reduce this endogeneity problem by using lagged time-variant explanatory variables. Therefore, the lagged cognitive abilities are likely to be strongly associated with current unhealthy behaviours, but cannot be affected by current unhealthy behaviours. The Instrumental Variable (IV) method would be a good way to solve for the potential reverse causality problem, but this study does not use this method, because possible IVs, such as the respondent's cognitive abilities in childhood and the parents' cognitive abilities, are not available in the CHARLS. This study also uses the fixed effects model and includes more control variables in the regression, which may mitigate the unobserved heterogeneity issue.

As for policy implication, we suggest that education on cognitive abilities to middle-aged and elderly people in China may be effective in reducing the prevalence of unhealthy behaviours. Health promotion campaigns may be beneficial. Also, this study proves that risky behaviours are prevalent among the people with lower cognitive abilities, so public health advice could pay more attentions to this group, and make recommendations that are applicable for them.

In further studies, we plan to test the influence of peer effect on risky behaviours in China. Also, we will compare the determinants of unhealthy behaviours in China with that in developed countries.

Table 2.1 Descriptive statistics

variable	mean	S.D.	min	p50	max	N
smoke	0.276	0.447	0	0	1	9088
drink	0.115	0.319	0	0	1	9088
inact	0.215	0.411	0	0	1	9088
smokenum	3.707	8.342	0	0	40	8199
drinknum	4.541	13.46	0	0	157	8986
male	0.474	0.499	0	0	1	9088
age	61.40	8.832	45	61	101	9088
rural	0.634	0.482	0	1	1	9088
above	0.339	0.474	0	0	1	9088
primary	0.407	0.491	0	0	1	9088
illiterate	0.254	0.435	0	0	1	9088
rincome	10679	17378	0	5579	438240	9088
single	0.00682	0.0823	0	0	1	9088
married	0.888	0.315	0	1	1	9088
divorced	0.00781	0.0880	0	0	1	9088
widowed	0.0938	0.291	0	0	1	9088
chronics	1.362	1.391	0	1	9	9088
slfhvg	0.0924	0.290	0	0	1	9088
slfhg	0.162	0.368	0	0	1	9088
slfhf	0.551	0.497	0	1	1	9088
slfhp	0.194	0.396	0	0	1	9088
social	0.551	0.792	0	0	6	9088
CESD	8.028	5.971	0	7	30	9088
M	3.617	1.687	0	3.500	10	9088
MI	7.650	3.020	0	8	11	9088
noworking	0.285	0.452	0	0	1	9088
workagri	0.525	0.499	0	1	1	9088
workemp	0.119	0.324	0	0	1	9088
workslf	0.0703	0.256	0	0	1	9088
hsptim1y	0.146	0.453	0	0	3	9088
rwealth	50534	95535	0	19177	2051900	9088

Note: *S.D.* stands for the standard deviations. *P50* represents the median value. *N* stands for the number of observations. Income (*rincome*) and wealth (*rwealth*) are expressed in YUAN. Precise definitions of all variables can be found in Appendix 2.1.

Table 2.2 Simultaneous Probit models

	(1) smoking	(2) drinking	(3) inactivity
$male_t$	0.432***	0.232***	-0.004
	(59.08)	(25.04)	(-0.44)
age_t	0.020***	0.008*	-0.027***
	(3.12)	(1.65)	(-4.76)
$age2_t$	-0.000***	-0.000	0.000***
	(-3.43)	(-1.64)	(5.28)
$rural_t$	0.024**	0.014*	0.022**
	(2.21)	(1.82)	(2.02)
$above_{t-1}$	-0.014	-0.015	-0.032**
	(-0.85)	(-1.22)	(-2.08)
$primary_{t-1}$	0.012	-0.001	0.003
	(0.90)	(-0.09)	(0.24)
$income_{t-1}$	-0.003	-0.000	-0.004**
	(-1.64)	(-0.03)	(-2.07)
$married_{t-1}$	0.017	-0.012	-0.012
	(0.48)	(-0.45)	(-0.31)
$divorced_{t-1}$	0.080	-0.008	0.017
	(1.41)	(-0.19)	(0.28)
$widowed_{t-1}$	0.090**	-0.025	-0.006
	(2.38)	(-0.84)	(-0.13)
$chronics_{t-1}$	-0.015***	-0.009***	0.007**
	(-4.20)	(-3.44)	(2.15)
$slfhvg_{t-1}$	-0.002	0.027***	-0.000
	(-0.16)	(2.69)	(-0.02)
$slfhg_{t-1}$	-0.007	0.014*	-0.004
	(-0.65)	(1.64)	(-0.34)
$slfhp_{t-1}$	-0.009	-0.033***	0.024**
	(-0.73)	(-3.28)	(2.02)
$social_{t-1}$	0.011**	-0.002	-0.013**
	(2.11)	(-0.55)	(-2.23)
$CESD_{t-1}$	0.003***	0.000	-0.001
	(3.46)	(0.37)	(-1.05)
M_{t-1}	0.001	-0.005**	-0.008***
	(0.42)	(-2.30)	(-2.69)
MI_{t-1}	-0.005**	-0.000	-0.003
	(-2.53)	(-0.12)	(-1.45)
$workagri_{t-1}$	0.003	0.022**	-0.121***
	(0.25)	(2.29)	(-10.86)
$workemp_{t-1}$	-0.016	0.018	-0.042***
	(-1.05)	(1.47)	(-2.62)
$workslf_{t-1}$	-0.027	0.009	-0.032*
	(-1.51)	(0.61)	(-1.69)
$hsptim1y_{t-1}$	-0.041***	-0.032***	0.012
	(-4.04)	(-3.86)	(1.28)
$wealth_{t-1}$	-0.006**	0.002	-0.006***
	(-2.39)	(0.98)	(-2.63)

/atanhrho_12	0.215***		
/atanhrho_13	-0.010		
/atanhrho_23	-0.076**		
N	9088	9088	9088
r2_p			

Note: Columns 1 to 3 report the estimates of the probability of smoking, heavy drinking, and physical inactivity, respectively. Marginal effects are reported, with *z* statistics in parentheses. *N* denotes the number of observations. Illiterate, single, fair self-reported health status and unemployed are omitted categories. Precise definitions of all variables can be found in Appendix 2.1. Income and wealth are expressed in logarithms. Province and year dummies are included, but their marginal effects are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Table 2.3 Fixed effects models

	(1) smokenum	(2) drinknum
$male_t$	0.000	0.000
	(.)	(.)
age_t	0.009	-0.204
	(0.06)	(-0.51)
$age2_t$	-0.000	0.002
	(-0.24)	(0.60)
$rural_t$	0.000	0.000
	(.)	(.)
$above_{t-1}$	0.434	-0.975
	(1.06)	(-0.87)
$primary_{t-1}$	0.041	0.352
	(0.15)	(0.46)
$income_{t-1}$	0.031	0.024
	(1.63)	(0.47)
$married_{t-1}$	1.080	0.600
	(1.52)	(0.32)
$divorced_{t-1}$	1.688*	-0.562
	(1.73)	(-0.23)
$widowed_{t-1}$	1.269*	0.609
	(1.66)	(0.30)
$chronics_{t-1}$	0.028	-0.088
	(0.01)	(-0.01)
$slfhvg_{t-1}$	-0.161	0.163
	(-1.13)	(0.42)
$slfhg_{t-1}$	0.063	-0.042
	(0.60)	(-0.15)
$slfhp_{t-1}$	0.060	0.129
	(0.57)	(0.43)
$social_{t-1}$	0.106**	-0.015
	(1.98)	(-0.10)
$CESD_{t-1}$	0.006	-0.010
	(0.69)	(-0.43)
M_{t-1}	-0.027	-0.025
	(-1.05)	(-0.34)
MI_{t-1}	0.005	-0.003
	(0.24)	(-0.05)
$workagri_{t-1}$	0.069	-0.069
	(0.56)	(-0.20)
$workemp_{t-1}$	-0.132	-0.657
	(-0.81)	(-1.49)
$workslf_{t-1}$	-0.231	-0.747
	(-1.23)	(-1.45)
$hsptim1y_{t-1}$	0.123	-0.101
	(1.43)	(-0.42)
$wealth_{t-1}$	0.015	-0.022
	(0.60)	(-0.33)
$\_cons$	2.357	10.344

	(0.39)	(0.59)
N	18636	21357
r2_p		

Note: Columns 1 and 2 report the estimates of the number of cigarettes and alcohol, respectively. *t* statistics in parentheses. *N* denotes the number of observations. Illiterate, single, fair self-reported health status and unemployed are omitted categories. Precise definitions of all variables can be found in Appendix 2.1. Income and wealth are expressed in logarithms. Province and year dummies are included, but their coefficients are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Table 2.4 Simultaneous Probit models among married respondents

	(1) smoking	(2) drinking	(3) inactivity
$male_t$	0.482***	0.265***	-0.029**
	(39.08)	(20.59)	(-2.02)
age_t	0.019**	0.006	-0.032***
	(2.47)	(0.93)	(-3.50)
$age2_t$	-0.000***	-0.000	0.000***
	(-2.72)	(-0.91)	(3.78)
$rural_t$	0.019	0.013	0.046***
	(1.56)	(1.40)	(2.81)
$smokesp_{t-1}$	0.061***		
	(3.83)		
$drinksp_{t-1}$		0.042***	
		(2.58)	
$inactsp_{t-1}$			0.036**
			(2.32)
$above_{t-1}$	-0.004	-0.014	-0.007
	(-0.19)	(-0.98)	(-0.32)
$primary_{t-1}$	0.016	-0.001	0.031*
	(1.00)	(-0.04)	(1.70)
$income_{t-1}$	-0.003	0.001	-0.009***
	(-1.22)	(0.38)	(-2.83)
$chronics_{t-1}$	-0.018***	-0.009***	0.006
	(-4.43)	(-2.79)	(1.24)
$slfhvg_{t-1}$	-0.004	0.032***	-0.007
	(-0.23)	(2.74)	(-0.29)
$slfhg_{t-1}$	-0.009	0.016*	-0.004
	(-0.73)	(1.73)	(-0.25)
$slfhp_{t-1}$	-0.001	-0.041***	0.018
	(-0.08)	(-3.41)	(1.06)
$social_{t-1}$	0.011*	-0.003	-0.005
	(1.82)	(-0.54)	(-0.55)
$CESD_{t-1}$	0.004***	0.000	-0.000
	(3.76)	(0.24)	(-0.07)
M_{t-1}	0.001	-0.004*	-0.005
	(0.16)	(-1.78)	(-1.21)
MI_{t-1}	-0.006***	-0.001	-0.002
	(-2.71)	(-0.36)	(-0.74)
$workagri_{t-1}$	0.012	0.034***	-0.141***
	(0.85)	(3.02)	(-8.42)
$workemp_{t-1}$	-0.011	0.023	-0.021
	(-0.63)	(1.63)	(-0.90)
$workslf_{t-1}$	-0.012	0.020	-0.017
	(-0.60)	(1.23)	(-0.62)
$hsptim1y_{t-1}$	-0.048***	-0.033***	0.012
	(-4.09)	(-3.49)	(0.86)
$wealth_{t-1}$	-0.003	0.003	-0.006
	(-1.07)	(1.29)	(-1.46)

```
/atanhrho_12      0.238***
/atanhrho_13     -0.001
/atanhrho_23     -0.089*
--------------------------------------------------
N                4112         4112         4112
r2_p
--------------------------------------------------
```

Note: Columns 1 to 3 report the estimates of the probability of smoking, heavy drinking, and physical inactivity among married respondents, respectively. Marginal effects are reported, with *z* statistics in parentheses. *N* denotes the number of observations. Illiterate, single, fair self-reported health status and unemployed are omitted categories. Precise definitions of all variables can be found in Appendix 2.1. Income and wealth are expressed in logarithms. Province and year dummies are included, but their marginal effects are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Table 2.5 Fixed effects models among married respondents

	(1) smokenum	(2) drinknum
$male_t$	0.000	0.000
	(.)	(.)
age_t	0.143	-0.237
	(0.77)	(-0.47)
$age2_t$	-0.001	0.002
	(-0.89)	(0.45)
$rural_t$	0.000	0.000
	(.)	(.)
$smokesp_{t-1}$	0.470*	
	(1.83)	
$drinksp_{t-1}$		-0.180
		(-0.37)
$above_{t-1}$	0.475	-1.446
	(1.02)	(-1.14)
$primary_{t-1}$	0.083	0.696
	(0.27)	(0.80)
$income_{t-1}$	0.016	0.045
	(0.72)	(0.74)
$chronics_{t-1}$	0.013	-0.057
	(0.00)	(-0.01)
$slfhvg_{t-1}$	-0.157	0.203
	(-0.97)	(0.46)
$slfhg_{t-1}$	0.090	-0.077
	(0.74)	(-0.24)
$slfhp_{t-1}$	0.103	-0.028
	(0.84)	(-0.08)
$social_{t-1}$	0.126**	-0.136
	(2.01)	(-0.81)
$CESD_{t-1}$	0.000	-0.020
	(0.03)	(-0.78)
M_{t-1}	-0.020	-0.061
	(-0.67)	(-0.74)
MI_{t-1}	0.008	-0.008
	(0.37)	(-0.13)
$workagri_{t-1}$	0.086	0.133
	(0.61)	(0.33)
$workemp_{t-1}$	-0.213	-0.469
	(-1.14)	(-0.93)
$workslf_{t-1}$	-0.314	-0.455
	(-1.45)	(-0.77)
$hsptim1y_{t-1}$	0.123	-0.091
	(1.24)	(-0.33)
$wealth_{t-1}$	0.020	-0.022
	(0.66)	(-0.26)
_cons	-0.853	13.309
	(-0.12)	(0.66)

N	15083	17714
r2_p		

Note: Columns 1 and 2 report the estimates of the number of cigarettes and alcohol among married respondents, respectively. *t* statistics in parentheses. *N* denotes the number of observations. Illiterate, single, fair self-reported health status and unemployed are omitted categories. Precise definitions of all variables can be found in Appendix 2.1. Income and wealth are expressed in logarithms. Province and year dummies are included, but their coefficients are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Appendices

This section is laid out as follows. Appendix 2.1 describes the definitions of variables that are used in this chapter. Appendices 2.2 reports the empirical results of the multivariable Probit model for the probability of smoking, heavy drinking, and physical inactivity.

Appendix 2.1 Definitions of all variables

Variables	Definitions
smoke	Dummy variable equal to 1 if the respondent is a current smoker, and 0 otherwise
drink	Dummy variable equal to 1 if the respondent drinks heavily according to the NIAAA and CDC standard, and 0 otherwise
inact	Dummy variable equal to 1 if the respondent does not meet the moderate or high physical activity standard, and 0 otherwise
M	Continuous variable averaging the immediate memory and delayed memory
MI	Continuous variable summing the correct answers from 11 questions. These questions include successively subtracting 7 from 100 (up to five times), awareness of the date, month, year (using either solar or lunar calendar), the day of the week and season of the year, and the ability to redraw a picture of two overlapping pentagons[9]
rincome	Continuous variable representing the annual per capita income of the household. This includes the main respondent's and spouse's wage income, fringe benefits, unpaid family business income, pension insurance income, net agricultural

[9] This part draws on my first year proposal (Fu, 2016)

	income, net self-employed business income, household public transfer income, household capital asset income, and other household members' income
income	Continuous variable representing the natural logarithm of rincome
illiterate	Dummy variable equal to 1 if the respondent is illiterate, and 0 otherwise.
primary	Dummy variable equal to 1 if the respondent has attained primary school level, and 0 otherwise
above	Dummy variable equal to 1 if the respondent has attained secondary school or above level, and 0 otherwise
male	Dummy variable equal to 1 if the respondent is male, and 0 otherwise
age	Continuous variable representing the respondent's age at the time of the interview
age2	Age squared
married	Dummy variable equal to 1 if the respondent is married, and 0 otherwise
divorced	Dummy variable equal to 1 if the respondent is divorced, and 0 otherwise
widowed	Dummy variable equal to 1 if the respondent is widowed, and 0 otherwise
rural	Dummy variable equal to 1 if the respondent is living in a rural area when interviewed, and 0 otherwise
chronics	Sum of chronic diseases the respondent has, including hypertension, dyslipidaemia, diabetes, cancer, chronic lung diseases, liver disease, heart problems, stroke, kidney disease, stomach disease, emotional problems, memory-related disease, arthritis, and asthma
slfhvg	Dummy variable equal to 1 if the respondent's health status is very good, and 0 otherwise

slfhg	Dummy variable equal to 1 if the respondent's health status is good, and 0 otherwise
slfhf	Dummy variable equal to 1 if the respondent's health status is fair, and 0 otherwise.
slfhp	Dummy variable equal to 1 if the respondent's health status is poor, and 0 otherwise
social	Continuous variable representing the sum of social activities conducted by the respondent in the last month of the wave, including interacting with friends, playing cards, providing help to family or friends, going to a park, taking part in a community-related organisation, doing voluntary work, caring for a sick adult who does not live with them, attending an educational course, investing in stocks, using the internet, and other activities
CESD	Sum of the answers provided by the respondent to the 10 questions from the Center for Epidemiologic Studies. A higher score means a higher level of depression
noworking	Dummy variable equal to 1 if the respondent does not work, and 0 otherwise.
workagri	Dummy variable equal to 1 if the respondent is engaging in agricultural work, and 0 otherwise
workemp	Dummy variable equal to 1 if the respondent is engaging in employed work, and 0 otherwise
workslf	Dummy variable equal to 1 if the respondent is engaging in self-employed work, and 0 otherwise
hsptim1y	Number of times the respondents received inpatient care during the past year

rwealth	Continuous variable representing the amount of per capita wealth. Wealth is defined by the sum of household assets, including present market value of the house, vehicles, fixed capital assets (such as tractor, refrigerator, washing machine, computer, valuable decorations, treasures, and so on), annual rent income of the land, current value of all livestock, and financial assets (such as cash, deposit, stocks, mutual funds, bonds, public housing fund, jizikuan, unpaid salary)
wealth	Continuous variable representing the natural logarithm of rwealth

Appendix 2.2 Multivariate Probit model

Table A. Multivariate Probit model

	(1) smoking	(2) drinking	(3) inactivity
$male_t$	1.843***	1.537***	-0.016
	(39.31)	(24.34)	(-0.44)
age_t	0.086***	0.052*	-0.099***
	(3.11)	(1.66)	(-4.74)
$age2_t$	-0.001***	-0.000*	0.001***
	(-3.42)	(-1.65)	(5.26)
$rural_t$	0.102**	0.094*	0.079**
	(2.21)	(1.81)	(2.02)
$above_{t-1}$	-0.058	-0.101	-0.115**
	(-0.85)	(-1.20)	(-2.07)
$primary_{t-1}$	0.052	-0.005	0.011
	(0.90)	(-0.07)	(0.25)
$income_{t-1}$	-0.014	-0.000	-0.016**
	(-1.64)	(-0.02)	(-2.07)
$married_{t-1}$	0.070	-0.078	-0.045
	(0.47)	(-0.44)	(-0.31)
$divorced_{t-1}$	0.341	-0.059	0.063
	(1.40)	(-0.20)	(0.28)
$widowed_{t-1}$	0.384**	-0.168	-0.021
	(2.38)	(-0.85)	(-0.13)
$chronics_{t-1}$	-0.065***	-0.062***	0.026**
	(-4.17)	(-3.44)	(2.15)
$slfhvg_{t-1}$	-0.010	0.182***	-0.001
	(-0.16)	(2.69)	(-0.02)
$slfhg_{t-1}$	-0.031	0.091*	-0.015
	(-0.65)	(1.65)	(-0.34)
$slfhp_{t-1}$	-0.037	-0.221***	0.087**
	(-0.72)	(-3.26)	(2.03)
$social_{t-1}$	0.047**	-0.015	-0.048**
	(2.11)	(-0.55)	(-2.23)
$CESD_{t-1}$	0.012***	0.002	-0.003
	(3.45)	(0.37)	(-1.05)
M_{t-1}	0.005	-0.033**	-0.028***
	(0.42)	(-2.29)	(-2.68)
MI_{t-1}	-0.020**	-0.001	-0.010
	(-2.53)	(-0.14)	(-1.44)
$workagri_{t-1}$	0.013	0.142**	-0.442***
	(0.25)	(2.26)	(-10.69)
$workemp_{t-1}$	-0.068	0.116	-0.154***
	(-1.05)	(1.47)	(-2.62)
$workslf_{t-1}$	-0.116	0.056	-0.116*
	(-1.50)	(0.61)	(-1.70)
$hsptim1y_{t-1}$	-0.173***	-0.210***	0.043

	(-4.02)	(-3.86)	(1.27)
$wealth_{t-1}$	-0.025**	0.012	-0.023***
	(-2.38)	(0.96)	(-2.62)
rho21	0.209***		
rho31	-0.002		
rho32	-0.080***		
N	9088	9088	9088
r2_p			

Note: Columns 1 to 3 report the estimates of the probability of smoking, heavy drinking, and physical inactivity, respectively. Marginal effects are reported, with *z* statistics in parentheses. *N* denotes the number of observations. Illiterate, single, fair self-reported health status and unemployed are omitted categories. Precise definitions of all variables can be found in Appendix 2.1. Income and wealth are expressed in logarithms. Province and year dummies are included, but their marginal effects are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Chapter Three: What Determines Smoking Cessation? Evidence from the Middle-aged and Elderly Chinese Population

3.1 Introduction

Cigarette smoking is prevalent around the world, but quitting seems to be very difficult for current smokers. Babb *et al.* (2017) use data from the 2000, 2005, 2010, and 2015 National Health Interview Surveys, and report that only 5.7% of American respondents quit smoking within the last one year when the survey was conducted in 2000, and the corresponding rate of smoking cessation increased to 7.4% in 2015. Similarly, the prevalence of smoking cessation is also very low in China. Based on the 2006 International Tobacco Control China Survey, Jiang *et al*. (2010) show that 75.6% of the current smokers had no plan to quit smoking in their lifetime. Furthermore, Li *et al.* (2011a) undertake a study based on the Global Adult Tobacco Survey (GATS) conducted in China in 2010. They show that 16.9% of respondents who have smoked at some time had quit smoking successfully when the survey is conducted.

Although it is well known that smoking is harmful to people's health, there is a misconception of smoking cessation among elderly people. Some of them believe that the damage caused by cigarettes smoking cannot be recovered, and it is too late to quit (Allen, 2008; Appel and Aldrich, 2003; Kerr *et al*., 2006). However, a large amount of studies shows that smoking cessation is necessary and beneficial, even for elderly smokers. Doll *et al.* (2004) compare the risks of mortality when cigarette smoking is stopped at different ages among male British doctors from 1951 to 2001, and consider excess mortality by smoking due to vascular, neoplastic, and respiratory diseases. They show that men who smoked either currently or sometime in their lifetime die on average about 10 years younger than lifelong non-smokers. However, smoking cessation at age 50 reduces the percentage of mortality by 50 percent, and cessation at 30 eliminates almost all of it. In line with these findings, using the 1997 to 2004 waves of the US National Health Interview Survey, Jha *et al*. (2013) show that quitting at age 25 to 34, 35 to 44, and 45 to 54 helps to extend the respondents' life in 10, 9, and 6 years, respectively.

Furthermore, Lam *et al.* (2007) examine the relationship of smoking with all-cause and major cause-specific mortality in elderly Chinese in Hong Kong from 1998 to 2000. The results show that, compared with persistent smoking, smoking cessation decreases the risks of death significantly from lung cancer, other cancers, stroke, and cardiovascular diseases. Lim *et al.* (2013) and Chang *et al.* (2014) confirm the benefits of quitting smoking among older Chinese smokers. Therefore, quitting smoking improves the quality of life, and increases length of life for the older persons in China.

Smoking cessation among the middle-aged and elderly people is really important in China. Firstly, older people are as willing as younger adults to quit smoking (Andrews *et al.,* 2004; Lee *et al*., 2016), but the difficulty of quitting for them is higher than young adults because of longstanding habits and nicotine dependence (Choi and DiNitto, 2014; Kim *et al*., 2013; Messer *et al*., 2008). Secondly, China is facing an ageing population. It is expected that one-fifth to one-third of China's population will be aged 65 and over by the end of 2050 (Mai *et al.,* 2013). Additionally, aged population have a higher medical economic burden than younger generation, because of worse health condition. For example, based on the medical expense data in 2014 in the Liaoning Province, China, Yang *et al.* (2016) document that there are 4.88 million elderly people in Liaoning province, which accounts for 11.52% of the total Liaoning population. Furthermore, they show that the medical expenses related to the aged population constitutes 29.71% of the total Liaoning medical expenses. Compared with other age groups, the per capita medical expense of the elderly people is nearly three times that of adults aged from 14 to 65, and four times that of children aged from 0 to 14. A majority of medical expenses among the aged population are spent treating diseases, such as cardiovascular diseases, respiratory diseases, and digestive diseases. Therefore, promoting middle-aged and elderly people to stop smoking reduces the risk of contracting these diseases, which ultimately releases the medical economic burden of the aged people.

Many studies have identified time to the first cigarette (TTFC) after waking in the morning as a determinant of smoking cessation among adults in countries, such as the United States, Australia, the United Kingdom, Canada, Poland, Russia, and other 15 member countries in the European Commission. They find that higher TTFC contributes to increasing likelihood of smoking cessation (Baker *et al.*, 2007; Fagerstrom *et al.*, 2003; Hyland *et al.*, 2004; Mercincavage *et al.*, 2013; West *et al.* 2001).

Some studies examine the determinants of smoking cessation among the middle-aged and elderly people in China, and show that age, gender, marital status, type of housing, educational level, income, occupation, living with others, drinking status, quitting drinking, duration of smoking history, cigarettes smoked per day (CPD), health-related problems, self-reported health status, hospitalization, emergency visit, newly diagnosis of diseases, and a smoke-free home are important predictors of quitting smoking (Abdullah *et al*., 2006; Abdullah *et al*., 2008; Tsai *et al*., 2012; Lee *et al*., 2016). However, to the best of our knowledge, no study has been conducted in China examining the relationship between TTFC and smoking cessation, especially among the middle-aged and elderly people. This chapter fills this gap in the literature by investigating the extent to which the TTFC may be associated with the probability of smoking cessation with a nationally representative dataset.

Using the 2011, 2013, and 2015 waves of the China Health and Retirement Longitudinal Survey (CHARLS), we find that the TTFC is an effective predictor of transitions out of smoking for middle-aged and elderly people in China. Also, the association between TTFC and the probability of smoking cessation is higher among light smokers than heavy smokers. These results are robust to using alternative measure of the TTFC.

Our results suggest that more health education should be provided targeting middle-aged and elderly people in China to increase the awareness of benefits of smoking cessation. Moreover, in clinical trials, a wide range of pharmacotherapy methods have been shown to be effective in smoking cessation, such as nicotine replacement therapy, varenicline, bupropion (Zwar *et al.,* 2014). However, advising people to delay the time to their first cigarette in the morning after waking up is also an effective method for smoking cessation, which may help people reduce their dependence level to nicotine. Also, it is easy to implement for policy makers with low costs.

The remainder of this chapter is laid out as follows. Section 3.2 surveys the existing literature. Section 3.3 introduces our dataset. Section 3.4 presents our hypotheses. Section 3.5 describes our specifications and estimation methodology. Section 3.6 illustrates some descriptive statistics. Section 3.7 presents the empirical results, and section 3.8 concludes.

3.2 Literature review

3.2.1 Studies on countries other than China

Many studies show that price of cigarettes is an important factor explaining the probability of smoking cessation. For example, Levy *et al.* (2005) examine the role of demographic variation, the number of cigarettes smoked, and the presence of tobacco control polices on recent smoking cessation in the US. Using data from the 1998 to 1999 Tobacco Use Supplement, which consists of 27,115 daily smokers aged 25 or older, they find that those living in a state with higher average cigarette prices are more likely to make quit attempts, but there is no relationship between the prices of cigarettes and remaining abstinent in the three months after quitting. However, price increases are significantly associated with an increased probability of remaining abstinent for more than three months. Including the interaction term between price and age shows that the effect of price on cessation does not differ by age. Levy *et al.* (2005) explain that increased prices lead to higher costs to consumers, which results in a higher likelihood of cessation. Reed *et al.* (2008) use the California Tobacco Surveys, in which 5,888 and 3,595 smokers aged 25 or older are interviewed in the 1996 and 1999 waves, respectively. From November 1998 to January 1999, the price of cigarettes per pack in California increased by 95 cents. Accordingly, the rate of quit attempts and the rate of successful smoking cessation increase immediately in the months following the price increases. This finding is consistent with Levy *et al.* (2005). Similarly, Tabuchi *et al.* (2016) estimate the impact of tobacco price increase in 2010 on smoking cessation in Japan. Using combined data from the Comprehensive Survey of Living Conditions of People on Health and Welfare and the National Health and Nutritional Survey in 2007 and 2010, 2,702 smokers aged 20 to 79 are assessed. Controlling for gender, age, and SES, the authors suggest that the increase in tobacco price is significantly associated with a higher likelihood of smoking cessation in Japan, especially among people who smoke more than 11 cigarettes per day. Moreover, consistent with Levy *et al.* (2005), they find no difference in the effect of price on cessation by age.

Brega *et al*. (2008) test the influence of cognition on successful smoking cessation among older adults in the US. Data are collected from the San Luis Valley Health and Ageing study with 1,338 individuals aged 60 and older in two counties in rural southern Colorado in 1993. After controlling for age, education, and ethnicity, they show that older persons with higher executive cognitive ability have a higher likelihood of stopping smoking. Also, the averaged level of the cognitive ability among current smokers is lower than that among former smokers. Brega *et al*. (2008) explain that higher cognitive ability implies a good ability to make a deliberate effort and good memory, which are essential abilities to inhibiting smoking behaviours. Hence, smokers with higher cognitive ability are more likely to recognise the risk of smoking-related diseases and to quit smoking.

Depression and drinking behaviour are known as critical determinants for the maintenance of smoking behaviour among the aged population. For example, Kenney *et al*. (2009) investigate the association between depression and smoking cessation in the US. A cohort of 442 participants aged 55 or older are assessed over a decade. They find that respondents with depressive symptoms are less likely to stop smoking. For those respondents with drinking problems, the relationship between depression and smoking cessation is strengthened. It is possible that individuals who experience depression and drinking problems are more likely to alleviate their emotional distress through smoking, because smoking is often seen as a method of self-medication. Additionally, depressed individuals are not concerned about the damage to the body caused by smoking and drinking. Hence, the rate of smoking cessation is lower among depressed respondents with drinking problems in the US. Kim *et al*. (2013) reach a similar conclusion by investigating the factors correlated with giving up smoking among elderly Koreans. 1,118 Koreans aged 65 years or older living in Seongnam, Korea, participate in the Korean Longitudinal Study on Health and Ageing between 2005 and 2006. The authors find that those respondents who are depressed or drink alcohol have a lower probability of smoking cessation.

Health-related problems are associated with a higher likelihood of smoking cessation. Chaaya *et al.* (2006) investigate the predictors of smoking cessation in the elderly population in Lebanon. Cross-sectional data is collected from the Urban Health Study between 2002 and 2003. 740 interviewees aged from 60 to 96 are taken from three poor urban communities in Beirut, Lebanon. Adjusting for age, gender, education, income, work status, subjective health evaluation, and depression, Chaaya *et al.* (2006) find that having at least one chronic disease and suffering from disability is significantly associated with an increased rate of smoking cessation. One possible reason for this finding is that Lebanon is a developing country, and the majority of the elderly poor individuals do not have a primary care physician and do not visit clinics regularly for care. Hence, they can only receive advice on quitting from a physician if they have serious health problems. Newly diagnosed chronic diseases are also proved to be a significant determinant of smoking cessation. For instance, Keenan (2009) examine the influence of recent diagnosed health events on quitting smoking and weight loss in the US. To this end, they use 7,764 middle-aged and older respondents interviewed in the Health and Retirement Study (HRS) from 1992 to 2000. Health events include new diagnoses of stroke, cancer, lung disease, heart disease, and diabetes. Controlling for age, gender, race, education, BMI, health status, marital status, employment status, income, and wealth, the authors find that those who are diagnosed with any one of the above diseases is 2.9 times more likely to give up smoking, and those who are diagnosed with several diseases are 6.1 times more likely to stop smoking. This can be explained considering that those with new diagnoses are more motivated to adopt healthier habits, because they wish to get well or avoid worsening symptoms.

Social contextual factors are important causes of quitting smoking. Holahan *et al*. (2012) examine the influence of social factors on smoking status, smoking level, smoking cessation, and smoking relapse among women in the US. Making use of 37,027 respondents aged 50 to 79 in the Women's Health Initiative Observational Study from 1993 to 1998, they analyse two social factors, including social support and living with a smoker. Controlling for age, education, and race, their results show that a higher social support is positively associated with smoking cessation, whilst living with a smoker is negatively associated with cessation. Holahan *et al*. (2012) explain that social support generally comes from family members and friends, who may explicitly endorse healthy behaviours instead of health-risky ones. Also, social support may reduce the level of stress and depression of the respondents, which may make it easier for them to adopt healthier behaviours. By contrast, living with a smoker increases the availability of cigarettes, and, as such, provides a barrier to stop smoking. Choi and DiNitto (2014) investigate the effect of new diagnosis of chronic disease, social isolation, and depression on smoking cessation over 1-year period among the older American adults. Using data from the 2011 and 2012 waves of the US National Health and Ageing Trends Study, which includes 412 smokers aged 65 or older at baseline, Choi and DiNitto (2014) find that 11.1% of the smokers quit smoking successfully in the survey. Social isolation is a dummy variable and is defined as 1 if the respondent has no people to talk about important things, and 0 otherwise. Controlling for sociodemographic characteristics and number of chronic illnesses at baseline, older male respondents with lower CPD, who show social isolation at baseline have a higher probability of giving up smoking. This finding is opposite to Holahan *et al*. (2012). It can be explained considering that socially isolated older smokers may have serious health problems, which forces them to quit smoking. Additionally, socially isolated older smokers are less likely to live with other smokers and can stop smoking easily.

3.2.2 Studies on China

Few studies have investigated the determinants of smoking cessation among the middle-aged and elderly people in China. Abdullah *et al*. (2006) examine the factors associated with smoking cessation and intention to quit among Chinese elderly smokers in Hong Kong. 1,318 respondents who are older than 60 are interviewed in 2001. Controlling for demographic characteristics, the authors find that living with others, needing assistance in mobility, being non-drinkers, having smoked for a shorter duration, and smoking a higher number of CPD are significant predictors of successfully quitting. In accordance with Holahan *et al*. (2012), respondents who are not living alone are more likely to receive quitting support from other household members or to be influenced by the household member's non-smoking status. Among the current smokers, those respondents who have at least one health problem, have smoked for shorter duration, and smoke fewer CPD show a strong desire to quit in the future. In line with the finding by Chaaya *et al.* (2006) in Lebanon, the authors explain that smokers with poor health show a higher likelihood of smoking cessation, because they are more likely to suffer from smoking attributable illness and to have received advice on smoking from their physicians. They are also more likely to quit to alleviate their existing disease condition. Consistent with literature on developed countries (Kenney *et al*., 2009; Kim *et al*., 2013), those who do not drink alcohol may smoke less and quit smoking more easily than drinkers. Additionally, the duration of smoking represents the addiction level to nicotine. As a result, smokers with shorter duration are not seriously addicted to nicotine and can quit successfully. However, contrary to the literature on developed countries (Choi and DiNitto, 2014), Abdullah *et al*. (2006) find that the respondents who smoke fewer CPD have a strong desire to quit in the future, but have a low rate of successful cessation, whilst those respondents who smoke more cigarettes have a high probability of successful cessation. A possible explanation is that those smoke fewer CPD are less dependent on nicotine and intend to quit, whilst those smoke a higher number of CPD may suffer from health problems which force them to quit smoking.

Abdullah *et al*. (2008) identify the predictors of quitting smoking with data from a mobile smoking cessation service between 2002 and 2004 in Hong Kong. 365 elderly smokers aged 60 or older are provided with individual health talks and four weeks' free supply of nicotine replacement therapies. Follow-up surveys are arranged at 1, 3, 6 months after the baseline survey. The authors suggest that light smokers who smoke less than 11 cigarettes per day have a higher likelihood of quitting at the 6 months follow up. This result is consistent with the studies in developed countries (Choi and DiNitto, 2014), and shows a negatively relationship between CPD and smoking cessation.

Tsai *et al*. (2012) investigate the sociodemographic and health-related factors associated with smoking cessation among male Taiwanese smokers who are 50 to 66 years old. Using the 1996, 1999, 2003, and 2007 waves of the Survey of Health and Living Status of the Elderly in Taiwan, 673 respondents are included in the analyses. In line with the literature (Abdullah *et al*., 2006; Abdullah *et al*., 2008; Choi and DiNitto, 2014; Holahan *et al*., 2012; Keenan, 2009), they find that older age, living with a spouse, having a new functional impairment, hospitalization, emergency visit, new diagnosis of heart disease, and quitting drinking are positively related with smoking cessation. By contrast, factors, such as a longer smoking history, higher CPD, and a higher educational level, are negatively related to it.

Lee *et al*. (2016) examine the prevalence and predictors of smoking cessation among older smokers in Taiwan. Their data is taken from the 2008, 2009, 2010, and 2012 waves of the Taiwan Adult Smoking Behaviour Survey, which includes 4,081smokers aged 50 and above. They show that the annual successful cessation rate is 4.7%. In addition, those with a higher educational level, poorer health status, and smoke-free home are more likely to make quit attempts, and those who are older, have higher income, and live in smoke-free homes are more likely to quit successfully. Lee *et al*. (2016) explain that older smokers are motivated by health-related problems to make quit attempts. Hence, those who perceive poor health have an intention to quit. Also, they discuss that higher income is not associated with quit attempts but is significantly associated with successful cessation. A possible reason is that affluent people have an easy access to smoking cessation trials, whilst people with lower income find it difficult to afford the cost. Moreover, the effect of educational level on smoking cessation is mixed in the literature (Kim *et al*., 2013; Tsai *et al*., 2012), the authors explain that different countries and population groups may contribute to the differences.

3.2.3 Relationship between TTFC and smoking cessation

A number of studies find a significantly positive association between TTFC and smoking cessation among adults. For example, West *et al.* (2001) examine determinants of making quit attempts, as well as the factors influence the success of these quit attempts. 1,012 smokers aged 16 or over in England and Wales are interviewed in 1996, and the follow-up study is conducted one year later in 1997. West *et al.* (2001) identify two determinants of quit attempts that one is the respondent know that smoking is harmful to his/her future health, and the other is the respondent has a partner who dislikes his/her smoking. Both these two determinants are positively related to the chance of making a quit attempt. However, those respondents who enjoy smoking at baseline are less likely to make a quit attempt. In addition, higher TTFC and starting smoking at older ages are found to be positively associated with the likelihood of quitting smoking successfully. A possible reason is that lower TTFC may be due to a higher dependence level of nicotine, and cigarette addition prevents smokers from quitting successfully.

The association between TTFC and quitting smoking is also examined by Fagerstrom *et al.* (2003). A cohort of 10,295 smokers from all the 15 member countries of the European Commission and Poland and Russia are surveyed in 2003. Fagerstrom *et al.* (2003) divide the respondents in two groups. The high dependence group consists of respondents who smoke their first cigarette after waking up within 30 minutes, while the low dependence group includes respondents who smoke their first cigarette after 30 minutes. Similarly, two groups of respondents are categorized by CPD's cut-off point at 20, which consist of those who smoke 20 or over cigarettes per day and those who smoke less than 20 cigarettes per day. The authors find that both CPD and TTFC predict dependence well. Specifically, smokers with TTFC higher than 30 minutes can give up smoking easily and show a longer duration of no smoking the last time they quit. Smokers who smoke less than 20 cigarettes per day tend to quit smoking easily and show a longer duration of quit attempt the last time they quit. The authors explain their findings considering that around half of the nicotine is converted into cotinine after two hours of smoking, but the normal sleeping duration is up to 8 hours, which means that there is hardly any nicotine left in the body in the morning. Hence, all smokers wake up in a state of nicotine deprivation. A lower TTFC represents a higher level of dependence to nicotine and predicts the difficulty of smoking cessation.

Consistent with Fagerstrom *et al.,* (2003) and West *et al.,* (2001), Hyland *et al.* (2004) identify the predictors of smoking cessation in the US. Data are collected from the National Cancer Institute Community Intervention Trial for Smoking Cessation study in 1988, 1993, and 2001 with 6,603 subjects aged 25 to 64. Hyland *et al.* (2004) report that nicotine dependence is negatively associated with smoking cessation. Consequently, a lower TTFC and higher CPD are associated with a lower likelihood of smoking cessation.

Baker *et al.* (2007) investigate the relationship between TTFC and smoking cessation among adults. Based on data derived from three large smoking cessation trials conducted by the University of Wisconsin in 2000, one clinical trial conducted by Yale University in 2006, and the International Tobacco Control Survey (ITC) conducted between 2002 and 2003 covering respondents from Canada, the UK, the US, and Australia, 8,628 respondents are considered in their study. Controlling for demographic variables, intention to quit, past quit attempts, longest past quit attempt, smoking frequency, opinion about smoking, self-efficacy, worries about health and quality of life, and favourable attitudes about smoking, Baker *et al.* (2007) suggest that TTFC significantly predicts smoking cessation both in the short time period of one week and in a long time period of six months. Specifically, in comparison with those who smoke within five minutes of waking, those respondents smoking 60 minutes or more after waking have a higher quit rate. A possible reason for this phenomenon is that the smoking ban at home delays the time to first cigarette early in the morning, which ultimately reduces their dependence to nicotine. Also, in a family with the smoking ban, family members are more likely to advise smokers to quit. These family supports could promote smoking cessation.

Mercincavage *et al.* (2013) confirm the positive relationship between TTFC and quitting smoking. They examine this relationship among American adolescents by using 1,167 adolescent smokers aged 14 to 19 participate in the American Lung Association's Not-On-Tobacco cessation programme from 1997 to 2008. Controlling for age, gender, motivation to quit smoking, confidence in ability to quit smoking, and the number of cigarettes smoked consumed every day, they find that adolescents who smoke the first cigarette within 30 minutes of waking up have a low probability of quitting smoking at the end of the three month programme, compared with adolescents who smoke the first cigarette after 30 minutes of waking. Consistent with the previous explanations (Fagerstrom *et al.,* 2003; Hyland *et al.*, 2004; West *et al.* 2001), it is possible that TTFC is a behavioural indicator of nicotine dependence. Smokers with a higher nicotine dependence level may experience more severe overnight nicotine abstinence, which drives them to smoke the first cigarette earlier in the morning and increases the difficulty of smoking cessation.

3.2.4 Contribution

We contribute to the literature by examining the determinants of smoking cessation among middle-aged and elderly people in China, focusing on the role of TTFC. To this end, we will use a national representative panel dataset. To the best of our knowledge, no empirical study has been conducted investigating the influence of TTFC on smoking cessation in the Chinese context. Furthermore, all the previous studies analysing the association between TTFC and smoking cessation have used cross-sectional datasets, while we are the first to use panel dataset in the Chinese context.

3.3 Data

3.3.1 The China Health and Retirement Longitudinal Survey

We use the 2011, 2013, and 2015 waves of the CHARLS. In order to make use of information on the number of cigarettes smoked per day, and the price per pack of cigarettes, we only consider those smoking filtered cigarettes and unfiltered cigarettes as smokers, and exclude smoking a pipe, self-rolled cigarettes, cigars, and/or water cigarettes. Data such as whether the respondents have ever smoked in the past, whether the respondents are still smoking, the age at which the respondents quit smoking, the age at which the respondents started smoking, the number of cigarettes smoked per day, and the price per pack of cigarettes are available at all the three waves. However, TTFC after waking up is only available in the first two waves. As we study smoking cessation, lifetime abstainers and former smokers are excluded from the analysis. In order to control for the potential effect of extreme values and faulty data, we drop observations below the 1st percentile and above the 99th percentile for each continuous main regression variable. A total of 3,096 observations for respondents aged 45 to 87 old are included in our dataset. Our panel dataset is balanced, in which each respondent is observed in all the waves. This accounts for unobserved heterogeneity by each individual and reduces attrition bias by missing values.

3.3.2 Variables

3.3.2.1 Smoking-related variables

Based on the questionnaire, respondents are asked whether they have ever smoked in the past, as well as whether they are still smoking when the survey is conducted. If the answers to both questions are 'yes', they are defined as smokers. If any answer is 'No', they are defined as non-smokers. The variable *smoke_stop* represents a dummy variable equals to 1 if the respondent was a smoker in the previous wave and is no longer a smoker in the current wave, and equals to 0 if the respondent is a smoker in both the lagged and current waves.

The variable *smokenum* indicates the number of cigarettes smoked per day. The variable *smoknumD* is a dummy equal to 1 if the respondent is a heavy smoker, and 0 otherwise. A heavy smoker is defined as a person smoking more than 20 cigarettes per day (Lei *et al.*, 2014).

3.3.2.2 TTFC variables

TTFC represents the time until the respondents' first cigarette after waking up in the morning. It is based on the question: 'How soon after you wake up do you smoke your first cigarette?'. There are four options the respondents can choose from, namely 'Within 5 minutes', 'Within 6 to 30 minutes', 'Within 31 to 60 minutes', and 'More than 1 hour'. The variable *TTFC* equals to 1 if the respondent chooses 'Within 5 minutes', equals to 2 if he/she chooses 'Within 6 to 30 minutes', equals to 3 if he/she chooses 'Within 31 to 60 minutes', and equals to 4 if he/she chooses 'More than 1 hour'.

We also construct four categorical variables relative to the TTFC. Specifically, *T1* equals to 1 if the respondent smokes his/her first cigarette within 5 minutes after waking up, and 0 otherwise. *T2* equals to 1 if the respondent smokes his/her first cigarette within 6 to 30 minutes after waking up, and 0 otherwise. *T3* equals to 1 if the respondent smokes his/her first cigarette within 31 to 60 minutes after waking up, and 0 otherwise. *T4* equals to 1 if the respondent smokes his/her first cigarette more than 1 hour after waking up, and 0 otherwise.

3.4 Hypotheses

We expect a higher TTFC to be a predictor of smoking cessation, because TTFC represents the level of dependence to nicotine (Fagerstrom *et al.*, 2003; Hyland *et al.*, 2004; Mercincavage *et al.*, 2013; West *et al.* 2001). Specifically, the urgency to smoke after a long period of withdrawal overnight implies a higher addiction level to nicotine.

Moreover, it has been proved that around half of the nicotine disappears after two hours of smoking (Benowitz *et al.*, 1982). If we assume a normal sleeping duration is 8 hours, this means that there is hardly any nicotine left in the smokers' body after waking up in the morning. Hence, all the smokers are in a state of nicotine deprivation at the moment of waking up in the morning. The TTFC in the morning after waking up represents the urgency to smoke after the longest nicotine deprivation during the day, which is up to 8 hours. Hence, the TTFC can estimate the dependence level to nicotine accurately.

Therefore, the TTFC is expected to be positively associated with smoking cessation. According to the study by Fu *et al*. (2011) and Muscat *et al*. (2009), respondents who experience an urgency to smoke in the morning are more likely to have a higher cotinine level, controlling for cigarette consumption. This negative relationship between the TTFC and cotinine level represents their response to overnight withdrawal. Fu *et al*. (2011) and Muscat *et al*. (2009) explain that those respondents smoking their first cigarettes earlier in the morning may inhale more nicotine, and have a higher addiction to nicotine, which makes it more difficult for them to quit.

A positive relationship between the TTFC and smoking cessation has been found in existing studies among adults in countries, such as the US, Australia, the United Kingdom, Canada, Poland, Russia, and 15 member countries of the European Commission (Baker *et al.*, 2007; Fagerstrom *et al.*, 2003; Hyland *et al.*, 2004; Mercincavage *et al.*, 2013; West *et al.* 2001). We expect to observe a similar positive association between the TTFC and smoking cessation among middle-aged and elderly people in China. The following Hypothesis summarises these arguments:

H1: Ceteris paribus, a higher TTFC is associated with a higher probability of smoking cessation.

Tabuchi *et al.* (2016) argue that the association between cigarette price and smoking cessation differs according to the number of cigarettes smoked per day. To the best of our knowledge, the effect of CPD on the association between TTFC and smoking cessation is unclear. We expect the association between the TTFC and smoking cessation will also vary by number of cigarettes smoked per day. Specifically, we expect the effect of the TTFC on smoking cessation to be more powerful among light smokers, but less powerful among heavy smokers. This is because heavy smokers are more dependent on nicotine, and cigarette smoking becomes a natural behaviour every day. Therefore, quitting smoking by delaying the TTFC after waking up will be harder for frequent smokers. Hypothesis two follows:

H2: Ceteris paribus, the association between the TTFC and the probability of smoking cessation is higher among light smokers than that among heavy smokers.

Baker *et al*. (2007) show that a possible reason for the positive relationship between TTFC and cessation is a home smoking ban. They show that if there is a smoking ban at home, people have no choice but to smoke their first cigarette after they leave house in the morning, which ultimately reduces their dependence to nicotine. Also, in a family with a smoking ban, family members are more likely to advise smokers to quit. These family supports could promote smoking cessation. As a result, we expect respondents to be more likely to quit if their spouses are not smoking, because a home smoking ban may exist in their family. Conversely, we expect that living with spouse who is a smoker leads to a lower chance of smoking cessation, because these respondents have a higher availability of cigarettes in the house (Holahan *et al.*, 2012). Therefore, we propose the following hypothesis:

H3: Ceteris paribus, a respondent whose spouse is also a smoker will show a lower probability of smoking cessation.

3.5 Specification and estimation methodology

3.5.1 Baseline specification

To document the effect of the TTFC on smoking cessation among middle-aged and elderly people in China, we estimate the following baseline specifications:

$$smoke\_stop_{i,t} = \beta_0 + \beta_1 D_{i,t} + \beta_2 X_{i,t-1} + \beta_3 TTFC_{i,t-1} + \varepsilon_i + v_p + v_t \qquad (1)$$

where subscript i indexes individuals and t indicates waves. The dependent variable *smoke_stop* and *TTFC* in Eq. 1 have been defined above. $D_{i,t}$ is a vector of control variables at time t, $X_{i,t-1}$ is a vector of control variables at time t-1, ε_i is the error term controlling for an individual effect, v_p is the error term for a province specific effect, and v_t is the error term for a time-specific effect. We account for the province specific effect including provincial dummies in the model and for the time-specific effect including time dummies in the model.

Our second Equation includes the change of control variables between waves into the baseline specifications, to take into account the fact that what happens between waves may also affect the dependent variable. The following model is therefore estimated:

$$smoke\_stop_{i,t} = \beta_0 + \beta_1 D_{i,t} + \beta_2 X_{i,t-1} + \beta_3 TTFC_{i,t-1} + \beta_4 dX_{i,t} + \varepsilon_i + v_p + v_t \quad (2)$$

where $dX_{i,t}$ indicates the change of control variables between waves. According to Hypothesis 1, $'\beta_3'$ in Eq. 1 and Eq. 2 are expected to be significant and positive, implying that a delayed time to the first cigarette after waking up is significantly positively associated with the probability of smoking cessation when all the other variables are kept constant.

In addition, the third Equation includes the interactions between TTFC and a dummy indicating whether the respondent is a heavy smoker into the baseline specifications, to take into account the fact that the relationship between TTFC and smoking cessation may vary by number of cigarettes smoked per day. The following model is estimated:

$$smoke\_stop_{i,t} = \beta_0 + \beta_1 D_{i,t} + \beta_2 X_{i,t-1} + \beta_3 TTFC_{i,t-1} + \beta_4 TTFC_{i,t-1} * smoknumD_{i,t-1} + \varepsilon_i + v_p + v_t \quad (3)$$

The variable *smoknumD* implies whether the respondent is a heavy smoker or not. In line with Hypothesis 2, $'\beta_4'$ in Eq. 3 is expected to be significant and negative: compared with light smokers, heavy smokers show a weaker association between TTFC and smoking cessation, holding other variables constant.

Furthermore, living with spouse who is a smoker may be associated with smoking cessation. In order to test the spousal effect, we estimate the Equation 4:

$$smoke\_stop_{i,t} = \beta_0 + \beta_1 D_{i,t} + \beta_2 X_{i,t-1} + \beta_3 TTFC_{i,t-1} + \beta_4 smokesp_{i,t-1} + \varepsilon_i + v_p + v_t \quad (4)$$

where *smokesp* is a dummy equal to 1 if the respondent' spouse is a smoker, and 0 otherwise. Consistent with Hypothesis 3, $'\beta_4'$ in Eq. 4 is assumed to be significantly negative: having a smoker spouse is associated with a lower probability of quitting smoking.

The control variables at time *t* include gender, age, age squared, living in a rural area or not. Also, the control variables at time *t-1* are the number of cigarettes smoked per day, the price of cigarettes, the duration of smoking, the educational level, income per capita, the number of chronic diseases, marital status, self-reported health status, the number of social activities, the depression level, cognitive abilities, work type, the number of times the respondent used inpatient care, wealth, and heavy drinking status. Additionally, we include the following change of status control variables: new retirement, new diagnosis of chronic diseases, quitting drinking, change in per capita income, change in per capita wealth, change in marital status, accident, change in self-reported health status, change in cognitive ability, change in social activities, change in depression status, and change in inpatient care. The definitions of all these variables are shown in the Appendix 3.1.

Following Lei and Liu (2018), we include two measures of cognitive abilities in our models, namely memory and mental intactness. Specifically, the variable *M* is a continuous variable averaging immediate memory and delayed memory. *MI* represents mental intactness. It is a continuous variable that sums the correct answers from 11 questions, including successively subtracting 7 from 100 (up to five times), awareness of the date, month, year (using either solar or lunar calendar), the day of the week and season of the year, and the ability to redraw a picture of two overlapping pentagons.

According to the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and the Center for Disease Control and Prevention (CDC) standard (Lin *et al.*, 2011; Shaw and Agahi, 2012), heavy drinking is typically defined as consuming 15 drinks or more per week for men. While, for women, heavy drinking is defined as consuming 8 drinks or more per week. For people who are older than 65 years old, heavy drinking is also defined as consuming 8 drinks or more per week for both genders. Each standard drink contains roughly 14 grams of pure alcohol. Hence, *drinker* equals to 1 if the respondent is a heavy drinker, and 0 otherwise. *qdrink* equals to 1 if the respondent quit heavy drinking between two consecutive waves, and 0 otherwise.

Age is expected to be positively related to smoking cessation, because older smokers are more likely to be motivated to quit by health problems (Choi and DiNitto, 2014; Lee *et* *al*., 2016; Tsai *et al*., 2012). Following the study by Choi and DiNitto (2014), the variable *male* is assumed to have a positive relationship with cessation. Also, income and wealth are expected to be positively associated with cessation, because affluent people have easy access to smoking cessation services and medicines, whilst poor people find it difficult to afford the costs (Lee *et al*., 2016). Living in rural area is expected to be negatively associated with quitting, because the urban population have better health awareness than the rural population, and the urban population are more likely to be exposed to more antismoking campaigns (Yang *et al.,* 2009).

Mixed results about the relationship between education and smoking have been shown in the literature (Kim *et al*., 2013; Lee *et al*., 2016; Tsai *et al*., 2012). We expect more educated respondents to be more likely to quit, because they are more likely to recognise the risk of smoking-related diseases than less educated ones. We expect married respondents to be more likely to quit, because they are more likely to receive smoking cessation advice from their spouse and they have more support from their spouse (Abdullah *et al.,* 2006). We expect being unemployed or newly retired to be negatively associated with cessation, because people in these categories do not have job to distract their focus, and may feel bored and lonely, which may be the reason why they are highly dependent on cigarettes and find it hard to quit (Chaaya *et al*., 2006; Keenan, 2009; Abdullah *et al*., 2006).

We expect health-related problems to be associated with a higher likelihood of smoking cessation because smokers are more likely to suffer from smoking attributable illness and receive cessation advice on smoking from their physicians. Also, smokers with health-related problems may need to quit to alleviate their existing disease conditions (Abdullah *et al*., 2006; Chaaya *et al*., 2006; Lee *et al*., 2016). For similar reasons, newly diagnosed health-related problems are also expected to be a significant factor for smoking cessation (Choi and DiNitto, 2014; Keenan, 2009; Tsai *et al*., 2012). Hence, the variables *chronics, slfhp, hsptim1y, nchronics, accident, slfhworse, and dhsptim1y* are expected to be positively related with the probability of smoking cessation, whilst *slfhvg, slfhg,* and *slfhbetter* are expected to be negatively associated with the probability of quitting.

According to previous studies (Levy *et al*., 2005; Reed *et al*., 2008; Tabuchi *et al*., 2016), higher cigarette prices and recent price increases are expected to be predictors of smoking cessation, because higher price increases the costs to consumers, who are more likely to stop smoking to prevent the higher costs. Cognitive ability is expected to be positively correlated to quitting smoking, because higher cognitive ability implies a good ability to make a deliberate effort and good memory that are essential abilities to inhibiting smoking behaviours (Brega *et al*., 2008). Depression is assumed to have a negative relationship with smoking cessation, because depressed people are more likely to smoke cigarettes to alleviate their emotional distress, and are less concerned about the damage to their body caused by cigarettes (Chaaya *et al*., 2006; Kenney *et al*., 2009; Kim *et al*., 2013). Thus, quitting smoking is harder for depressed respondents. We expect drinking to be negatively associated with smoking cessation as both drinking and smoking are health-compromising behaviours, and people tend to adopt both of them simultaneously (Abdullah *et al*., 2006; Kenney *et al*., 2009; Kim *et al*., 2013; Tsai *et al*., 2012). For the same reason, those respondents who quit drinking recently are more likely to quit smoking soon.

However, the literature about the relationship between social activities and smoking cessation is unclear. Some studies (Abdullah *et al*., 2006; Holahan *et al.*, 2012; Tsai *et al*., 2012; West *et al.*, 2001) show that this relationship is positive, while other studies (Choi and DiNitto, 2014) report that this relationship is negative. We expect the relationship to be positive, because respondents who have more social activities with family and friends are more likely to receive smoking cessation advice from them, and to show a lower level stress, which makes smoking cessation easier. CPD and the duration of smoking history are expected to have a negative association with smoking cessation, as the higher nicotine addiction among smokers who consume more cigarettes per day and have a longer smoking history prevents them from quitting (Abdullah *et al*., 2006; Abdullah *et al*., 2008; Choi and DiNitto, 2014; Fagerstrom *et al.,* 2003; Tsai *et al*., 2012).

3.5.2 Estimation methodology

We use both linear probability models (LPM) and Probit models to estimate Eq. 1, 2, 3, and 4, as the dependent variable is dichotomous. For Probit models, we report and discuss marginal effects (Baltagi, 1995).

3.6 Summary statistics

Table 3.1 presents descriptive statistics of the main regression variables. Column 1 refers to the full sample, column 2 to the successful stoppers, column 3 to the continuous smokers, and column 4 to the difference between successful stoppers and continuous smokers.

Seven percent of the full sample stopped smoking. For continuous smokers, the average number of cigarettes smoked per day is 17.4. 50.7 percent of the full sample are heavy smokers. The corresponding value among continuous smokers is 56.9 percent.

Consistent with Hypothesis 1, the mean value of TTFC among stoppers is significantly higher than that among continuous smokers at the 5 percent level. In the full sample, the mean value of TTFC is 2.51, and the corresponding values among stoppers and continuous smokers are 2.869 and 2.477. Additionally, stoppers show a significant higher percentage of smoking their first cigarette after one hour of waking up, whilst continuous smokers show a significant higher percentage of smoking their first cigarette within 6 to 30 minutes after waking up. Specifically, in the full sample, the prevalence of smoking their first cigarette within 5 minutes, within 6 to 30 minutes, within 31 to 60 minutes, and within more than one hour are 33.2, 19.8, 9.75, and 37.2 percentage, respectively. The corresponding percentages among stoppers are 29.5, 8.2, 8.2, and 54.1 percent, respectively. The corresponding prevalence among continuous smokers are 33.6, 20.8, 9.9, and 35.7 percent, respectively.

In the full sample, 91 percent of the respondents are males. The average age is 60.02. Around 65 percent of respondents live in rural areas. The average price per pack of cigarettes is 6.143 YUAN. Compared with the price of cigarettes in the previous wave, it increases by around 1 YUAN. The average years of smoking history are 33.86. The percentage of illiterate people, those with a primary school equivalent level, and those with a secondary school or above level are 12.8, 44.7, and 42.4 percent, respectively. The average annual income per capita is 10,263 YUAN, and the average per capita wealth is 50,085 YUAN. The growth in the income and wealth between consecutive waves are 392.4 YUAN and 5,814 YUAN, respectively. 91.1 percent of the respondents are married, whilst only 1.1 percent of the respondents are divorced, and 6.36 percent of the respondents are widowed. Only 0.517 percent of respondents get married within the sample period, 0.485 percent of respondents get divorced, and 1.29 percent of respondents get widowed. As for health status, 11.6 percent of the respondents consider that they have a very good health status. 15 percent of the respondents consider their health status as good, and 58.9 percent as fair. 14.5 percent of the respondents consider that they have a poor health status. Compared with the health status in the previous wave, 7.84 percent of respondents consider their health status improved, whilst 43.7 percent of them feel their health status became worse. The average number of chronic diseases is 1.291. 5.31 percent of the respondents have been in a traffic accident since last wave. The average number of impatient visits is 0.127. The increases in the average number of chronic diseases and average number of impatient visits between consecutive waves are 0.0949 and 0.0365, respectively. 26.1 percent of respondents are heavy drinkers, and only 7.88 percent of them quit heavy drinking between waves. The mean number of social activities undertaken by respondents is 0.579, and the average increase in the score between waves is 0.0264. The average CESD score is 7.098, and the average reduction in the score between waves is 0.17. The average memory score is 3.634 and the average mental intactness score is 8.037. The average reduction in the two scores between waves is respectively 0.118 and 0.462. 22 percent of the respondents are unemployed or retired.

43.9 percent of the people surveyed are working in agriculture, and 23.8 percent of them are employees, whilst 10.3 percent are self-employed. Around 3 percent of the respondents are newly retired. 9.39 percent of the respondents' spouses are smokers.

Compared with continuous smokers, the stoppers are older, females, unemployed or retired, have more chronic diseases, a higher depression level, and are more likely to have a smoker spouse. They also have a shorter smoking history, lower income, poorer health status, fewer social activities, a lower memory level, a lower mental intactness level, have cheaper cigarettes, and are less likely to be heavy drinkers. Additionally, the stoppers show a worsening in self-reported health status, are more likely to be newly retired, show more inpatient hospital visits, are more likely to have been diagnosed with a new chronic disease, and to have experienced a wealth loss.

3.7 Empirical Results

This section reports results of Probit estimations and LPM regressions for the probability of smoking cessation. Two measures of TTFC are used for robustness tests.

3.7.1 Probit and LPM regressions for the probability of smoking cessation

Table 3.2 presents the estimates of the probability of smoking cessation as a function of TTFC. Columns 1 and 2 show the marginal effects of Probit models, and columns 3 and 4 show the coefficients of LPM regressions. Two measures of TTFC are used in the analyses. Specifically, columns 1 and 3 use the variable $TTFC_{t-1}$, whilst columns 2 and 4 use the category variables $T1_{t-1}$, $T2_{t-1}$, and $T3_{t-1}$.

In columns 1 and 3, TTFC is positively and statistically significantly associated with the likelihood of smoking cessation. The results suggest that smoking the first cigarette later in the morning after waking up is associated with a 1.2 percentage point higher probability of quitting smoking. In columns 2 and 4, the variable $T1_{t-1}$and $T2_{t-1}$ are negatively and statistically significantly associated with the likelihood of smoking cessation. There is no significant association between $T3_{t-1}$ and the likelihood of smoking cessation. Specifically, the figures in column 2 are -0.037 and -0.024 for variable $T1_{t-1}$and $T2_{t-1}$, respectively, meaning that compared with those respondents who smoke their first cigarettes more than 1 hour after waking up in the morning, the respondents who smoke within 5 minutes after waking have a 3.7 percentage point lower probability of quitting smoking, and the respondents who smoke within 6 to 30 minutes after waking have a 2.4 percentage point lower probability of quitting smoking. Similar results can be found in column 4. The results based on $T1_{t-1}$, $T2_{t-1}$, and $T3_{t-1}$ also confirm the positive relationship between TTFC and smoking cessation (columns 2 and 4). These results are consistent with Hypothesis 1 and the existing literature (Baker *et al.*, 2007; Fagerstrom *et al.*, 2003; Hyland *et al.*, 2004; Mercincavage *et al.*, 2013; West *et al.* 2001).

Consistent with our expectations and the literature (Abdullah *et al.*, 2006; Abdullah *et al.*, 2008; Chaaya *et al.*, 2006; Choi and DiNitto, 2014; Lee *et al.*, 2016; Tsai *et al.*, 2012), in all four columns, the number of cigarettes smoked per day and the duration of smoking history are negatively and significantly associated with the probability of giving up smoking, while the number of chronic diseases is positively associated with giving up smoking.

However, compared with unmarried respondents, the married and widowed respondents are less likely to quit smoking, which is different from our expectation. These results might be explained considering that married people have more responsibilities for the family (Yang *et al*., 2009). Hence, the respondents tend to have more family and financial pressures. Cigarette smoking may help them release their pressures. For respondents who are widowed, they are more likely to smoke cigarettes to alleviate their missing feelings for their deceased spouse. Also, different from our expectation, poor self-reported health status is negatively associated with cessation, which means that people who consider their health status as poor are less likely to stop smoking. It is reasonable that the middle-aged and elderly people in China are still holding the old-school concept about smoking cessation, and believe that the damage of cigarette smoking in the young age has already been done and it is too late to quit smoking (Allen, 2008; Appel and Aldrich, 2003; Kerr *et al.*, 2006). Therefore, more health education should be provided targeting middle-aged and elderly people to increase the awareness of smoking cessation, and inform the older people about the benefits of stopping smoking even in their time. Furthermore, the number of social activities is shown to be negatively related to cessation. A possible explanation is that the respondents' friends are likely to be smokers, thus, staying with smokers gives the respondents higher availability of cigarettes which prevents them from quitting (Holahan *et al.*, 2012).

3.7.2 Probit and LPM regressions for the probability of smoking cessation accounting for differences in relevant variables between waves

Table 3.3 analyses the effect of TTFC in the previous wave on smoking cessation in the following wave, controlling for changes in control variables. Columns 1 and 2 show the marginal effects of Probit models, and columns 3 and 4 show the coefficients of LPM regressions.

Consistent with the results in Table 3.2, the coefficients/marginal effects associated with TTFC in all the four columns of Table 3.3 show a positive relationship between TTFC and smoking cessation. Specifically, the figures in columns 1 and 3 indicate that if the respondent smokes the first cigarette later after waking, he/she would show a 1.4 percentage point higher probability of quitting. These findings provide further support for our Hypothesis 1. Also, CPD, being married or widowed show negative relationships with cessation, although the association between marital status and cessation is only significant in columns 3 and 4.

Different from the results in Table 3.2, the significant effects of some variables disappear in Table 3.3, including the duration of smoking, number of chronic diseases, self-reported health status, and social activities. Moreover, the coefficients/marginal effects associated with many variables become significant in Table 3.3. Specifically, attending inpatient hospital and new diagnosis of chronic diseases are now significantly and positively associated with smoking cessation in columns 3 and 4. These are consistent with our expectations and the literature (Abdullah *et al*., 2006; Chaaya *et al*., 2006; Choi and

DiNitto, 2014; Keenan, 2009; Lee *et al*., 2016; Tsai *et al*., 2012). Furthermore, the price of a cigarette pack and the price increase between waves are significantly negatively correlated to cessation, which is opposite with our expectation and the literature (Levy *et al*., 2005; Reed *et al*., 2008; Tabuchi *et al*., 2016). A possible reason for the difference is that we use the real cigarette price per pack that varies by brands and regions, instead of the use of the average state level cigarette prices in the previous studies (Levy *et al*., 2005; Reed *et al*., 2008; Tabuchi *et al*., 2016). This can be explained considering that respondents tend to consume more expensive and high-quality cigarettes to reduce the harm of cigarettes on their health, rather than quitting.

When not considering the newly diagnosed diseases, some variables, such as the number of cigarettes smoked per day, the duration of smoking history, the number of chronic diseases, marital status, self-reported poor health status, and the number of social activities, are significant predictors of cessation, as shown in Table 3.2. However, newly diagnosed diseases are a strong trigger for smokers. Hence, when considering newly diagnosed diseases in Table 3.3, control variables, *smokeage*, *chronics*, *slfhp*, and *social*, which are significant in Table 3.2 are insignificant now.

3.7.3 Probit and LPM regressions for the probability of smoking cessation with interaction term between TTFC and a dummy indicating whether the respondent is a heavy smoker

Table 3.4 examines the effect of TTFC on smoking cessation depending on the quantity of cigarettes smoked every day. Column 1 presents the marginal effects of Probit model, and column 2 presents the coefficients obtained from the LPM model.

For heavy smokers, the variable $TTFC_{t-1}$ is positively and statistically significantly associated with cessation in both columns. In column 1, the data indicates that smoking the first cigarette later after waking up is associated with 7.8 percent point higher probability of quitting. Similarly, the figures in column 2 also suggest that a higher TTFC is related to a 0.7 percent point higher probability of quitting.

In comparison, among light smokers, the variable $TTFC_{t-1}$ is also positively and statistically significantly associated with cessation in both columns. Yet, the magnitude of the relationship between TTFC and cessation is higher than that among heavy smokers. For example, in column 1, the data indicates that smoking the first cigarette later after waking up is associated with a 22.4 percent point higher probability of cessation. The corresponding figure in column 2 is 3 percentage points. Both these figures are higher than the counterparts among heavy smokers. This implies that light smokers have a higher association between TTFC and the likelihood of smoking cessation than heavy smokers. A possible reason is that light smokers are less addicted to nicotine. These finding supports our Hypothesis 2.

Consistent with the results in Table 3.2, the duration of smoking history, being married, being widowed, having a poor self-rated health status, and the number of social activities are negatively associated with the chance to quit smoking. However, the number of chronic diseases is positively associated with giving up smoking. This suggests that married or widowed respondents with long smoking history, poor self-rated health status, and more social activities are less likely to give up smoking, while those people with more chronic diseases are more likely to stop smoking.

3.7.4 Probit and LPM regressions for the probability of smoking cessation among married respondents

Table 3.5 reports the Probit regression results for smoking cessation (columns 1 and 2), and the LPM regression results (columns 3 and 4). $smokesp_{t-1}$ is insignificantly associated with quitting in all the four columns. This finding is inconsistent with our Hypothesis 3. It suggests that spousal smoking status has no effect on the respondent's smoking cessation behavior. A possible reason is that the majority of middle-aged and elderly people are more convinced of the doctor's words than the spouse's words.

Consistent with our results in Table 3.2, CPD, duration of smoking history, and poor self-rated health status are negatively associated with the probability of quitting, whilst TTFC is positively associated with it. However, the marginal effects/coefficients associated with the number of social activities and the number of chronic diseases in the previous wave are insignificant in all the columns in Table 3.5.

Similar to Table 3.5, Table 3.6 also investigates the influence of spousal smoking status on cessation among marred respondents. The only difference is that Table 3.6 adds the changes in control variable between the previous and current waves. We also do not find significant association between spousal smoking status and the respondent's smoking cessation. This confirms the results in Table 3.5.

In line with the results in Table 3.3 among the middle-aged and elderly people in China, Table 3.6 reveals that CPD, cigarette prices, and cigarette price increases are negatively associated with the probability of quitting, whilst the TTFC, the number of inpatient visits, and the increase in the number of impatient visits are positively associated with quitting. However, the coefficients/marginal effects associated with the new diagnosis of chronic disease are insignificant in all the columns. In addition, the poor self-reported health status shows a significantly negative relationship with cessation in columns 1 and 2, but not in columns 3 and 4. The difference between married respondents and the full sample may contributes to these different results between Table 3.3 and Table 3.6.

Additionally, Appendix 3.2 adds a temporal dimension to the smoking habits. The negative binomial regression model is used to examine the effect of spousal smoking status on the number of waves the respondents are observed smoking among married respondents. The results are reported in Table A and Table B in Appendix 3.2.

3.8 Conclusion and limitations

We analyse the determinants of smoking cessation among the middle-aged and elderly in China, focusing on the role of TTFC. Making use of the 2011, 2013, and 2015 waves of CHARLS, we find that TTFC is positively associated with the probability of smoking cessation in the Chinese context. Also, the association between TTFC and the probability of smoking cessation is higher among light smokers than that among heavy smokers. These results are robust to using alternative measure of TTFC.

Our results suggest that more health education should be provided targeting middle-aged and elderly people in China to increase their awareness of the benefits of smoking cessation. Moreover, this chapter advises people to delay the time to their first cigarette in the morning after waking up, because this method increases the probability of smoking cessation.

Future research will focus on studying the effect of smoking reduction on smoking cessation after more follow-up waves of CHARLS are published. Also, it is worth studying the determinants of smoking cessation among adolescents and adults in China.

Table 3.1 Descriptive statistics

Variables	Full sample (1)	Stoppers (2)	Continuous smokers (3)	Difference (4)
smoke_stop	0.0704	1	0	
smokenum	15.50	0	17.40	17.397***
male	0.910	0.862	0.914	0.051**
age	60.02	60.92	59.96	-0.966*
rural	0.647	0.624	0.649	0.0250
smoknumD	0.507	0	0.569	0.569***
price	6.143	5.241	6.227	0.986***
smokeage	33.86	29.45	34.20	4.750***
above	0.424	0.413	0.425	0.0120
primary	0.447	0.459	0.446	-0.0120
iliterat	0.128	0.128	0.128	0
rincome	10263	8012	10000	2419.632**
chronics	1.291	1.576	1.269	-0.307***
married	0.911	0.904	0.911	0.00800
divorced	0.0110	0.0140	0.0110	-0.00300
widowed	0.0636	0.0550	0.0640	0.00900
slfhvg	0.116	0.133	0.115	-0.0180
slfhg	0.150	0.148	0.150	0.00200
slfhf	0.589	0.493	0.596	0.103***
slfhp	0.145	0.227	0.139	-0.087***
social	0.579	0.443	0.589	0.146**
CESD	7.098	8.060	7.027	-1.033**
M	3.634	3.273	3.661	0.388***
MI	8.037	7.613	8.068	0.455**
noworking	0.220	0.298	0.215	-0.083***
workagri	0.439	0.465	0.437	-0.0280
workemp	0.238	0.172	0.243	0.071**
workslf	0.103	0.0650	0.105	0.040*
hsptim1y	0.127	0.294	0.114	-0.180***
rwealth	50085	42000	51000	9153
drinker	0.261	0.194	0.266	0.073**
TTFC	2.510	2.869	2.477	-0.392**
T1	0.332	0.295	0.336	0.0410
T2	0.198	0.0820	0.208	0.126**
T3	0.0975	0.0820	0.0990	0.0170
T4	0.372	0.541	0.357	-0.184***
smokesp	0.0939	0.180	0.0870	-0.092***
dprice	1.003	0.170	1.081	0.911***
drincome	392.4	-399.1	451.7	850.8
nchronics	0.0949	0.161	0.0900	-0.071***
accident	0.0531	0.0730	0.0520	-0.0220
newmarried	0.00517	0.00500	0.00500	0.00100

newdiv	0.00485	0.00900	0.00500	-0.00500
newwid	0.0129	0.00900	0.0130	0.00400
slfhbetter	0.0784	0.0800	0.0780	-0.00200
slfhworse	0.437	0.493	0.433	-0.060*
dsocial	0.0264	0.00500	0.0280	0.0230
dCESD	-0.170	0.0950	-0.190	-0.285
dM	-0.118	-0.281	-0.106	0.175
dMI	-0.462	-0.462	-0.462	0
newretire	0.0302	0.0520	0.0290	-0.024*
dhsptim1y	0.0365	0.190	0.0250	-0.164***
drwealth	5814	-3600	6527	10000*
qdrink	0.0788	0.0740	0.0790	0.00500
N	3,096	218	2,878	

Note: Mean values of each variable are presented in this table. In the last row, the number of observations in each group is presented. It should be noted that, the observation number of the variables *smokenum, smokesp,* and the majority of change variables between the lagged and current waves are smaller than 3,096, 218, and 2,878 among the full sample, stoppers, and continuous smokers. *price*, *rincome*, *rwealth*, *dprice*, *drincome*, and *drwealth* are expressed in YUAN. Precise definitions of all variables can be found in Appendix 3.1. Column 4 contains the p-values associated with the t-test for equality of mean percentages among stoppers and continuous smokers. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Table 3.2 Probit and LPM regressions for the probability of smoking cessation

	(1) Probit1	(2) Probit2	(3) LPM1	(4) LPM2
$male_t$	-0.001	-0.001	-0.005	-0.005
	(-0.06)	(-0.06)	(-0.25)	(-0.26)
age_t	0.010	0.010	0.008	0.008
	(1.35)	(1.35)	(1.04)	(1.04)
$age2_t$	-0.000	-0.000	-0.000	-0.000
	(-1.33)	(-1.33)	(-0.97)	(-0.98)
$rural_t$	-0.006	-0.006	-0.010	-0.010
	(-0.58)	(-0.59)	(-0.86)	(-0.87)
$smokenum_{t-1}$	-0.003***	-0.003***	-0.003***	-0.003***
	(-4.67)	(-4.67)	(-4.70)	(-4.64)
$price_{t-1}$	-0.001	-0.001	-0.000	-0.000
	(-0.51)	(-0.51)	(-0.23)	(-0.23)
$smokeage_{t-1}$	-0.001*	-0.001*	-0.001**	-0.001**
	(-1.81)	(-1.81)	(-2.34)	(-2.33)
$above_{t-1}$	0.018	0.018	0.023	0.023
	(0.99)	(0.98)	(1.18)	(1.18)
$primary_{t-1}$	0.017	0.017	0.021	0.020
	(1.04)	(1.02)	(1.21)	(1.20)
$income_{t-1}$	-0.003	-0.003	-0.003	-0.003
	(-1.21)	(-1.21)	(-1.29)	(-1.29)
$chronics_{t-1}$	0.008**	0.008**	0.008**	0.008**
	(2.14)	(2.14)	(2.02)	(2.02)
$married_{t-1}$	-0.054*	-0.054*	-0.086**	-0.086**
	(-1.70)	(-1.70)	(-2.06)	(-2.06)
$divorced_{t-1}$	-0.064	-0.064	-0.088	-0.089
	(-1.15)	(-1.15)	(-1.52)	(-1.54)
$widowed_{t-1}$	-0.070*	-0.070*	-0.105**	-0.105**
	(-1.86)	(-1.87)	(-2.24)	(-2.24)
$slfhvg_{t-1}$	0.009	0.009	0.011	0.011
	(0.58)	(0.57)	(0.72)	(0.71)
$slfhg_{t-1}$	-0.001	-0.001	0.000	0.001
	(-0.05)	(-0.05)	(0.02)	(0.05)
$slfhp_{t-1}$	-0.024*	-0.024*	-0.018	-0.018
	(-1.72)	(-1.71)	(-1.29)	(-1.28)
$social_{t-1}$	-0.015**	-0.015**	-0.011*	-0.011*
	(-2.31)	(-2.32)	(-1.82)	(-1.84)
$CESD_{t-1}$	0.001	0.001	0.001	0.001
	(1.46)	(1.47)	(1.07)	(1.07)
M_{t-1}	-0.001	-0.001	-0.001	-0.001
	(-0.29)	(-0.28)	(-0.24)	(-0.21)
MI_{t-1}	-0.003	-0.003	-0.003	-0.003
	(-1.19)	(-1.18)	(-1.40)	(-1.37)
$workagri_{t-1}$	-0.005	-0.005	-0.002	-0.002
	(-0.39)	(-0.39)	(-0.14)	(-0.14)
$workemp_{t-1}$	0.012	0.012	0.009	0.009
	(0.69)	(0.69)	(0.49)	(0.49)

$worksIf_{t-1}$	-0.022	-0.022	-0.025	-0.026
	(-1.01)	(-1.01)	(-1.16)	(-1.17)
$hsptim1y_{t-1}$	0.009	0.008	0.005	0.005
	(0.75)	(0.73)	(0.40)	(0.37)
$wealth_{t-1}$	-0.002	-0.002	-0.001	-0.001
	(-0.65)	(-0.64)	(-0.44)	(-0.42)
$drinker_{t-1}$	-0.015	-0.015	-0.013	-0.013
	(-1.27)	(-1.27)	(-1.10)	(-1.12)
$TTFC_{t-1}$	0.012***		0.012***	
	(3.00)		(3.05)	
$T1_{t-1}$		-0.037***		-0.037***
		(-2.94)		(-3.04)
$T2_{t-1}$		-0.024*		-0.028**
		(-1.73)		(-2.05)
$T3_{t-1}$		-0.017		-0.026
		(-0.99)		(-1.52)
$\_cons$			-0.009	0.037
			(-0.04)	(0.15)
N	3096	3096	3111	3111
r2_p				

Note: All the columns report the estimates of the probability of smoking cessation as a function of TTFC. Columns 1 and 2 report results from Probit models, whilst columns 3 and 4 report results from LPMs. Columns 1 and 3 use the variable $TTFC_{t-1}$, whilst columns 2 and 4 use the category variables $T1_{t-1}$, $T2_{t-1}$, and $T3_{t-1}$. Marginal effects of Probit models are reported, with *t* statistics in parentheses. *N* denotes the number of observations. The variables $T4_{t-1}$, illiterate, single, fair self-reported health status and unemployed are omitted categories. Precise definitions of all variables can be found in Appendix 3.1. Income and wealth are expressed in logarithms. Province and year dummies are included in all models, but their coefficients/marginal effects are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Table 3.3 Probit and LPM regressions for the probability smoking cessation accounting for differences in relevant variables between waves

	(1) Probit1	(2) Probit2	(3) LPM1	(4) LPM2
$male_t$	-0.018	-0.019	-0.033	-0.035
	(-0.66)	(-0.70)	(-0.96)	(-1.01)
age_t	0.004	0.003	0.002	0.002
	(0.28)	(0.26)	(0.12)	(0.13)
$age2_t$	-0.000	-0.000	-0.000	-0.000
	(-0.32)	(-0.30)	(-0.15)	(-0.16)
$rural_t$	-0.016	-0.017	-0.010	-0.010
	(-0.89)	(-0.91)	(-0.44)	(-0.45)
$smokenum_{t-1}$	-0.003***	-0.003***	-0.003***	-0.003***
	(-3.19)	(-2.90)	(-3.16)	(-3.07)
$price_{t-1}$	-0.008***	-0.008**	-0.007**	-0.007**
	(-2.62)	(-2.40)	(-2.47)	(-2.39)
$smokeage_{t-1}$	0.000	0.000	0.000	0.000
	(0.40)	(0.46)	(0.14)	(0.15)
$above_{t-1}$	0.024	0.022	0.027	0.027
	(0.79)	(0.74)	(0.81)	(0.81)
$primary_{t-1}$	0.017	0.016	0.035	0.035
	(0.65)	(0.62)	(1.20)	(1.18)
$income_{t-1}$	-0.007	-0.007	-0.007	-0.007
	(-1.58)	(-1.56)	(-1.48)	(-1.47)
$chronics_{t-1}$	0.006	0.006	0.009	0.009
	(1.00)	(1.04)	(1.29)	(1.32)
$married_{t-1}$	-0.067	-0.070	-0.117*	-0.120*
	(-1.43)	(-1.44)	(-1.85)	(-1.89)
$divorced_{t-1}$	-0.116	-0.118	-0.128	-0.133
	(-1.12)	(-1.12)	(-1.37)	(-1.41)
$widowed_{t-1}$	-0.070	-0.073	-0.126*	-0.130*
	(-1.26)	(-1.29)	(-1.73)	(-1.78)
$slfhvg_{t-1}$	-0.004	-0.005	0.010	0.009
	(-0.14)	(-0.21)	(0.42)	(0.35)
$slfhg_{t-1}$	-0.013	-0.012	-0.010	-0.009
	(-0.61)	(-0.58)	(-0.52)	(-0.45)
$slfhp_{t-1}$	-0.029	-0.029	-0.016	-0.015
	(-1.37)	(-1.33)	(-0.72)	(-0.67)
$social_{t-1}$	-0.019	-0.019	-0.013	-0.013
	(-1.45)	(-1.44)	(-0.95)	(-1.01)
$CESD_{t-1}$	0.002	0.003	0.001	0.001
	(1.46)	(1.48)	(0.73)	(0.74)
M_{t-1}	0.002	0.002	0.001	0.001
	(0.29)	(0.34)	(0.13)	(0.13)
MI_{t-1}	0.002	0.003	0.001	0.001
	(0.58)	(0.63)	(0.26)	(0.31)
$workagri_{t-1}$	-0.004	-0.003	0.009	0.009
	(-0.16)	(-0.11)	(0.36)	(0.35)
$workemp_{t-1}$	0.019	0.022	0.028	0.029

	(0.68)	(0.78)	(0.96)	(0.97)
$workslf_{t-1}$	-0.002	-0.001	-0.005	-0.004
	(-0.06)	(-0.02)	(-0.14)	(-0.13)
$hsptim1y_{t-1}$	0.037	0.035	0.056**	0.055*
	(1.48)	(1.37)	(2.00)	(1.94)
$wealth_{t-1}$	0.001	0.001	0.004	0.004
	(0.29)	(0.28)	(0.78)	(0.78)
$drinker_{t-1}$	-0.006	-0.006	-0.006	-0.005
	(-0.29)	(-0.29)	(-0.25)	(-0.21)
$TTFC_{t-1}$	0.014**		0.014**	
	(2.28)		(2.16)	
$T1_{t-1}$		-0.047**		-0.046**
		(-2.25)		(-2.36)
$T2_{t-1}$		-0.020		-0.018
		(-0.95)		(-0.85)
$T3_{t-1}$		-0.050*		-0.045*
		(-1.70)		(-1.73)
$dprice_t$	-0.008***	-0.008***	-0.006***	-0.006***
	(-2.87)	(-2.60)	(-2.88)	(-2.83)
$dincome_t$	-0.001	-0.001	-0.000	-0.000
	(-0.69)	(-0.67)	(-0.09)	(-0.08)
$nchronics_t$	0.017	0.016	0.028*	0.027*
	(1.21)	(1.12)	(1.85)	(1.78)
$accident_t$	-0.002	-0.001	0.006	0.006
	(-0.06)	(-0.02)	(0.22)	(0.22)
$newmarried_t$	0.053	0.050	0.064	0.065
	(0.49)	(0.45)	(0.64)	(0.65)
$newdiv_t$	0.079	0.079	0.087	0.089
	(1.00)	(0.98)	(0.96)	(0.98)
$newwid_t$	-0.056	-0.057	-0.038	-0.039
	(-0.67)	(-0.67)	(-0.55)	(-0.56)
$slfhbetter_t$	-0.007	-0.007	-0.010	-0.010
	(-0.24)	(-0.26)	(-0.37)	(-0.37)
$slfhworse_t$	0.013	0.013	0.017	0.017
	(0.85)	(0.85)	(1.09)	(1.09)
$dsocial_t$	-0.013	-0.012	-0.008	-0.008
	(-1.21)	(-1.12)	(-0.82)	(-0.84)
$dCESD_t$	0.002	0.002	0.001	0.001
	(1.18)	(1.18)	(0.56)	(0.57)
dM_t	0.000	0.000	-0.001	-0.001
	(0.07)	(0.08)	(-0.17)	(-0.18)
dMI_t	0.003	0.003	0.002	0.002
	(0.85)	(0.84)	(0.60)	(0.58)
$newretire_t$	0.027	0.028	0.021	0.022
	(0.69)	(0.72)	(0.50)	(0.52)
$dhsptim1y_t$	0.058***	0.058***	0.079***	0.078***
	(4.19)	(3.35)	(4.49)	(4.46)
$dwealth_t$	-0.000	-0.000	0.000	0.000
	(-0.16)	(-0.22)	(0.39)	(0.34)
$qdrink_t$	0.004	0.002	-0.002	-0.003
	(0.13)	(0.08)	(-0.08)	(-0.11)
$\_cons$			0.253	0.299
			(0.55)	(0.65)

N	1548	1548	1549	1549
r2_p				

Note: All the columns report the estimates of the probability of smoking cessation as a function of TTFC. Columns 1 and 2 report results from Probit models, whilst columns 3 and 4 report results from LPMs. Columns 1 and 3 use the variable $TTFC_{t-1}$, whilst columns 2 and 4 use the category variables $T1_{t-1}$, $T2_{t-1}$, and $T3_{t-1}$. Marginal effects of Probit models are reported, with t statistics in parentheses. *N* denotes the number of observations. The variables $T4_{t-1}$, illiterate, single, fair self-reported health status, unemployed, keeping the same marital status, having the same self-reported health status, and keeping the same employment status between the previous and current waves are omitted categories. Precise definitions of all variables can be found in Appendix 3.1. Income, wealth, change in income (*dincome*), and change in wealth (*dwealth*) are expressed in logarithms. Province and year dummies are included in all models, but their coefficients/marginal effects are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Table 3.4 Probit and LPM regressions for the probability of smoking cessation with interaction term between TTFC and a dummy indicating whether the respondent is a heavy smoker

	(1) Probit	(2) LPM
$male_t$	-0.060	-0.010
	(-0.42)	(-0.53)
age_t	0.081	0.008
	(1.26)	(1.06)
$age2_t$	-0.001	-0.000
	(-1.21)	(-0.97)
$rural_t$	-0.054	-0.010
	(-0.58)	(-0.88)
$price_{t-1}$	-0.003	-0.000
	(-0.31)	(-0.18)
$smokeage_{t-1}$	-0.009**	-0.002**
	(-1.97)	(-2.43)
$above_{t-1}$	0.156	0.024
	(1.02)	(1.25)
$primary_{t-1}$	0.130	0.021
	(0.97)	(1.21)
$income_{t-1}$	-0.023	-0.003
	(-1.24)	(-1.28)
$chronics_{t-1}$	0.066**	0.008**
	(2.19)	(2.07)
$married_{t-1}$	-0.424	-0.082*
	(-1.62)	(-1.96)
$divorced_{t-1}$	-0.530	-0.086
	(-1.16)	(-1.48)
$widowed_{t-1}$	-0.575*	-0.101**
	(-1.84)	(-2.17)
$slfhvg_{t-1}$	0.071	0.012
	(0.55)	(0.75)
$slfhg_{t-1}$	-0.014	-0.000
	(-0.13)	(-0.03)
$slfhp_{t-1}$	-0.192*	-0.020
	(-1.67)	(-1.40)
$social_{t-1}$	-0.141**	-0.013**
	(-2.54)	(-2.03)
$CESD_{t-1}$	0.010	0.001
	(1.30)	(1.01)
M_{t-1}	-0.009	-0.001
	(-0.33)	(-0.35)
MI_{t-1}	-0.023	-0.003
	(-1.31)	(-1.49)
$workagri_{t-1}$	-0.051	-0.003
	(-0.45)	(-0.20)
$workemp_{t-1}$	0.094	0.009

	(0.67)	(0.52)
$workslf_{t-1}$	-0.207	-0.028
	(-1.13)	(-1.27)
$hsptim1y_{t-1}$	0.068	0.007
	(0.71)	(0.52)
$wealth_{t-1}$	-0.015	-0.002
	(-0.68)	(-0.52)
$drinker_{t-1}$	-0.134	-0.014
	(-1.39)	(-1.27)
$smoknumD_{t-1}$	0.199	0.033
	(0.90)	(1.30)
$TTFC_{t-1}$	0.224***	0.030***
	(4.10)	(4.60)
$smoknumD_{t-1}\#TTFC_{t-1}$		
	-0.146**	-0.023***
	(-2.14)	(-2.83)
$\_cons$	-3.599*	-0.097
	(-1.78)	(-0.39)
/		
lnsig2u	-3.465	
	(-1.02)	
N	3096	3111
r2_p		

Note: Both columns report the estimates of the probability of smoking cessation with interactions between TTFC and a dummy indicating whether the respondent is a heavy smoker. Column 1 reports the result from Probit model, whilst columns 2 reports the result from LPM. Marginal effects of Probit models are reported, with *t* statistics in parentheses. *N* denotes the number of observations. Illiterate, single, fair self-reported health status and unemployed are omitted categories. Precise definitions of all variables can be found in Appendix 3.1. Income and wealth are expressed in logarithms. Province and year dummies are included in all models, but their coefficients/marginal effects are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Table 3.5 Probit and LPM regressions for the probability of smoking cessation among married respondents

	(1) Probit1	(2) Probit2	(3) LPM1	(4) LPM2
$male_t$	-0.004	-0.005	-0.011	-0.011
	(-0.21)	(-0.22)	(-0.46)	(-0.48)
age_t	0.006	0.006	0.004	0.004
	(0.68)	(0.68)	(0.46)	(0.47)
$age2_t$	-0.000	-0.000	-0.000	-0.000
	(-0.68)	(-0.67)	(-0.40)	(-0.41)
$rural_t$	0.006	0.005	0.002	0.002
	(0.46)	(0.44)	(0.15)	(0.13)
$smokesp_{t-1}$	0.024	0.024	0.028	0.027
	(1.22)	(1.27)	(1.35)	(1.33)
$smokenum_{t-1}$	-0.002***	-0.002***	-0.002***	-0.002***
	(-2.70)	(-2.95)	(-3.82)	(-3.75)
$price_{t-1}$	-0.001	-0.001	-0.001	-0.001
	(-0.75)	(-0.75)	(-0.48)	(-0.51)
$smokeage_{t-1}$	-0.001*	-0.001*	-0.002**	-0.002**
	(-1.90)	(-1.93)	(-2.37)	(-2.37)
$above_{t-1}$	0.024	0.024	0.027	0.027
	(1.10)	(1.13)	(1.31)	(1.30)
$primary_{t-1}$	0.023	0.023	0.026	0.026
	(1.17)	(1.20)	(1.41)	(1.40)
$income_{t-1}$	-0.001	-0.001	-0.001	-0.001
	(-0.36)	(-0.37)	(-0.49)	(-0.47)
$chronics_{t-1}$	0.004	0.004	0.004	0.004
	(0.95)	(0.97)	(0.96)	(0.94)
$slfhvg_{t-1}$	0.004	0.004	0.004	0.003
	(0.27)	(0.25)	(0.23)	(0.20)
$slfhg_{t-1}$	-0.003	-0.003	-0.004	-0.004
	(-0.25)	(-0.25)	(-0.32)	(-0.29)
$slfhp_{t-1}$	-0.032*	-0.032*	-0.025	-0.024
	(-1.66)	(-1.77)	(-1.63)	(-1.63)
$social_{t-1}$	-0.012	-0.012	-0.009	-0.009
	(-1.47)	(-1.57)	(-1.37)	(-1.39)
$CESD_{t-1}$	0.001	0.001	0.001	0.001
	(1.16)	(1.18)	(0.77)	(0.80)
M_{t-1}	-0.001	-0.001	-0.001	-0.001
	(-0.34)	(-0.27)	(-0.32)	(-0.25)
MI_{t-1}	-0.003	-0.003	-0.003	-0.003
	(-1.16)	(-1.17)	(-1.40)	(-1.36)
$workagri_{t-1}$	-0.015	-0.015	-0.011	-0.011
	(-0.96)	(-0.97)	(-0.71)	(-0.71)
$workemp_{t-1}$	-0.005	-0.005	-0.005	-0.005
	(-0.29)	(-0.26)	(-0.26)	(-0.24)
$workslf_{t-1}$	-0.032	-0.032	-0.035	-0.035
	(-1.21)	(-1.27)	(-1.53)	(-1.52)
$hsptim1y_{t-1}$	0.015	0.015	0.015	0.014

	(1)	(2)	(3)	(4)
	(1.15)	(1.15)	(1.06)	(1.02)
$wealth_{t-1}$	-0.002	-0.002	-0.002	-0.002
	(-0.65)	(-0.63)	(-0.59)	(-0.56)
$drinker_{t-1}$	-0.013	-0.014	-0.010	-0.010
	(-1.02)	(-1.08)	(-0.82)	(-0.86)
$TTFC_{t-1}$	0.012**		0.012***	
	(2.00)		(2.88)	
$T1_{t-1}$		-0.036**		-0.037***
		(-2.11)		(-2.86)
$T2_{t-1}$		-0.028*		-0.033**
		(-1.67)		(-2.26)
$T3_{t-1}$		-0.029		-0.033*
		(-1.40)		(-1.79)
$\_cons$			0.019	0.064
			(0.07)	(0.22)
N	2609	2609	2619	2619
r2_p				

Note: All the columns report the estimates of the probability of smoking cessation as a function of TTFC among married respondents. Columns 1 and 2 report results from Probit models, whilst columns 3 and 4 report results from LPMs. Columns 1 and 3 use the variable $TTFC_{t-1}$, whilst columns 2 and 4 use the category variables $T1_{t-1}$, $T2_{t-1}$, and $T3_{t-1}$. Marginal effects of Probit models are reported, with *t* statistics in parentheses. *N* denotes the number of observations. The variables $T4_{t-1}$, illiterate, fair self-reported health status and unemployed are omitted categories. Precise definitions of all variables can be found in Appendix 3.1. Income and wealth are expressed in logarithms. Province and year dummies are included in all models, but their coefficients/marginal effects are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Table 3.6 Probit and LPM models for smoking cessation among married respondents accounting for differences in relevant variables between waves

	(1) Probit1	(2) Probit2	(3) LPM1	(4) LPM2
$male_t$	-0.047	-0.048	-0.070*	-0.070*
	(-1.41)	(-1.37)	(-1.66)	(-1.68)
age_t	0.006	0.007	0.001	0.001
	(0.41)	(0.44)	(0.04)	(0.08)
$age2_t$	-0.000	-0.000	-0.000	-0.000
	(-0.43)	(-0.46)	(-0.08)	(-0.12)
$rural_t$	-0.010	-0.011	-0.000	-0.000
	(-0.50)	(-0.51)	(-0.00)	(-0.01)
$smokesp_{t-1}$	-0.004	-0.001	-0.021	-0.020
	(-0.12)	(-0.03)	(-0.61)	(-0.60)
$smokenum_{t-1}$	-0.002**	-0.002**	-0.002**	-0.002**
	(-2.19)	(-2.04)	(-2.10)	(-2.00)
$price_{t-1}$	-0.007**	-0.007**	-0.006**	-0.006**
	(-2.14)	(-1.99)	(-2.03)	(-1.99)
$smokeage_{t-1}$	0.000	0.000	0.000	0.000
	(0.20)	(0.25)	(0.31)	(0.31)
$above_{t-1}$	0.024	0.022	0.027	0.027
	(0.69)	(0.65)	(0.75)	(0.75)
$primary_{t-1}$	0.009	0.008	0.027	0.027
	(0.30)	(0.27)	(0.84)	(0.82)
$income_{t-1}$	-0.003	-0.004	-0.003	-0.003
	(-0.73)	(-0.76)	(-0.61)	(-0.58)
$chronics_{t-1}$	0.000	0.000	0.002	0.002
	(0.01)	(0.07)	(0.29)	(0.29)
$slfhvg_{t-1}$	-0.026	-0.028	-0.011	-0.013
	(-0.90)	(-0.95)	(-0.44)	(-0.50)
$slfhg_{t-1}$	-0.020	-0.018	-0.021	-0.020
	(-0.84)	(-0.76)	(-1.01)	(-0.94)
$slfhp_{t-1}$	-0.048*	-0.048*	-0.021	-0.021
	(-1.91)	(-1.79)	(-0.93)	(-0.90)
$social_{t-1}$	-0.020	-0.021	-0.014	-0.015
	(-1.42)	(-1.42)	(-1.02)	(-1.05)
$CESD_{t-1}$	0.003	0.003	0.002	0.002
	(1.51)	(1.52)	(0.85)	(0.87)
M_{t-1}	0.000	0.001	0.003	0.003
	(0.07)	(0.13)	(0.44)	(0.46)
MI_{t-1}	0.003	0.003	0.000	0.001
	(0.66)	(0.70)	(0.08)	(0.11)
$workagri_{t-1}$	-0.026	-0.025	-0.017	-0.017
	(-1.01)	(-0.97)	(-0.65)	(-0.65)
$workemp_{t-1}$	-0.020	-0.017	-0.013	-0.013
	(-0.64)	(-0.54)	(-0.43)	(-0.42)
$workslf_{t-1}$	-0.008	-0.007	-0.018	-0.018
	(-0.21)	(-0.19)	(-0.51)	(-0.51)
$hsptim1y_{t-1}$	0.040	0.036	0.083***	0.080***

	(1.35)	(1.20)	(2.70)	(2.62)
$wealth_{t-1}$	-0.005	-0.005	-0.003	-0.003
	(-0.84)	(-0.87)	(-0.45)	(-0.46)
$drinker_{t-1}$	-0.009	-0.010	-0.010	-0.010
	(-0.38)	(-0.42)	(-0.43)	(-0.42)
$TTFC_{t-1}$	0.016**		0.015**	
	(2.30)		(2.27)	
$T1_{t-1}$		-0.051**		-0.049**
		(-2.18)		(-2.41)
$T2_{t-1}$		-0.029		-0.029
		(-1.22)		(-1.28)
$T3_{t-1}$		-0.061*		-0.045*
		(-1.77)		(-1.68)
$dprice_t$	-0.006**	-0.006**	-0.004**	-0.004**
	(-2.19)	(-2.00)	(-2.06)	(-1.99)
$dincome_t$	-0.001	-0.000	0.000	0.000
	(-0.48)	(-0.46)	(0.22)	(0.24)
$nchronics_t$	-0.001	-0.002	-0.006	-0.007
	(-0.06)	(-0.10)	(-0.39)	(-0.45)
$accident_t$	-0.044	-0.043	-0.022	-0.021
	(-1.20)	(-1.16)	(-0.76)	(-0.73)
$newdiv_t$	0.000	0.000	-0.048	-0.049
	(.)	(.)	(-0.40)	(-0.40)
$newwid_t$	0.000	0.000	-0.117	-0.117
	(.)	(.)	(-1.40)	(-1.40)
$slfhbetter_t$	0.002	0.000	-0.006	-0.005
	(0.05)	(0.01)	(-0.20)	(-0.17)
$slfhworse_t$	0.008	0.008	0.016	0.016
	(0.45)	(0.44)	(0.98)	(0.98)
$dsocial_t$	-0.017	-0.017	-0.011	-0.011
	(-1.44)	(-1.33)	(-1.10)	(-1.08)
$dCESD_t$	0.003	0.003	0.002	0.002
	(1.57)	(1.58)	(1.32)	(1.32)
dM_t	0.000	0.000	0.000	0.000
	(0.02)	(0.01)	(0.05)	(0.05)
dMI_t	0.004	0.004	0.001	0.001
	(1.01)	(1.01)	(0.37)	(0.35)
$newretire_t$	0.036	0.036	0.039	0.038
	(0.87)	(0.86)	(0.89)	(0.88)
$dhsptim1y_t$	0.071***	0.070***	0.100***	0.099***
	(4.59)	(3.42)	(5.24)	(5.20)
$dwealth_t$	-0.000	-0.000	0.000	0.000
	(-0.26)	(-0.32)	(0.27)	(0.22)
$qdrink_t$	0.015	0.013	0.011	0.010
	(0.46)	(0.40)	(0.35)	(0.30)
$\_cons$			0.242	0.275
			(0.44)	(0.50)
N	1220	1220	1295	1295
r2_p				

Note: All the columns report the estimates of the probability of smoking cessation as a function of TTFC among married respondents. Columns 1 and 2 report results from Probit models, whilst columns 3 and 4 report results from LPMs. Columns 1 and 3 use the variable $TTFC_{t-1}$, whilst columns 2 and

4 use the category variables $T1_{t-1}$, $T2_{t-1}$, and $T3_{t-1}$. Marginal effects of Probit models are reported, with *t* statistics in parentheses. *N* denotes the number of observations. The variables $T4_{t-1}$, illiterate, fair self-reported health status, unemployed, keeping the same marital status, having the same self-reported health status, and keeping the same employment status between the previous and current waves are omitted categories. Precise definitions of all variables can be found in Appendix 3.1. Income, wealth, change in income (*dincome*), and change in wealth (*dwealth*) are expressed in logarithms. Province and year dummies are included in all models, but their coefficients/marginal effects are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Appendices

This section includes Appendices 3.1 and 3.2. Specifically, Appendix 3.1 describes the definitions of all variables used in this chapter. Appendix 3.2 presents the descriptive statistics and empirical results of an analysis which estimates the impact of the spousal smoking status on the number of years the respondent is observed smoking.

Appendix 3.1 Definitions of all variables

Variable	Definition
smoke_stop	Dummy variable equal to 1 if the respondent is a smoker in the previous wave and a non-smoker in the current wave, and 0 if the respondent is a smoker in both the previous and current waves
male	Dummy variable equal to 1 if the respondent is male, and 0 otherwise
age	Continuous variable representing the respondent's age
age2	Age squared
rural	Dummy variable equal to 1 if the respondent is living in a rural area, and 0 otherwise
smokenum	Continuous variable representing the number of cigarettes smoked
price	Continuous variable representing the price of a pack of cigarettes
smokeage	Continuous variable representing the duration of smoking
iliterat	Dummy variable equal to 1 if the respondent is illiterate, and 0 otherwise
primary	Dummy variable equal to 1 if the respondent completes primary education, and 0 otherwise

above	Dummy variable equal to 1 if the respondent completes secondary school or a higher level of education, and 0 otherwise
rincome	Continuous variable representing the annual per capita income of the household. This includes the main respondent's and spouse's wage income, fringe benefits, unpaid family business income, pension insurance income, net agricultural income, net self-employed business income, household public transfer income, household capital asset income, and other household members' income
income	Continuous variable representing the natural logarithm of *rincome*
chronics	Continuous variable representing the total number of diseases the respondent has, including hypertension, dyslipidaemia, diabetes, cancer, chronic lung diseases, liver disease, heart problems, stroke, kidney disease, stomach disease, emotional problems, memory-related disease, arthritis, and asthma
married	Dummy variable equal to 1 if the respondent is married, and 0 otherwise
divorced	Dummy variable equal to 1 if the respondent is divorced, and 0 otherwise
widowed	Dummy variable equal to 1 if the respondent is widowed, and 0 otherwise
slfhvg	Dummy variable equal to 1 if the respondent's health status is very good, and 0 otherwise
slfhg	Dummy variable equal to 1 if the respondent's health status is good, and 0 otherwise

slfhf	Dummy variable equal to 1 if the respondent's health status is fair, and 0 otherwise
slfhp	Dummy variable equal to 1 if the respondent's health status is poor, and 0 otherwise
social	Continuous variable representing the sum of social activities conducted by the respondent in the last month of the wave, including interacting with friends, playing cards, providing help to family or friends, going to a park, taking part in a community-related organisation, doing voluntary work, caring for a sick adult who does not live with them, attending an educational course, investing in stocks, using the internet, and other activities
CESD	Continuous variable based on the sum of the answers given to 10 questions from the Center for Epidemiologic Studies website. Higher scores mean higher levels of depression
M	Continuous variable averaging the immediate memory and delayed memory
MI	Continuous variable summing the correct answers from 11 questions. These questions include successively subtracting 7 from 100 (up to five times), awareness of the date, month, year (using either solar or lunar calendar), the day of the week and season of the year, and the ability to redraw a picture of two overlapping pentagons
noworking	Dummy variable equal to 1 if the respondent does not work, and 0 otherwise

workagri	Dummy variable equal to 1 if the respondent is engaging in agricultural work, and 0 otherwise
workemp	Dummy variable equal to 1 if the respondent is engaging in employed work, and 0 otherwise
workslf	Dummy variable equal to 1 if the respondent is engaging in self-employed work, and 0 otherwise
hsptim1y	Continuous variable representing how many times the respondents received inpatient care in the past year of the wave
rwealth	Continuous variable representing the amount of per capita wealth. Wealth is defined by the sum of household assets, including present market value of the house, vehicles, fixed capital assets (such as tractor, refrigerator, washing machine, computer, valuable decorations, treasures, and so on), annual rent income of the land, current value of all livestock, and financial assets (such as cash, deposit, stocks, mutual funds, bonds, public housing fund, jizikuan, unpaid salary)
wealth	Continuous variable representing the natural logarithm of *rwealth*
drinker	Dummy variable equal to 1 if the respondent is a heavy drinker, and 0 otherwise
TTFC	Categorical variable representing how soon after the respondent wakes up he/she smokes his/her first cigarette. It equals to 1 if the first cigarette is smoked within 5 minutes; 2 if it is smoked within 6-30 minutes; 3 if it is smoked within 31-60 minutes; 4 if it is smoked after more than 1 hour

T1	Dummy variable equal to 1 if the respondent smokes his/her first cigarette within 5 minutes after waking up, and 0 otherwise
T2	Dummy variable equal to 1 if the respondent smokes his/her first cigarette within 6-30 minutes after waking up, and 0 otherwise
T3	Dummy variable equal to 1 if the respondent smokes his/her first cigarette within 31-60 minutes after waking up, and 0 otherwise
T4	Dummy variable equal to 1 if the respondent smokes his/her first cigarette more than 1 hour after waking up, and 0 otherwise
smokesp	Dummy variable equal to 1 if the respondent' spouse is smoking, and 0 otherwise
dprice	Continuous variable representing the difference in the price per pack of cigarettes between the current and the previous waves
dM	Continuous variable representing the difference in memorial scores between the current and the previous waves
dMI	Continuous variable representing the difference in mental intactness scores between the current and the previous waves
drincome	Continuous variable representing the difference in the real amount of per capita income between the current and the previous waves
dincome	Continuous variable representing the natural logarithm of *drincome*
nchronics	Continuous variable representing the changes in the numbers of chronic diseases between the current and the previous waves
accident	Dummy variable equal to 1 if the respondent has been in a traffic accident since last wave, and 0 otherwise

newmarried	Dummy variable equal to 1 if the respondent got married since last wave, and 0 otherwise
newdiv	Dummy variable equal to 1 if the respondent got divorced since last wave, and 0 otherwise
newwid	Dummy variable equal to 1 if the respondent got widowed since last wave, and 0 otherwise
slfhbetter	Dummy variable equal to 1 if the respondent says his/her health is better compared with his/her health in the last interview, and 0 otherwise
slfhworse	Dummy variable equal to 1 if the respondent says his/her health is worse compared with his/her health in the last interview, and 0 otherwise
dsocial	Continuous variable representing the difference in the number of frequent social activities between the current and the previous waves
dCESD	Continuous variable representing the changes in the depression score between the current and the previous waves
newretire	Dummy variable equal to 1 if the respondent retired since last interview, and 0 otherwise
dhsptim1y	Continuous variable representing the change in the times the respondents received inpatient care between the current and previous waves
drwealth	Continuous variable representing the difference in the real amount of wealth per capita between the current and the previous waves
dwealth	Continuous variable representing the natural logarithm of *drwealth*
qdrink	Dummy variable equal to 1 if the respondent quit heavy drinking between the previous and current waves, and 0 otherwise

obwaves	Continuous variable representing the number of waves the respondent is observed smoking over the three waves of the CHARLS. It includes 0, 1, 2, and 3

Appendix 3.2 Negative binomial regression model

Table A in Appendix 3.2 presents descriptive statistics of the variables used in the regressions contained in Table B in Appendix 3.2. The total number of observations is 8,432. This number is higher than the corresponding number (3,096) shown in Table 3.1, because in Table 3.1, the observed respondents are restricted to the respondents who are current smokers in the last wave, while Appendix 3.2 considers all the married respondents who are observed in the CHARLS.

The prevalence of some variables is similar between Table 3.1 and Table A in Appendix 3.2, including the rural status (*rural*), number of chronic diseases (*chronics*), memory status (*M*), mental intactness (*MI*), and number of inpatient cares in the past year (*hsptim1y*).

However, there are also differences in the prevalence of variables in tables mentioned above. In Table A in Appendix 3.2, the average number of waves the respondent is observed smoking is 0.881, which is not presented in Table 3.1. The average age is 58.11, compared with 60.02 in Table 3.1. It means the sample in Table A in Appendix 3.2 is around two years younger than the corresponding sample in Table 3.1. This decrease in the average age can be caused by dropping those widowed elderly respondents. Male respondents and female respondents are equally observed. By contrast, 91 percent of the respondents are males in Table 3.1. This difference can be explained by the fact that Appendix 3.2 only focuses on married respondents, while Table 3.1 focuses on the respondents who are current smokers in the last wave. Based on the same reason, Table

A in Appendix 3.2 shows that around 28 percent of respondents' spouses are current smokers, which is not presented in Table 3.1. As for employment status, 25.3 percent of the respondents are unemployed or retired. 62.3 percent of the respondents are working in agriculture. 7.71 percent of them are employees, and 4.67 percent are self-employers. These results imply that more people are not working or working in agriculture compared with the main sample in Table 3.1. 10.7 percent of the respondents surveyed are heavy drinkers, while 26.1 percent are heavy drinkers in the main sample. This finding means that married people behave healthier in terms of alcohol consumption.

Moreover, there are several aspects that married respondents behave worse than the respondents in Table 3.1. For example, in Table A in Appendix 3.2, the percentage of education at illiterate level, primary school level, and secondary school or above level are 24.2, 41.1, and 34.8 percent, respectively. These figures show that the educational level among married respondents is lower than that in the main sample. Also, the average annual income and wealth per capita are lower at 9,426 and 44,530 YUAN. 7.66 and 18.2 percent of the respondents have very good and good self-rated health status, respectively, while 51.9 percent of the respondents consider their health status as fair, and 22.2 percent as poor. These results mean that married respondent's self-reported health status is not as good as that in Table 3.1. The mean number of social activities undertaken by the respondents is lower at 0.475. Furthermore, the mean value of CESD is 7.993, compared with 7.098 in Table 3.1. The higher value of CESD represents higher depression level among married respondents.

Although there are differences between Table 3.1 and Table A in Appendix 3.2, the results of the negative binomial regression model are not biased. The main reason for the differences is the different populations. Appendix 3.2 studies married respondents, while Table 3.1 studies the respondents who are current smokers in the last wave. The negative binomial regression model adds a temporal dimension to the smoking habits and estimates the spousal effect on the number of years the person is observed smoking, which is not presented in the main analyses.

Table B in Appendix 3.2 shows the results of the negative binomial regression model, which estimates the effect of the spousal smoking status on the number of waves the respondents are observed smoking over the three waves of datasets. The observed values of the dependent variable *obwaves* are 0, 1, 2, and 3, which means that the individual can be observed as a smoker for 0, 1, 2, and 3 waves, respectively. The regression includes values of the independent variables in the first wave.

The negative binomial regression model is used, because this is an appropriate model for modelling count data in the presence of overdispersion. Table A in Appendix 3.2 shows that the variance (1.315^2=1.729225) of the dependent variable *obwaves* is higher than its mean value (0.881). This means that the variable *obwaves* shows overdispersion. Moreover, the overdispersion parameter */lnalpha* displayed in Table B in Appendix 3.2 shows that the alpha test is significant at 1% level. This test further proves the existence of overdispersion. Therefore, the negative binomial model is more appropriate than the Poisson model in this study.

Table B in Appendix 3.2 reports the marginal effects. The spouse's smoking status is positively and statistically significantly associated with the number of waves the respondents are observed smoking. The result suggests that if the respondent's spouse was a current smoker, then the number of waves the respondent was observed smoking would be 0.141 higher, holding all other control variables constant. This can be explained considering that people's smoking behaviours are more likely to be affected by their family members, and one smoking person in a family will increase the probability of smoking by the other family members in the current or future waves (Li *et al*., 2013; Stimpson *et al*., 2006; Jurj *et al*., 2006). However, stopping smoking needs firm determination and perseverance, and peer effect from spouse may only play a limited impact on successful smoking cessation. Therefore, Table 3.5 and 3.6 do not show a significant association between spousal smoking status and smoking cessation.

Consistent with the literature (Lee *et al*., 2016; Chan and Leung, 2015; Cui *et al*., 2012; Chou, 2008; Abdullah *et al*., 2006), in Table B in Appendix 3.2, depressed older male respondents who drink heavily are more likely to be observed smoking in several waves, while those who have chronic diseases are less likely to smoke in several waves. However, some variables have inconsistent results with previous studies (Chan and Leung, 2015; Gao *et al*., 2013). Specifically, the number of social activities and the people who work in agriculture are associated with a higher number of waves observed smoking. The possible explanation is that the Chinese cigarette culture promotes the prevalence of smoking in social networking. Also, those respondents who are smokers are more likely to build friendship with smokers, therefore, more social activities between smokers contribute to the continuation of smoking behaviour. As for the significantly positive association between working in agriculture and the number of waves observed smoking, it can be explained considering that agricultural job is quite heavy and stressful (Wang *et al*., 2018), and farmers are likely to be influenced by other farmers' smoking habits. Therefore, smoking is prevalent in agriculture.

Table A. Descriptive statistics for married respondents

Variables	mean	sd	min	max	N
obwaves	0.881	1.315	0	3	8432
male	0.502	0.500	0	1	8432
age	58.11	8.356	45	92	8432
rural	0.638	0.481	0	1	8432
smokesp	0.279	0.448	0	1	8432
above	0.348	0.476	0	1	8432
primary	0.411	0.492	0	1	8432
iliterat	0.242	0.428	0	1	8432
rincome	9426	11996	0	179400	8432
chronics	1.364	1.387	0	10	8432
slfhvg	0.0766	0.266	0	1	8432
slfhg	0.182	0.386	0	1	8432
slfhf	0.519	0.500	0	1	8432
slfhp	0.222	0.416	0	1	8432
social	0.475	0.734	0	7	8432
CESD	7.993	6.166	0	30	8432
M	3.629	1.642	0	9.500	8432
MI	7.903	2.890	0	11	8432
noworking	0.253	0.435	0	1	8432
workagri	0.623	0.485	0	1	8432
workemp	0.0771	0.267	0	1	8432
workslf	0.0467	0.211	0	1	8432
hsptim1y	0.109	0.386	0	3	8432
rwealth	44530	70987	0	898000	8432
drinker	0.107	0.309	0	1	8432

Note: Mean values, standard deviations (*sd*), minimum values (*min*), maximum values (*max*), and number of observations (*N*) of each variable are presented in this table. *rincome* and *rwealth* are expressed in YUAN. Precise definitions of all variables can be found in Appendix 3.1.

Table B. Negative binomial regression for the number of waves the respondents are observed smoking among married respondents

	obwaves
$male_1$	2.360***
	(41.18)
age_1	0.054**
	(2.17)
$age2_1$	-0.000**
	(-2.28)
$rural_1$	0.064
	(1.51)
$smokesp_1$	0.141**
	(2.39)
$above_1$	-0.048
	(-0.74)
$primary_1$	0.040
	(0.70)
$income_1$	-0.007
	(-0.75)
$chronics_1$	-0.054***
	(-3.79)
$slfhvg_1$	0.077
	(1.22)
$slfhg_1$	0.020
	(0.44)
$slfhp_1$	0.062
	(1.25)
$social_1$	0.053**
	(2.22)
$CESD_1$	0.009***
	(2.62)
M_1	0.003
	(0.24)
MI_1	-0.010
	(-1.24)
$workagri_1$	0.162***
	(3.23)
$workemp_1$	0.079
	(1.09)
$workslf_1$	0.034
	(0.38)
$hsptim1y_1$	-0.073
	(-1.58)
$wealth_1$	-0.009
	(-0.75)
$drinker_1$	0.261***
	(5.74)
$\_cons$	-3.640***

	(-4.78)
/lnalpha	-0.265***
	(-4.29)
N	8432
r2_p	0.136

Note: This table reports the estimate of the number of waves the respondents are observed smoking as a function of the spousal smoking status among married respondents. Marginal effects of the negative binomial regression model are reported, with *t* statistics in parentheses. *N* denotes the number of observations. Illiterate educational level, fair self-reported health status and unemployment are omitted categories. Precise definitions of all variables can be found in Appendix 3.1. Income and wealth are expressed in logarithms. Province dummies are included in the model, but their marginal effects are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Chapter Four: Time Discounting and Obesity: Evidence from the China Family Panel Studies

4.1 Introduction

In recent years, overweight and obesity have become significant concerns. According to the World Health Organisation (WHO, 2018), worldwide obesity has nearly tripled from 1975 to 2016. Specifically, in 2016, there were more than 1.9 billion overweight adults, and over 650 million obese adults. The prevalence of overweight was 39 percent, and that of obesity was 13 percent. As a result, the prevalence of overweight/obesity has exceeded that of underweight, and is related to more deaths worldwide. Although a large number of existing studies are focused on developed countries (e.g. Baum and Ruhm, 2009; Pouliou and Elliott, 2010), the prevalence of both overweight and obesity has increased significantly over the past decades in developing countries as well, particularly in China. According to data from eight waves of the China Health and Nutrition Survey (CHNS) over the period 1991-2011, the prevalence of obesity increased from 2.88% to 11.8% among Chinese men, and from 4.55% to 11.0% among women. Similarly, the prevalence of overweight and obesity combined has risen from 16.5% to 46.1% among men, and from 24.3% to 39.3% among women during this same period (Mi *et al*., 2015).

Obesity is a major risk factor for people's health, because it is associated with a group of non-communicable diseases, such as cardiovascular diseases, diabetes, musculoskeletal disorders, and some cancers (WHO, 2018). Obesity also causes a high level of stress on medical care services. Using longitudinal data from the 2000–2009 waves of the CHNS, Qin and Pan (2016) provide a nationally representative estimate of the medical cost attributable to obesity and overweight in China. They find that, focusing on a single medical event (illness or injury), overweight and obese people have to pay an additional 5.3% of their total expected spending on medical care as ‘penalties’, compared to their counterparts with healthy weight. This is because obesity is not only a medical condition by itself but is also linked to increased risks of a variety of acute and chronic diseases. As a result, a large amount of medical costs is found to be associated with obesity. This translates into an annual cost of 24.35 billion Yuan on the national scale, accounting for 2.46% of China's national health care expenditure.

Many studies have looked at the determinants of obesity in developing countries. A common finding is that the more affluent tend to be more likely to be obese in low-income and middle-income countries, whilst poor people are more likely to be obese in high-income countries (Swinburn *et al*., 2011). Ford *et al*. (2017) report that changes in the global diet and declining physical activity are the leading contributors to obesity in developing countries. Processed foods, snacks, soft drinks, and alcoholic beverages are important sources of caloric intake in many parts of the developing world. Ford *et al*. (2017) also suggest that emerging risk factors such as chronic stress associated with poverty, unemployment, crime, lack of safety, and poor social networks could be contributing to rising obesity in low- and middle-income countries.

Other studies have identified high time discounting as a determinant of obesity in developed countries, such as the United States, Australia, the United Kingdom, and the Netherlands. They find that high time preferences contribute to increasing body weight, higher body mass index (BMI)[10], and a higher probability of being obese (Smith *et al*., 2005; Komlos *et al*., 2004; Zhang and Rashad, 2008; Chabris *et al*., 2008; Courtemanche *et al*., 2015; Dodd, 2014; Brown and Biosca, 2016; Borghans and Golsteyn, 2006). A similar positive association has been found in a developing county, namely Cambodia (Rieger, 2015).

With reference to China, existing studies found that age, socioeconomic status, living in an urban area, gender, marital status, occupation, depression, smoking, drinking, physical activity, and self-reported health are key determinants of obesity (Li *et al*., 2012; Gao and Shen, 2017; Du *et al*., 2004; Zhang *et al*., 2008; Mi *et al*., 2015; Yu *et al*., 2011; Zhang *et al*., 2016; Wang, 2015, Ng *et al*., 2010; Zhang and Rashad, 2008;). However, to the best of our knowledge, the role of time preferences in explaining obesity in China has not been investigated with national representative datasets. We fill this gap in the literature by investigating the extent to which high time preferences may be associated with the probability of being overweight or obese, and BMI.

Using the 2014 wave of the the China Family Panel Studies (CFPS), we find that time preferences are positively associated with the probability of being overweight, being obese and BMI. Additionally, the association between time preferences and the probability of overweight is higher than the corresponding relationship with obesity. Our results suggest that policies aimed at educating people to be more far-sighted could contribute to reducing the prevalence of overweight and obesity in China.

[10] BMI is defined as weight in kilograms over the square of height in meters.

The remainder of the chapter is organised as follows. Section 4.2 summarises the existing related studies. Section 4.3 introduces our dataset. Section 4.4 presents our hypotheses. Section 4.5 illustrates our specification and estimation methodology. Section 4.6 presents summary statistics. Section 4.7 presents the empirical results, and section 4.8 concludes.

4.2 Literature review

4.2.1 Determinants of obesity

4.2.1.1 Studies on countries other than China

A large number of studies have explored the determinants of obesity and BMI worldwide. For instance, Baum and Ruhm (2009) investigate the predictors of BMI and obesity among the youth in the US considering age and socio-economics status (SES). Using panel data from the National Longitudinal Survey of Youth (NLSY) for the years 1981-2004, 70,908 observations are included in their analyses. Years of education and family income are used as proxies of SES. After controlling for gender, race/ethnicity, marital status, number of children, alcohol consumption, cigarette smoking, physical activity, and job-related physical demands, they show that both BMI and the prevalence of obesity are positively associated with age, whilst are negatively associated with SES.

Similarly, Sobal and Stunkard (1989) and McLaren (2007) study the relationship between social standard and body weight, and suggest that this relationship varies in accordance with the level of economics development of the country concerned. In other words, there is a positive relationship between SES and body weight in developing countries among all population, while the corresponding relationship changes to be negative in developed countries among females, and insignificant among males. Additionally, the body weight difference among women in higher- versus lower-development countries became smaller in 2004 compared with sixteen years ago. The possible reason is the increase in globalisation that reduces the difference between developing countries and developed countries gradually.

Additionally, Pouliou and Elliott (2010) explore the individual and socio-environmental determinants of overweight and obesity in urban Canada. Based on the dataset from the 2003 Canadian Community Health Survey as well as spatial information from the Desktop Mapping Technologies Incorporation, a total of 115,548 participants aged 20 or older are included in the analyses. They investigate the individual determinants of body weight, including age, gender, marital status, whether the respondent is a recent immigrant, race, chronic disease status, education, income, employment status, home ownership, smoking, drinking, physical activity, and fruits and vegetable consumption. Also, the effects of broader environments on body weight are estimated, such as geographic and social environments. The authors report that age, male, and chronic condition are significantly positively associated with BMI, whilst Asian race, educational level, smoking status, and energy expenditure are significantly negatively associated with BMI. Furthermore, Pouliou and Elliott (2010) demonstrate the important role of built environment in overweight and obesity in Urban Canada, and show that a walkable built environment is negatively related to body weight. Considering on the combined information on residential density, street network connectivity and distribution of several land-use types, including residential, commercial, industrial, institutional and open space, people are more likely to walk in a walkable built environment instead of driving or taking public transpotations. This may be the reason why the people's BMI values are lower in this area.

Yu *et al*. (2010) perform a systematic review to investigate the association between childhood intelligence and adulthood obesity. Using meta-analysis, Yu *et al*. (2010) organise 883 papers and find that childhood intelligence is negatively associated with adulthood obesity and BMI. It is plausible that children with higher intelligence scores are more likely to have healthy behaviours, such as frequent intake of fruits and vegetables or low intake of high calorie foods.

Also, Plurphanswat and Rodu (2014) explore the relationship between smoking status and BMI according to gender in the US. The dataset consists of 11,123 men and 10,949 women aged 25 to 64 from 1990 to 2012, and is taken from the National Health and Nutrition Examination Surveys (NHANES). They suggest that current smokers tend to have lower BMI than people who never smoked. They also tend to have higher probabilities of being underweight and having normal weight, but lower probabilities of being obesity. These findings apply to both men and women. The plausible explanation for the negative relationship between smoking and BMI is that cigarettes may suppress appetite and increase metabolism. As for demographic characteristics, older people are found to have higher BMI, whilst education beyond high school is associated with lower BMI among both men and women. Married men have higher BMI, but married women have low BMI.

4.2.1.2 Studies on China

In China, an increasing number of studies have been conducted to investigate the determinants of obesity. In terms of demographic characteristics, significant disparities of body weight have been found by gender, age, and whether the respondent is living in urban or rural areas. Using the 2011 wave of the CHNS coving 12,249 Chinese adults aged 20 and above, Mi *et al*. (2015) indicate that big body sizes are more prevalent among males than females, based on the Chinese criteria[11], of relative body weight. Specifically, the combined prevalence of overweight and obesity is 46.1% among men and 39.3% among women. Similarly, the percentage of obesity is 11.8% among men and 11.0% among women. Also, the mean value of BMI among men is 23.8, compared with 23.4 among women. Li *et al*. (2012) make use of a cohort of 98,271 adults aged 18 and above from 31 provinces in mainland China. Using the Chinese criteria to define overweight and obesity, a big difference in body weight is found between urban residents and rural residents. The combined prevalence of overweight and obesity are 48.1% and 41.1% in urban and rural areas, respectively. Li *et al*. (2012) also find an inverted U-shaped relationship between age and obesity.

According to the literature (Sobal and Stunkard, 1989; McLaren, 2007), SES is considered to be an important positive predictor of body weight in China. Using the 1989, 1991, 1993, and 1997 waves of the CHNS with 5,783 subjects aged 20 to 45 from 3,129

[11] According to the Chinese criteria, respondents are defined as being obese if their BMI is 28 kg/m^2 or greater, whilst they are defined as being overweight if their BMI is greater or equal to 24 kg/m^2, and is smaller than 28 kg/m^2.

household, Du *et al*. (2004) find that income is positively associated with body weight, because the combined prevalence of overweight and obesity in the high-income tertile is about twofold higher than its counterpart in the low income tertile. For example, in 1997, the corresponding prevalences in the high- and low- income groups were 19.6% and 10.9%, respectively. The possible reason for this phenomenon is that China is experiencing a nutrition transition, in which people with high income are more likely to relinquish traditional staple foods of rice, wheat and wheat products, and consume more animal foods and processed foods that include higher energy and fat intakes. This finding is confirmed by Zhang *et al*. (2008) with data from the China National Nutrition and Health Survey (CNNHS) conducted in 2002. 56,442 Chinese consumers above the age of 18 years old are investigated. The authors find that the combined prevalence of overweight and obesity is higher among those Chinese who are living in urban areas, have higher educational level, and have higher income, because they are living in an environment in which a variety of energy-dense food products are supplied and promoted, and these consumers tend to adopt higher calorie diets.

Other studies reach different conclusions. Using panel data from CHNS over the period between 1989 and 2006 with 7,314 female and 6,492 male respondents, Jones-Smith *et al*. (2012) find that the prevalence of overweight doubles among female respondents, and triples among male respondents over seventeen years. Using educational attainment as a proxy of SES, they indicate that educated women have a lower level of overweight and BMI. By contrast, educated men have a higher level of overweight and BMI. Similarly, Xiao *et al*. (2013) estimate the association between SES and obesity in the Zhejiang province in China. With data from the 2010 Zhejiang metabolic syndrome prevalence survey, which includes 16,013 respondents aged 18 or older, they use education and income as indicators of SES, and use BMI and waist circumference (WC) as the measurements of general obesity and abdominal obesity, respectively. They find significant disparities in the SES-obesity association between males and females when measured by different SES indicators. Specifically, obesity is positively associated with income among men, whist it is negatively associated with education among women. To sum up, the findings indicate that the development of Chinese economy accelerates the transition of the SES-obesity association from the 'developing country pattern' towards the 'developed country pattern', and females show faster transitions than males. Also, the tolerance of male obesity is high because of Chinese culture, where, to some extent, a large body size among men is a sign of physical dominance and prowess. By contrast, the tolerance of females is low, and a thinner figure is socially valued. Therefore, women are more concerned about their physical appearance.

A large number of Chinese adults suffer from depressive symptoms recently (Kang *et al*., 2015). Therefore, many studies in China are conducted to estimate the association between depression and obesity. For example, Yu *et al*. (2011) use data from the Health Promotion Knowledge, Attitudes, and Performance Survey in 2002, and investigate the association between depression and BMI by gender among 20,385 Taiwanese adults aged between 18 and 64. Depression is defined with the Taiwanese Depression Questionnaire with the cut-off point at 19. Also, Chinese criteria are used to categorise weight groups, namely, underweight, normal weight, overweight, and obesity. The results suggest that men with depressive symptoms are likely to be underweights rather than show normal weight, whilst depressed women are less likely to be obese. The possible reason is that depressed people may have poor appetite, and thus suffer from a loss of body weight. Males are more responsive to the depression-obesity association than females. This result is confirmed by Zhang *et al*. (2016), who describe the association between BMI and depression among middle-aged and elderly people in China. They focus on 6,224 males and 6,883 females taken from the 2011 wave of the China Health and Retirement Longitudinal Study (CHARLS). The Chinese version of the Center for Epidemiologic Studies Depression scale (CESD) is used to measure depressive symptoms. A negative relationship is revealed between depression and the probability of being overweight and obese, while a positive relationship is found between depression and underweight. This association remains after controlling for age, education, marital status, hukou, whether respondents live in rural or urban areas, smoking habits, alcohol consumption, average hours of sleep per night, eating habits, social activity status, major accidental injury, chronic diseases, physical activity, and self-reported health status.

Moreover, lifestyles are considered as determinants of obesity. Consistent with the study by Plurphanswat and Rodu (2014) in the US, Wang (2015) uses the 1991 to 2011 waves of CHNS from China, and estimates the effect of cigarette smoking on body weight. In the 2011 wave, 19,000 observations from 4,400 households are covered in the survey. The author indicates that controlling for SES as well as demographic characteristics, smoking increases the likelihood of being underweight or having normal weight, but decreases the likelihood of overweight and obesity. A plausible reason is that smoking decreases the respondents' appetite and calorie intake. Besides, nicotine intake enhances metabolism. Hence, smoking reduces fat accumulation. However, a limitation of this study is that it ignores the roles of individuals' preferences toward health. Hence, Wang

(2015) concludes that his estimation of the relationship between smoking and body weight may be biased, because of the potential interaction between smoking and health preferences which may jointly contributes to the body weight. Additionally, Ng *et al.*, (2010) estimates the effects of diet, physical activity, drinking, and smoking on weight. They make use of the 1991, 1993, 1997, 2000, 2004, and 2006 waves of CHNS, which include 10,935 male adults aged from 18 to 55. They find that 5.4% of weight gain among adult men in China within the six waves is due to declines in physical activity, while 2.8–3.1% is due to dietary changes with increasing fat intake. Smoking and drinking have no significant effect on body weight. They explain that physical activity is a critical indicator of body weight because it increases energy expenditure and improves the body's metabolism.

4.2.2 The role of time preferences

4.2.2.1 Definitions of time preferences

Time discounting is also known as time preference or intertemporal choice. According to Smith *et al*. (2005) and Komlos *et al*. (2004), time preferences are defined as the willingness of trading current benefits for future benefits. They represent an individual's level of patience and self-control. Having low time preferences means the respondent values the future, i.e. he/she is patient and has good self-control. Having high time preferences means the respondent prefers satisfaction now and discounts the future heavily, i.e. he/she is impatient and lacks self-control (Zhang and Rashad, 2008).

There are several proxies for time discounting. The most popular measurement of time discounting is the rate of time preferences (Huffman *et al*., 2017; Chabris *et al*. 2008; Courteanche *et al*. 2015; Fuchs, 1982; Rieger, 2015). The respondent is confronted with a hypothetical scenario involving different amount of money at different points in time, and is asked to express his/her preferences among the options. Hence, the rate of time preferences can be revealed implicitly by his/her choice. Specifically, if we assume the rate of time preference is σ, then, the intertemporal discount rate is $1/(1+\sigma)$. Also, the intertemporal discount rate could be calculated by dividing the current benefits by future benefits. For example, an extreme case is when the respondent is indifferent between the present and the future benefits. In this case, the intertemporal discount rate equals to 1. Hence, the rate of time preferences σ equals to 0, which means the respondent values the future benefits the same as the current benefits. Similarly, if the respondent prefers a larger amount of money of £140 in the future instead of a lower amount of money of £100 at present, the intertemporal discount rate is 0.714. Hence, the rate of time preference is 0.401. The higher the rate of time preference, the lower the respondent is willing to trade the current benefits for future benefits, and the greater is the his/her impatience.

Another popular measurement of time preferences is the stated preference approach (Dodd, 2014; Borghans and Golsteyn, 2006; Zhang and Rashad, 2008; Ikeda *et al*. 2010; Fuchs, 1982). In this case, the respondent is asked attitudinal questions with only two answers, namely ‘Yes’ and ‘No’. For example, Fuchs (1982) measures the time preferences based on a question: ‘Do you agree or disagree with this statement: it makes more sense to spend your money now rather than save it for the future’. If the respondent agrees with this statement, he/she is defined as having high time preferences. Likewise, if the respondent disagrees with this statement, he/she is defined as having low time preferences.

Other indicators have been used as proxies of time discounting in the previous studies, such as the saving rate (Smith *et al*., 2005; Komolos *et al*., 2004; Brown and Biosca, 2016) and the household debt ratio (Komolos *et al*., 2004). People with high time preferences tend to spend rather than save in order to obtain satisfaction now. Hence, a high saving rate is associated with a low rate of time preference, whilst a low saving rate or dissaving is associated with high time preferences. Likewise, respondents with debt show a high rate of time preference, because they incur debt in order to finance current consumption as opposed to future consumption.

4.2.2.2 Time preferences and obesity

Fuchs (1982) is the first to link time preferences with the probability of being overweight. A pilot survey was conducted in 1979 with 500 men and women in the Nassau and Suffolk Counties in the US. In order to test for the possible effect of time preferences, overweight is regressed on intertemporal choices and several other control variables. Two measurements of intertemporal choices are used in this study, including the rate of time preferences and stated time preferences. The empirical results show that the effect of time preferences on weight is not statistically significant.

A significantly positive relationship between time preferences and obesity has been found in a number of studies. Almost all the studies explore the relationship between time discounting and obesity in developed countries, especially the US.

Smith *et al*. (2005) propose the hypothesis that high time preferences contribute to increasing body weight. They use individual data from the 1989 wave of the NLSY. All respondents are at least 24 years old, and the mean age is 28.1. The authors model obesity as a function of family income, age, age squared, ethnicity, and time preferences, proxied by saving. They find evidences that higher time preferences are associated with greater obesity among men, and to a lesser extent among women. Specifically, it is found that time preferences are positively associated with BMI among black and Hispanic men, as well as among black women. The authors suggest that the ethnic and gender differences in the effect of time preferences on BMI and obesity may relate to varying cultural beliefs and attitudes about body image. They further argue that more people are consuming more calories than they expend because they have become less willing to trade current pleasures for potential future health benefits.

Komlos *et al*. (2004) estimate the association between the rate of time preferences and obesity, using five waves of the NHANES from the 1970s until the end of the century in the US. Saving and customer debt are used as indicators of time preferences. The respondents are between the ages of 20 and 74. The prevalence of obesity shares similar trends with the debt to income ratio, whilst it shares opposite trends with the saving rate. Komlos *et al*. (2004) indicate that individuals with a high rate of time preference are more likely to be obese, because they tend to consume more high-calorie foods and adopt non-physically active life styles at the expense of lower levels of health and utility in the future.

Using the 2004 Roper Center Obesity data, complemented with 2003 data from the Behavioural Risk Factor Surveillance System, Zhang and Rashad (2008) analyse the possible effect that time preferences have on BMI in the United States. Lacking willpower is used as a proxy for the high rate of time preference, because it reflects the weak ability of self-control. After controlling for a variety of covariates, including marital status, education, age, race, ethnicity and residence in an urban area, and correcting for potential selectivity, they find that time preferences have a significantly positive impact on BMI and the probability of being obese, particularly for males. This finding is consistent with Smith *et al*. (2005).

Chabris *et al*. (2008) show an analysis of the influence of time preferences on individual behaviours, such as smoking, physical exercise, dieting status, BMI, nutrition, saving, borrowing, wealth, and gambling. Their sample includes 555 subjects from three new databases, including the Weight study, the Cognition study, and the Web study in the US. Different from previous proxies of time discounting, the discount rate is used, which is a popular and direct measurement. The discount rate is estimated with 27-choice delay-discounting laboratory tasks, which contains 27 choice trials in a fixed order. In each trial, the respondents are asked to express their preference for either a smaller immediate monetary reward, or a larger delayed monetary reward. Conditioning on gender, age, depression symptoms, education, and cognitive ability, Chabris *et al*. (2008) show that the discount rate is one of the most important variables in their model, positively related with BMI.

Using the NLSY dataset, Courtemanche *et al*. (2015) report consistent results with Smith *et al*. (2005), i.e. higher time preferences are associated with higher BMI, especially among males, after controlling for demographic characteristics, IQ, work hours, and risk preferences. Both a cross-sectional dataset in 2006, and a panel dataset from 1986 to 2006 are used in the analyses. The discount rate matched with price data from the Cost of Living Index is used as a proxy for time preferences.

A group of studies have also focused on other developed countries. For instance, Dodd (2014) explores how the body weight outcomes could be driven by intertemporal preferences in Australia. To this end, they make use of 1,868 observations obtained from the 2008 South Australian Health Omnibus Survey. Two stated preference indicators of intertemporal preferences are used, based on both binary choices over monetary values and health outcomes. Dodd (2014) uses Quantile regression analysis and estimates the relationships between time discounting and relative body weight across the distribution of BMI. With control variables, such as age, gender, income, education, country of origin, and functional health literacy, Dodd (2014) finds that the indicator of time preference in the monetary domain has a positive association with BMI at most of the quantiles, especially at the overweight and obese ranges. However, the indicator of time preference elicited from health outcomes shows no significant relationship with BMI. This supports the earlier studies which found that indicators of time preference for money were significant determinants of the probability of being overweight/obese (Smith *et al*., 2005;

Komlos *et al*., 2004). The possible reason for the difference between the two indicators is that the health-related questions used in the survey are too cognitively difficult for the respondents, and the respondents are more familiar with the monetary domain questions.

Brown and Biosca (2016) are the first to explore the relationship between body fatness and time preferences in the UK. They use three different ways to measure body fatness objectively, including BMI, the percentage of body fatness, and WC. Based on 15,591 individuals from the second wave (2010/2011) of the Understanding Society Survey, they show that there is a negative relationship between the three measures of body fatness and saving money. This is consistent with previous results (Komlos *et al*., 2004; Chabris *et al*., 2008; Dodd, 2014). However, opposite with the earlier results about the gender difference (Smith *et al*., 2005; Zhang and Rashad, 2008; Courtemanche *et al*., 2015), a stronger association is found for women than men. Brown and Biosca (2016) argue that this could be due to the fact that their study uses objective measures of fatness, whist the previous studies used self-reported measures of BMI, which could underestimate the importance of the relationship between saving and BMI among women.

Pickering *et al*. (2017) further test the relationship between time preferences and probability of being overweight/obese among older English people, using saving as a proxy for time preferences. But they find that this relationship is unclear. They make use of 10,502 observations aged 50 and above from the 2004, 2008 and 2012 waves of the English Longitudinal Study of Ageing. The results show that saving is negatively and statistically significantly associated with BMI only when using the Generalised Method of Moments model, but not when using the Fixed effects model. Meanwhile, the authors consider two different types of savings, which are safe and risky savings. They reveal that choosing to invest in safe savings is negatively and statistically significantly associated with the likelihood of being obese only when using the random-effects Probit model, while investing in risky savings is not significantly related to the likelihood of being overweight/obese or BMI (Zhang, 2018). Additionally, the saving ratio that represents the proportion of saving relative to income is also found to be insignificantly related to the probability of being overweight/obese and BMI. The possible reason for the difference with the previous study by Brown and Biosca (2016) is due to the different target group of population, as current study focuses on older population whilst the previous study focuses on adults in the UK.

A majority of papers use cross-sectional datasets. Similar to Courtemanche *et al*. (2015), Borghans and Golsteyn (2006) use both cross-sectional and time series data in the Netherlands. They use data from the DNB household Survey for the period 1995-2004. Their regression results show that there is a significant relationship between time discounting and BMI, when time discounting is measured based on questions relating to the ability of managing expenditures, the ability of managing income, and saving attitude. However, the relationship between time preferences and BMI is absent using other proxies, which are based on the questions relating to risk aversion, future attitude, behaviours of spending money, and debt amount excluding mortgage. Borghans and Golsteyn (2006) conclude that the relationship between time preferences and relative body weight is sensitive to the exact proxies of time preferences. From their panel analyses, Borghans and Golsteyn (2006) show that the average individual discount rate does not change from 1995 to 2004, but there is a positive relationship between time preferences and BMI over the period. This implies that high discounters gain more weight than low discounters.

A few studies focus on developing countries. Among these, Dioikitopoulos *et al*. (2019) investigate the influence of the cultural origins of time preferences on obesity rates around the world. Using a panel of 132 developing and developed countries during the period 2000-2015, they indicate that lower obesity rates is found among those historically long-term oriented countries. This may be caused by the fact that historically long-term oriented culture helps people form better diet habits and exercise behaviours, through the improved ability of individuals to delay gratification.

Rieger (2015) studies the importance of risk aversion and time preference on nutrition of adults and children in rural Cambodia. They use experimental data with 369 adult observations from the Livelihood Enhancement and Association of the Poor Project in 2011. The discount rate is used as an indicator of time preference, and it is measured by a question that asks the respondents to choose between an amount of money today and a series of larger amounts next week. In the case of adults, after controlling for demographic characteristics, whether the respondent is a household head, and literacy, he suggests that weight and BMI are positively correlated with time preferences. In other words, impatient people are heavier.

Gao and Shen (2017) estimate whether the level of subjective well-being impacts on BMI and the likelihood of being overweight or obese. Using the 2013 wave of the Preference Parameters Study, 1318 respondents are covered in their analysis. Specifically, 818 respondents are from urban areas, namely, Beijing, Shanghai, Guangzhou, Chengdu, Wuhan, and Shenyang, whist 500 respondents come from rural areas, including Hunan, Hubei, Sichuan, and Liaoning. Control variables are considered, including time preferences, demographic characteristics, risk aversion, labour intensity, smoking status, drinking status, physical exercise, and eating habits. Happiness is proxied for the subjective well-being. The results suggest that happy female drinkers who have meals regularly and high labour intensity are less likely to have a high BMI among urban respondents. Moreover, respondents with less siblings, who have meals irregularly, and have more income tend to have high BMI among rural respondents. Age has an inverted U-shaped relationship with BMI. Debt is used as a proxy variable for time preference, and there is no evidence to support the significant relationship between time preference and BMI. One possible reason for this finding is that debt is not a suitable proxy for time preferences in the Chinese context, because it also measures risk preference and interest rate effects. Regarding lifestyle, consistent with the study by Ng *et al*., (2010), the combined prevalence of overweight and obesity is negatively associated with physical activity and regular eating habits in daily life. Moreover, smoking does not affect BMI, whilst drinking is negatively associated with BMI among urban respondents. The possible reason is that drinking alcohol is a substitute behavior for eating, and frequent drinking tend to happen together with infrequent eating, thus the respondents tend to have lower BMI.

4.2.3 Contribution

We contribute to the literature by investigating the extent to which time preferences may be associated with the probability of being overweight, being obese and with BMI, using a national representative dataset in China. Although, Gao and Shen (2017) also test the role of time preferences in explaining obesity in the Chinese context, this is not the main focus of their study, and they do not look at separate effects on the probability of being overweight and obese due to data limitations. Also, their study only covers six urban areas and four rural areas, while our study uses a national representative dataset. Moreover, our study makes use of different and better proxies for time preferences.

4.3 Data

4.3.1 The China Family Panel Studies

We use the 2014 wave of the China Family Panel Studies. This is a nationally representative, annual longitudinal survey. The CFPS is designed to collect comprehensive data at community, family and individual levels. Multiple domains are included at each level to help researchers to better understand the complicated national situation in China. The survey collects information on government policies, population distribution, and traffic situation at the community level. It covers family financial situation, family members and their relationships, and social networks at the family level. Finally, the survey includes individual educational background, employment status, marital status, physical and mental health conditions for all family members (Xie and Hu, 2014).

The CFPS was launched in 2010 by the Institute of Social Science Survey of Peking University, China. In the baseline survey, the CFPS successfully interviews approximately 30,000 individuals from approximately 15,000 families (Xie and Hu, 2014). Follow-up surveys were conducted in 2012, 2014, and 2016. This survey covers 25 provinces[12], excluding Hong Kong, Macao, Taiwan, Xinjiang, Tibet, Qinghai, Inner Mongolia, Ningxia, and Hainan, and represents 94.5 percent of the total population in China (Xie and Lu, 2015).

We focus on the 2014 wave of the CFPS in this study, because the variable of time preferences is only available in this wave. As time preferences are only available for adults, children and adolescents are excluded from this analysis. The time preference questions are not asked to every respondent in the survey. Instead, only the family members who are most familiar with the family's financial situation answer this question. In order to control for the potential effect of extreme values and faulty data, we drop observations below the 1st percentile and above the 99th percentile for each continuous main regression variable. A total of 3,890 individuals aged from 20 to 89 years old are included in this analysis.

[12] This part draws on my first year proposal (Fu, 2016)

4.3.2 Variables

4.3.2.1 Weight-related variables

BMI is widely used in the measurements of overweight and obesity (Gao and Shen, 2017). It is defined as weight in kilograms over the square of height in meters. The WHO (2018) defines the respondents as being overweight if their BMI is between 25 and 30 kg/m^2. Moreover, the respondents are defined as being obese if their BMI is greater than 30 kg/m^2. However, this standard is not applicable to Asians, as a higher prevalence of overweight and obesity has been found among Asians than white people of the same age, gender, and BMI. Therefore, it is recommended to use different BMI cut-off points in different Asian countries. For Chinese adults, we follow the standard from the Working Group on Obesity in China (WGOC) (Zhou, 2002), and set 24 and 28 kg/m^2 as cut-off points respectively for overweight and obesity. Using the CFPS, *obesity* is defined as 1 if the BMI is 28 kg/m^2 or greater, and 0 otherwise. Similarly, *overweight* is defined as 1 if the BMI is 24 kg/m^2 or greater, and 0 otherwise. It should be noted that the variable *overweight* represents those who are overweight or obese.

According to the WGOC, the optimal BMI range is between 18.5 and 24. A categorical variable *order_bmi* is used to represent the different levels of BMI. It equals to 1 if the BMI is less than 18.5, which represents the underweight category. It equals to 2 if the BMI is 18.5 kg/m^2 or greater, but no more than 24 kg/m^2, which represents the normal weight category. It equals to 3 if BMI is equal to or great than 24 kg/m^2, but less than 28 kg/m^2, indicating the overweight (but not obese) category. It equals to 4 if the BMI is 28 kg/m^2 or greater. This last category denotes obesity.

4.3.2.2 Time discounting variables

Following the stated preference approach (Dodd, 2014; Borghans and Golsteyn, 2006; Zhang and Rashad, 2008; Ikeda *et al*. 2010; Fuchs, 1982), the variable *TP* is defined as a dummy variable. It represents time preferences in the analysis. If the value of *TP* is 1, it means the respondent has high time preferences and prefers satisfaction now. If the value of *TP* is 0, it indicates low time preferences and the respondent values the future. In the 2014 wave of CFPS, the interviewer asked the respondent 'I intend to live in the present more and do not consider the future. Which degree do you think this opinion fits you?' There are five options following this question, namely 'Totally inapplicable', 'Somewhat inapplicable', 'Generally applicable', 'Somewhat applicable' and 'Totally applicable'. The variable *TP* equals to 1 if the respondent chooses 'Somewhat applicable' or 'Totally applicable', and 0 otherwise.

Based on the same question mentioned in the last paragraph, we define the ordinal variable *TPc,* which is an alternative measure of time preferences and is used in the estimations for robustness tests. *TPc* represents the degree of the respondent's time preferences. High values of *TPc* imply high time preferences. Specifically, *TPc* equals to 1 for 'Totally inapplicable', 2 for 'Somewhat inapplicable', 3 for 'Generally applicable', 4 for 'Somewhat applicable', and 5 for 'Totally applicable'.

4.4 Hypotheses

Overweight and obesity are typically caused by excess caloric intake and physical inactivity (Zhang and Rashad, 2008). Ford *et al*. (2017) report that increased income, urbanization, and shifts in food availability and pricing facilitate the increased consumption of processed modern foods. Also, they argue that changes in diets from traditional food to modern food have increased the total caloric intake in low- and middle-income countries. Relative to modern diets with high energy density, traditional diets are lower in animal-source foods, salt, and refined oils, sugars, and flours. This is the reason why obesity was not a major concern many years ago, but has become more widespread nowadays. Kang and Ikeda (2016) show that obesity is considered as an outcome of time preferences under trade-offs between a small immediate reward and a larger delayed reward. For instance, obesity may be a result of a trade-off between eating high-calorie meals and enjoying future good health. Furthermore, Smith *et al*. (2005) report that if people are less likely to trade current pleasures for potential future health benefits, they will consume more calories than they expend.

Time preferences are a measure of impatience and self-control. Therefore, a high value of time discounting means the respondents are impatient and discount the future heavily. It is expected that people with higher time preference are likely to eat calorie-rich food and enjoy a sedentary lifestyle at the expense of an increased body weight in the future. In contrast, people with a low value of time preferences are patient and have good self-control, and may sacrifice the current consumption of unhealthy foods in order to have good health conditions in the future.

A positive relationship between time preference and obesity has been found in existing studies in developed countries (Smith *et al*., 2005; Komlos *et al*., 2004; Zhang and Rashad, 2008; Chabris *et al*., 2008; Courtemanche *et al*., 2015; Dodd, 2014; Brown and Biosca, 2016; Borghans and Golsteyn, 2006) and in a developing county (Dioikitopoulos *et al*., 2019; Rieger, 2015). No significant association between time preferences and obesity has been found by Gao and Shen (2017) in China. This can be due to the proxy of time preferences used in their analysis, namely debt. Borghans and Golsteyn (2006) argue that the amount of debt is not a good proxy for time preferences. This may be caused by the fact that debt also measures risk preferences and interest rate effects. Therefore, we propose the following hypotheses:

H1. Higher time preferences are associated with a higher probability of being overweight.

H2. Higher time preferences are associated with a higher probability of being obese.

4.5 Specification and estimation methodology

4.5.1 Baseline specification

The following Equations are estimated:

$$Overweight_i = \beta_0 + \beta_1 X_i + \beta_2 TP_i + \varepsilon_i \quad (1)$$

$$Obesity_i = \beta_0 + \beta_1 X_i + \beta_2 TP_i + \varepsilon_i \quad (2)$$

where subscript i indexes individuals, $Overweight_i$ indicates the probability of overweight. It equals to 1 if the respondent is overweight, and 0 otherwise. $Obesity_i$ denotes the probability of obesity. It equals to 1 if the respondent is obese, and 0 otherwise. TP_i represents whether the respondent has a high time preference or not, X_i is a vector of control variables, and ε_i is the error term.

In accordance with Hypothesis 1, $'\beta_2'$ in Eq. 1 is expected to be significant and positive, which implies that high time preferences are significantly positively associated with the probability of being overweight when all the other variables keep constant. In line with Hypothesis 2, $'\beta_2'$ in Eq. 2 is also expected to be significant and positive, which shows there is a significantly positive relationship between time preferences and the prevalence of obesity holding other variables constant.

Several control variables are included in the estimation: Gender, whether the respondent is living in a rural area or urban area, whether the respondent is married, whether the respondent has problems with carrying out any of the daily activities, smoking status, drinking status are dummy variables according to the questionnaire. Continuous variables, such as age, age squared, time spent on doing physical exercises every week, years of education of the respondents, their father, mother, and spouse are also included in the analysis. The natural logarithm of net income per capita, family assets per capita, and food expenditure per capita are also considered. There are four categories of self-reported health status, including very good, good, fair, and poor. Similarly, four categories of work types are included among the control variables, namely, unemployed, working in agriculture, employed, and self-employed. The definitions of all these variables are shown in the Appendix 4.1. This study includes provincial fixed effects as control variables to capture the unobserved provincial effects.

Moreover, depression is a dummy variable equals to 1 if the respondent has a moderate or high level of psychological distress, and 0 if the respondent has a low level of psychological distress. According to Kang *et al*. (2015), the level of psychological distress is classified based on the 6-item Kessler psychological distress scale (K6), which is a useful tool in measuring mental illness among Chinese. Six questions are asked to the respondents, including how often they feel 'so depressed that noting could cheer you up', 'nervous', 'restless or fidgety', 'hopeless', 'everything is an effort', 'life is meaningless' during the past month. There are five options for each question, namely 'Almost daily' (4), 'Often' (3), 'Half of the time' (2), 'Sometimes' (1), and 'Never' (0). Therefore, adding all the six answers together, K6 ranges from 0 to 24. The respondent has a moderate or high level of psychological distress if K6 is 8 or higher, and the respondent has a low level of psychological distress if K6 is less than 8.

Cognitive ability consists of two parts. The first part is a word test. 34 words are ordered from simple to difficult levels. 1 represents the easiest word, and 34 stands for the most advanced word. The interviewer shows a card with a word on it, and asks the respondent to read it. If the respondent can read the word correctly, the interview will show him/her the next word, which is a little more difficult than the previous one. The test ends if the respondent fails to answer or answers incorrectly 3 questions in a row. The number of the last correct word represents the score of the word test of the respondent. The second part is math test. There are 24 math questions ordered from 1 as the easiest question to 24 as the most difficult question. Same as word test, the interviewer shows questions in order and asks the respondent to answer them. The number of the last correct answer is the score of the math test of the respondent. In this study, the variable *cognitive* is the sum of the scores of the word and math tests, and a higher value represents a higher level of cognitive ability.

Age is assumed to have an inverted u-shaped relationship with the probability of being overweight/obese (Li *et al*., 2012; Gao and Shen, 2017). According to Du *et al*. (2004) and Zhang *et al*. (2008), SES is expected to be positively correlated with fatness. Hence, the variables *edu_y, fedu_y, medu_y, sedu*_y, *incomepc* and *wealthpc* are expected to be positively and significantly associated with the probability of being overweight/obese. However, in the females sample, a negative relationship between SES and fatness is expected, because women prefer to have a good body appearance like the women in developed counties (Sobal and Stunkard, 1989; McLaren, 2007; Jones-Smith *et al*., 2012; Xiao *et al*., 2013). The variable *male* is expected to be positively associated with overweight/obesity, as a big body size among men is a sign of physical dominance and prowess (Mi *et al*., 2015). The marginal effects associated with *rural* in both Eq. 1 and Eq. 2 are expected to be significantly negative. Urban areas have a relative higher economics development than rural areas, and have a wide selection of restaurants with different characteristics, which increases energy and nutrient intakes among urban population (Li *et al*., 2012).

Good self-reported health status is expected to have a negative correlation with the probability of being overweight/obese, while poor self-rated health is expected to have a positive association with the probability of being overweight/obese (Gao and Shen, 2017). Obesity is widely known as a major risk factor for people's health and is associated with a group of diseases. Thus, if the respondents consider their health status as very good, they are less likely to be overweight/obese. Oppositely, if the respondents consider their health status as very poor, they have a higher probability of being overweight/obese. Following the literature (Yu *et al*., 2011; Zhang *et al*., 2016), depressed respondents are expected to have a lower probability of being overweight/obesity because of the poor appetite. Behaviours, such as smoking, drinking, and physical activity are expected to show negative correlations with the percentage of overweight/obesity. It is shown that cigarette smoking and alcohol drinking may decrease appetite, and behave as a substitute behavior for eating, thus decreasing the probability of being overweight/obese. Physical exercises increase energy expenditure and improve body's metabolism, which prevents people from obesity (Wang, 2015, Ng *et al*., 2010; Gao and Shen, 2017). A poor diet, characterized by more energy intake or irregular meals, could increase the probability of being overweight/obese (Ng *et al*., 2010; Gao and Shen, 2017). Using the amount of money spent on food consumption as a proxy of diet, it is expected that there is a positive relationship between variable *foodpc* and the probability of being overweight/obese, because with more money spent on food, the respondent is more likely to take a large amount of calories.

Furthermore, variable *good_mob* is expected to be negatively correlated with overweight/obesity (Pickering *et al*., 2017). Consistent with Zhang and Rashad (2008), it is expected that married respondents are less likely to be overweight/obese, because they are responsible for family members, and they have more family burdens than single respondents, thus they are more eager to have a good body condition. Based on Courtemanche *et al*. (2015), occupation types should be included in the estimation of obesity, because they reflect different labour market activities. According to the Chinese culture, occupation types includes four categories, namely unemployed, working in agriculture, employed, and self-employed. We expect employed workers to have a lower chance of being overweight/obese. Compared with farmers, unemployed, and self-employed people, employees need to pay more attention to their appearances, because they need to get along with colleagues and customers every day and a good body shape improves their confidence. Yu *et al*. (2010) show that higher intelligence in childhood is negatively related to be obese in adulthood. In our dataset, respondents are asked about their cognitive ability level, and we expect that a higher level of cognitive ability in adults is negatively associated with the probability of being overweight/obese, because intelligent people are more likely to have healthy behaviours, such as frequent intake of fruits and vegetables or low intake of high calorie foods.

4.5.2 Estimation methodology

We use both linear probability models (LPM) and Probit models to estimate Eq. 1 and 2, because our left-hand side variables are dichotomous. These two models are used to estimate the determinants of the probability of being overweight/obese. The dependent variables in the two equations are respectively a dummy variable equals to 1 if the respondent is overweight, and 0 otherwise, and a dummy variable equals to 1 if the respondent is obese, and 0 otherwise. For Probit model, we report and discuss marginal effects to show the magnitudes of the effects that the independent variables have on the probability of being overweight/obese.

In order to check the robustness of our main results, an ordinary least squares (OLS) regression model is used to test the relationship between time discounting and BMI. The specification of this model is similar to Eq. 1. The only difference is that the dependent variable in Eq. 1 is now BMI. Using the OLS model, we investigate whether there is a linear relationship between time preferences and body weight, via the BMI.

Next, following Dodd (2014), we use the Quantile regression model to investigate the relationship between time discounting and body weight across different points of the distribution of body weight. The independent variables in this model are the same as in Eq. 1, but the dependent variable is replaced with BMI. The Quantile regression model enables us to analyse the extent to which the time discounting effect on BMI differs along the distribution of BMI. An advantage of this estimator over OLS is that it studies the entire distributions instead of relying on the means (Mata and Machado, 1996). We include Quantile regressions because we expect the relationship between time preferences and BMI to vary across quantiles. We report detailed results from Quantile regression for the 10th, 30th, 50th, 70th, and 90th quantiles.

Finally, an Ordered Probit regression is used to estimate the relationships between time preferences and the probability of being in various BMI categories, namely underweight, normal weight, overweight, and obesity. It is different from the Probit model, because the dependent variable in Ordered Probit regression has a natural ordering. In this case, the dependent variable is replaced with *order_bmi,* which equals to 1 for underweight, 2 for normal weight, 3 for overweight, and 4 for obesity. This model keeps the same control variables as in Eq. 1. The Ordered Probit model is a good method because it estimates the relationships between time preferences and the probability of being underweight, normal weight, overweight and obese separately, in which way, we can compare the relationships among the four groups (Plurphanswat and Rodu, 2014). For Ordered Probit model, we report and discuss marginal effects. All analyses are performed with Stata SE version 15 (StataCorp, Texas, USA).

4.6 Summary statistics

4.6.1 Descriptive statistics

Table 4.1 presents descriptive statistics for the main regression models. Column 1 refers to the full sample, column 2 to the male sub-sample, column 3 to the sub-sample of females, and column 4 to the difference between the males and females. This distinction between males and females is important according to the literature (Mi, *et al*, 2015), because there is a gender difference in the probability of being overweight and obese.

Consistent with Mi *et al*. (2015), males have relatively higher BMI values than females. The average values of BMI are 23.61, 24.06, and 23.14 among the full sample, males, and females, respectively. The prevalence of overweight and obesity also shows higher values among males than females. Specifically, 43.5 percentage of the full sample are overweight or obese. The corresponding values among men and women are 50.1 and 36.7 percentage, respectively. The difference between these two figures is statistically significant at the 1 percent level. Regarding obesity, 10.1 percent of men are obese, while only 7 percent of women are obese. The difference between these two figures is also statistically significant at the 1 percent level.

Focusing on time discounting (*TP*), 31.4 percent of the people in the full sample have high time preferences and discount the future heavily. The corresponding figures for males and females are 31.3 and 31.6 percent, respectively. The variable *TPc* is an alternative measure of time preferences. The figure in column 1 shows that the average value of *TPc* is 2.783 in the full sample. No big difference of this ordinal variable is found by gender.

In the full sample, almost half of the respondents are males. The average age is 48.57. Only six percent of respondents live in rural areas. The average years of education are 10.24. The corresponding values of the respondents' father, mother, and spouse are 5.647, 3.503 and 9.740. The educational levels of parents are lower than the next generation, and males' educational level are higher than females. This phenomenon is reasonable in the Chinese context (Wu, 2011). The amount of net income per capita, wealth per capita, and food expenditure per capita are 22,248, 238,307, and 7,814 YUAN, respectively. 99.7 percent of the respondents are married. 27.5 percent of the respondents consider that they have a very good health status at the time when the survey is conducted. 44.6 percent of the respondents consider their health status as good, and 16.4 percent as fair. 11.5 percent of the respondents consider that they have a poor health status. 8.82 percent of people in the sample have moderate or high level of psychological distress. The average value of cognitive ability is 32.19. Focusing on work status, 32.9 percent of the respondents are unemployed or retired. 3.37 percent of the people surveyed are working in agriculture, and 51.4 percent of them are employees, whilst 12.4 percent are self-employed. 96.7 percent of the respondents have good mobility. 29.8 percent of them are current smokers, and 17.1 percent of them are heavy drinkers. The average time spent on physical activity in the past week of the survey is 4.042 hours.

In columns 2 and 3, males have significantly higher educational level, self-reported health status, cognitive ability, employment rate, smoking rate, drinking rate, and physical activity times than females, while females have higher depression level than males.

4.6.2 T-test results

Table 4.2 presents the percentage of overweight/obese respondents among the high time preference group and low time preference group. The high time preference group is defined when variable *TP* equals to 1, whilst the low time preference is defined when *TP* equals to 0. We can see that people with higher time preferences are significantly more likely to be overweight and obese.

The data shows that 47 percent of people with high time preferences are overweight, while the corresponding percentage among the low time preferences group is 41.9 percent. The difference between these two figures is statistically significant at the 1 percent level. Focusing on males, 53.5 percentage of men with high time preference are overweight, whilst 48.5 percentage of men with low time preference are overweight. The difference is statistically significant at the 5 percent level. Similarly, focusing on females, the prevalence of overweight is 40.3 percent among people with high time preference, while the corresponding prevalence is 35 percent among people with low time preference. The difference is also statistically significant at the 5 percent level. This distribution is consistent with Hypothesis 1.

Focusing on obesity, in the full sample, the prevalence of obesity in the high time preference group is higher than the prevalence in low time preference group, at 9.9 percent and 8.0 percent respectively. The difference between the two figures is statistically significant at the 10 percent level. This distribution is consistent with Hypothesis 2. The corresponding percentages among males are 11.1 percentage and 9.7 percentage in the high and low time preference groups. Yet, this difference is not statistically significant. Among females, the percentage are 8.6 percent in the high time discounting group, compared with 6.3 percent in the low time preference. The difference between these two figures is significant at the 10 percent level. In summary, compared with the differences in the percentages of overweight respondents among respondents with high and low time preferences, the differences in the percentages of obese respondents are smaller and less significant.

4.7 Empirical results

This section reports results of Probit estimations and LPM regressions for the probability of being overweight/obese, as well as results of OLS and Quantile regressions for BMI, and results of Ordered Probit regressions for different categories of BMI. Finally, an alternative measure of time discounting is used for robustness tests.

4.7.1 Probit and LPM regressions for the probability of being overweight/obese

Table 4.3 shows the marginal effects of the Probit models, which estimates the probability of being overweight/obese as a function of time preferences. Column 1 to 3 analyse the probability of being overweight for the full sample, males and females, respectively. Similarly, column 4 to 6 analyse the probability of being obese for the full sample, males and females, respectively.

In column 1 and 4, the marginal effects associated with *TP* are positive and statistically significant. They suggest that having high time preferences is associated with a 6.6 percentage point higher probability of being overweight and a 2.5 percentage point higher probability of being obese. This means that having high time preferences is positively associated with being overweight/obese in the full sample. An intrinsic willingness of enjoying calorie-rich foods and sedentary lifestyles may be the reason for the increased body weight in China. In addition, people with high time preference are more likely to be overweight rather than obese. A plausible explanation for the higher association between *TP* and overweight than the association between *TP* and obesity is that China is a developing county. Most developing countries are still struggling with the underweight problem, and, as such, the prevalence of obesity is lower. These results are consistent with Hypothesis 1, Hypothesis 2, and the existing literature (Zhang and Rashad, 2008; Chabris *et al*., 2008; Courtemanche *et al*., 2015; Dodd, 2014; Brown and Biosca, 2016;

Borghans and Golsteyn, 2006; Rieger, 2015). Yet, our findings differ from those in Gao and Shen (2017), who find no association between time preferences and the probability of being overweight/obese in China, but this difference could be due to the fact that their proxy for time preferences (debt) is inadequate.

Consistent with our expectations and the literature (Mi *et al*., 2015; Du *et al*., 2004; Zhang *et al*., 2008), in column 1, the variables *male, incomepc* and *wealthpc* are positively and significantly associated with the likelihood of being overweight, whereas *employed* is negatively associated with being overweight.

However, different with our expectation, *age* is positively related to overweight, which means that older people have a higher probability of being overweight. This might be explained considering that young Chinese generation pays more attention to their body appearances and adopts healthy life habits. Also, *good_mob* is positively associated with overweight, which is opposite with our expectation. It is possible that those people, who have no problem with daily activities, have more life pressures and may ignore the behavioural risk factors, such as eating fast foods for saving time. By contrast, a majority of people with problems in carrying out daily activities are 45 years old or above.

In line with our expectations and the literature (Mi *et al*., 2015; Gao and Shen, 2017; Yu *et al*., 2011; Zhang *et al*., 2016), in column 4, the variable *male* is positively associated with obesity, whist the marginal effects associated with the variables *slfhvg* and *depression* are significantly negative. However, compared with the results in column 1, the significant effects of income, wealth and good mobility, and employment become insignificant. It is possible that affluent employees with good mobility are more likely to enjoy high calorie foods, which contributes to higher body weight. Because the total amount of energy intake is higher than that of energy consumption. However, the difference between the total amount of energy intake and consumption is not big enough to cause obesity.

In addition, in contrast with our expectations, variables, such as *rural, married, cognitive, foodpc, smoker, drinker,* and *phyact*, have no significant association with the likelihood of being overweight/obese. It is possible that the significant relationships between these variables and the probability of being overweight/obese found in previous studies (Li *et al*., 2012; Ng *et al*., 2010; Wang, 2015) were due to the fact that they did not include time preferences in their models (omitted variable bias).

In the males sample, those affluent respondents with high time preferences are more likely to be overweight. Also, depressed respondents with good self-reported health status are less likely to be obese. In the females sample, older smokers with good mobility and high time preferences have a higher probability of being overweight, whilst those who are employed have a lower probability of being overweight. Moreover, depressed respondents with high income are less likely to be obese.

Table 4.4 shows the results of LPM models. These are consistent with the results of the Probit model, except for the health status in the column 1. The coefficient associated with poor self-rated health status becomes positively significant at the 10 percentage level. This result is consistent with our expectation and the literature (Gao and Shen, 2017), which suggests that the respondents who report a poor health status are more likely to be overweight. Similar to Table 4.3, the data in Table 4.4 indicates that if the respondent has high time preferences, he/she would show a 6.8 percentage point higher probability of being overweight and a 2.8 percentage point higher probability of being obese. In addition, the results in the column 1 also show that affluent older males with poor self-rated health, good mobility and high time preferences tend to be overweight, while employed respondents are less likely to be overweight. In terms of obesity, the figures in column 4 show that males who discount the future heavily tend to be obese, while depressed respondents with very good self-rated health are not likely to be obese. These findings provide further support for our two Hypotheses.

4.7.2 OLS regressions for BMI

Table 4.5 reports the OLS regression estimations of BMI for the full sample (column 1), males (column 2) and females (column 3). The variable *TP* is positively and statistically significantly associated with BMI in all three columns. The association between TP and BMI is stronger for males than females. This finding supports our Hypotheses 1 and 2, and is in line with the results we obtained for the probability of being obese/overweight.

Consistent with our previous estimations, the dummy variable *male* is significantly positively associated with BMI, and *depression* and *employed* are significantly negatively associated with BMI. This suggests that males tend to have a higher relative body weight than females, and depression and employment help the respondents keep away from weight gains.

age shows an inverted U-shaped relationship with BMI. This is consistent with the existing studies (Li *et al*., 2012; Gao and Shen, 2017). The coefficients in Table 4.5 suggest that the value of BMI in the full sample peaks at the age of 70.5 years. There is a positive relationship between age and BMI before 70.5 years and a negative relationship thereafter. In the female subsample (column 3), the BMI among females peaks at 45. We find no association between age and BMI among males. Moreover, apart from employees, farmers also have relatively lower BMI than the other work types. This can be explained considering that the nature of work in agriculture needs a lot of physical work, especially in China. China is still a developing county, and not every farmer can afford agricultural machines and equipment.

Surprisingly, the effects of income and wealth are insignificant in column 1 in Table 4.5. Also, self-reported health status is insignificant in the full sample in Table 4.5.

The R^2 values of the OLS regression estimations fall between 0.083 and 0.119, which are quite low, but low values are standard in cross-sectional regressions. Moreover, these values are similar to the R^2 values obtained in a similar analysis of BMI by Dodd (2014) and Ikeda *et al*. (2010).

4.7.3 Quantile regressions for BMI

Table 4.6 reports Quantile regression results for all the respondents. Results relative to the 10^{th}, 30^{th}, 50^{th}, 70^{th}, and 90^{th} quantiles appear in columns 1 to 6, respectively. The results show that there are significantly positive associations between *TP* and BMI, along all quartiles of the distribution of BMI. The coefficients increase from 0.312 in the lowest quantile to 0.672 in the upper highest. The significant coefficients in all the quantiles confirm the findings in the OLS regression estimations and show that respondents who have high time preferences have higher BMI than those with low time preferences. This further confirms Hypothesis 1.

As for the control variables, consistent with the OLS regression results, males have a significantly positive relationship with BMI throughout the distribution. There is an inverted U-shaped relationship between age and BMI. Also, income and wealth are only significant in the 30^{th}, 50^{th}, 70^{th} quantiles, which suggests that high SES is not associated with underweight and obesity. Yet, within the normal weight and overweight groups, people with higher SES have higher BMI than those with low SES. Moreover, farmers have lower BMI than the other work types at the lower tail of BMI because of the high demand of the job in terms of physical activity. Employees have lower BMI at the upper tail of BMI. The variable *depression* has significantly negative correlations with BMI, especially in the 70^{th} and 90^{th} quantiles. This indicates that depressed people are more likely to have lower BMI. A possible reason is that depressed people may have poor appetite, and thus suffer from a loss of body weight. The results from the Quantile regressions also show a non-linear association between self-rated health status and BMI across the distribution of relative body weight. Specifically, among the people with lower BMI, people with good self-rated health status have higher BMI. However, among the people with higher BMI, people with good self-rated health status have lower BMI. This may be the reason why we could not find a significant linear relationship between health status and BMI in the OLS model. It is possible that for those overweight and obese respondents, they are less likely to have higher BMI if they consider their health status as very good, because their body weight is relatively higher than normal people, and higher BMI for them will increase the probability of diseases. Thus, if they are confident in their health status, they are less likely to have higher BMI, instead, they will lose weight and protect themselves from diseases. Oppositely, for those people with relatively lower BMI within the optimal BMI group, they don't have the same concerns as overweight and obese respondents. Therefore, if they consider their health status as good, they are more likely to consume more food to make themselves strong and powerful. Consistent with the Probit model and LPM, the respondents with good mobility tend to have higher BMI than those with bad mobility, but this association is only significant in the 50^{th}, 70^{th}, and 90^{th} quantiles. This means that good mobility only matters for those with relatively higher body weight, but not those with relatively lower relative body weight.

Contrary to the previous regression results, there are more significant variables in the Quantile regression models. For instance, smoking has a significantly negative correlation with BMI in the 10^{th} quantile. This is consistent with the previous study by Wang (2015). It is possible that cigarette smoking decreases the appetite, which means that smokers tend to have lower BMI than non-smokers, especially among low weight respondents. Opposite to Gao and Shen (2017), among the low weight respondents, drinkers have higher BMI than non-drinkers.

Table 4.7 shows the Quantile regression results for males. As in Table 4.6, columns 1 to 6 represent the 10^{th}, 30^{th}, 50^{th}, 70^{th}, and 90^{th} quantiles of BMI. The data shows that the coefficients of *TP* are positively significant in all the quantiles, and the magnitudes of the coefficients vary between 0.31 and 0.823 throughout the distribution of BMI. This suggests that all other variables held constant, high time preferences are associated with higher BMI among men. Similarly, Table 4.8 shows the Quantile regression estimations for females. In this case, the positively significant relationships between *TP* and BMI only exist in the columns 3 and 5. This means that time preferences are not associated with BMI if the female respondents have a relatively lower weight. In China, a thinner figure is widely recognized as the standard of beauty among women. Women with relatively low weight are recognised as having a beautiful body shape, and they are more likely to keep their body appearances regardless of their time preferences. In addition, we can also find the non-significant coefficient in column 4, the possible reason is that for those ladies whose BMI is a little bit higher than the optimal BMI range. They are reluctant to have a higher BMI because any increase in BMI will result in a big change in their body shapes, such as having higher waist circumferences. Thus, they are more likely to keep their body appearances regardless of their time preferences. In conclusion, the Quantile regressions allow us to estimate the association between time discounting and BMI across the distribution of body weight. The results suggest that higher time preferences are associated with a higher BMI in the full sample, for males, and for females with relative higher body weights.

4.7.4 Ordered Probit regressions for different BMI categories

Table 4.9 reports the marginal effects of the Ordered Probit models for the underweight (column 1), normal weight (column 2), overweight (column 3), and obese (column 4) categories for the full sample. The marginal effects associated with *TP* are significant at the 1 percentage level in all four columns. In column 1, the marginal effect associated with *TP* is -0.013. It implies that if the respondent discounts the future heavily, he/she has 1.3 percentage points lower probability of being underweight. The marginal effect in column 2 suggests that if the respondent discounts the future heavily, he/she has 5.1 percentage points lower probability of having normal weight. The marginal effect in column 3 indicates that higher time preferences are associated with 3.9 percentage points higher probability of being overweight, whilst that in column 4 suggests that higher time preferences are associated with a 2.6 percentage points higher chance of being obese. These findings are consistent with the results in our Probit models and with Hypotheses 1 and 2. Also, the marginal effect associated with *TP* in the column 3 is larger than the figure in column 4, which means that the association between time preferences and the probability of being overweight is higher than the association with the probability of being obese. This implies that people with higher time preferences tend to be overweight rather than obese.

Also, consistent with the results in the Probit models, young depressed female employees with very good self-reported health status are less likely to be overweight/obese, but are more likely to be underweight or have normal weight.

Tables 4.10 and 4.11 show the marginal effects of the Ordered Probit models among males and females, respectively. Among men, affluent self-employed men with higher time preferences are more likely to be overweight/obese, but are less likely to be underweight or have normal weight. Among women, young, unmarried employees, who have low time preferences and are currently not smoking, are more likely to be underweight or have normal weight, but are less likely to be overweight or obese.

4.7.5 Robustness tests based on an alternative proxy of time preferences

In order to test the robustness of the previous results, the variable *TPc* is used as an alternative proxy of time discounting. Probit models, LPM, OLS regressions, Quantile regressions, and Ordered Probit models are used for the analyses. All our results are robust to replacing the time preference dummy with the ordinal time preference variable, and it is proved that time preferences are positively associated with the probability of being overweight, being obese, and with BMI. These results are presented in Table A to Table I in Appendices 4.2.

4.8 Conclusion and implications

The prevalence of overweight and obesity among the Chinese population has been increasing significantly in recent years. However, to the best of our knowledge, the role of time preferences in explaining obesity in China has not been investigated with national representative datasets. This study fills this gap in the literature by investigating the extent to which high time preferences may be associated with the likelihood of being overweight, being obese, and with BMI. Making use of the 2014 wave of the China Family Panel Studies, and of linear probability models and Probit models, we find that time preferences are positively associated with the probability of being overweight and obese in the Chinese context, which means that impatient people are prone to be overweight and obese.

In addition, the association between time preferences and the probability of being overweight is higher than the corresponding relationship with obesity. Based on the results of the OLS estimations, time preferences are also positively and statistically significantly associated with BMI. These results are robust to using an alternative measure of time preferences and different estimation methods, such as Quantile regressions and Ordered Probit models.

A limitation to this study is that we use the BMI to classify overweight and obesity. There is evidence that BMI is not an accurate measure of fatness because it does not distinguish fat from fat-free mass, such as muscle and bone (Burkhauser and Cawley, 2008). Waist circumference and the percent body fatness are better than BMI, but these two measures are not available in the 2014 wave of CFPS. Another limitation of our work is that it is only based on cross-sectional data, so we cannot explore the effect of changes in time preferences over time.

As for policy implications, direct intervention policies, such as educating people from being myopic to being more far-sighted may be effective in reducing the prevalence of overweight and obesity in China. In detail, emphasising the importance of nutrition balance and improving the awareness of calories content in food will be beneficial in China. Because good nutritional knowledge is positively associated with people's long-term oriented choices, such as adopting and maintaining healthy diets and keeping reasonable calorie intakes. Meanwhile, policies, such as improving physical activities, are needed. If people sacrifice the sedentary life now and invest their time in doing physical exercises, the prevalence of overweight and obesity in China will reduce.

Table 4.1 Descriptive statistics

Variables	Full sample (1)	Males (2)	Females (3)	Difference (4)
bmi	23.61	24.06	23.14	-0.914***
overweight	0.435	0.501	0.367	-0.134***
obese	0.0861	0.101	0.0700	-0.031***
TP	0.314	0.313	0.316	0.00400
TPc	2.783	2.770	2.798	0.02800
male	0.510	1	0	-1
age	48.57	49.18	47.94	-1.232***
rural	0.0609	0.0630	0.0590	-0.00300
edu y	10.24	10.71	9.749	-0.956***
fedu y	5.647	5.533	5.766	0.233
medu y	3.503	3.378	3.632	0.254*
sedu y	9.740	9.274	10.22	0.950***
rincomepc	22248	23000	22000	-630.1
married	0.997	0.996	0.998	0.00200
slfhvg	0.275	0.301	0.248	-0.053***
slfhg	0.446	0.458	0.433	-0.0260
slfhf	0.164	0.134	0.196	0.061***
slfhp	0.115	0.107	0.124	0.017*
depression	0.0882	0.0760	0.101	0.026***
cognitive	32.19	33.69	30.62	-3.075***
unemployed	0.329	0.247	0.415	0.168***
workagri	0.0337	0.0350	0.0330	-0.00200
employed	0.514	0.580	0.444	-0.136***
slfemploy	0.124	0.138	0.109	-0.030***
rwealthpc	238307	240000	230000	-8400
rfoodpc	7814	7856	7770	-86.27
good mob	0.967	0.968	0.965	-0.00300
smoker	0.298	0.556	0.0300	-0.526***
drinker	0.171	0.308	0.0290	-0.279***
phyact	4.042	4.238	3.838	-0.400*
N	3,890	1,983	1907	

Note: Mean values of each variable are presented in this table. In the last row, the observation numbers in each group are presented. *rincomepc*, *rwealthpc*, and *rfoodpc* are expressed in YUAN. Precise definitions of all variables can be found in Appendix 4.1.

Table 4.2 T-test results

	Overweight(%)			Obesity(%)		
	Full sample	Males	Females	Full sample	Males	Females
Time preference						
TP=1	47.0	53.5	40.3	9.9	11.1	8.6
TP=0	41.9	48.5	35.0	8.0	9.7	6.3
Mean diff	-5.1***	-5.1**	-5.3**	-1.9*	-1.4	-2.3*

Note: Precise definitions of all variables can be found in Appendix 4.1. *Mean diff* contains the p-values associated with the t-test for equality of mean percentages among respondents with high time preferences and low time preferences. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Table 4.3 Probit models for the probability of being overweight/obese

	(1) Tover	(2) Mover	(3) Fover	(4) Tobes	(5) Mobes	(6) Fobes
TP	0.066***	0.092***	0.051*	0.025**	0.029	0.024
	(2.99)	(2.92)	(1.78)	(2.03)	(1.57)	(1.54)
male	0.158***			0.034**		
	(6.01)			(2.18)		
age	0.017***	0.011	0.020**	-0.002	-0.002	-0.002
	(3.08)	(1.39)	(2.55)	(-0.64)	(-0.41)	(-0.53)
age2	-0.000***	-0.000	-0.000**	0.000	-0.000	0.000
	(-3.08)	(-1.37)	(-2.32)	(0.17)	(-0.26)	(0.47)
rural	0.004	0.024	-0.011	0.004	0.008	0.014
	(0.09)	(0.35)	(-0.17)	(0.17)	(0.22)	(0.46)
edu_y	-0.003	0.004	-0.008	-0.000	0.003	-0.003
	(-0.88)	(0.69)	(-1.56)	(-0.15)	(0.97)	(-1.05)
fedu_y	-0.001	-0.002	-0.000	-0.001	-0.003	-0.001
	(-0.23)	(-0.55)	(-0.06)	(-0.70)	(-1.32)	(-0.27)
medu_y	-0.002	0.003	-0.005	0.000	0.001	-0.000
	(-0.60)	(0.74)	(-1.21)	(0.19)	(0.61)	(-0.02)
sedu_y	0.001	0.005	-0.006	0.000	-0.001	-0.001
	(0.40)	(1.16)	(-1.42)	(0.20)	(-0.48)	(-0.38)
incomepc	0.030*	0.076***	-0.018	-0.003	0.016	-0.016*
	(1.95)	(3.33)	(-0.98)	(-0.38)	(1.18)	(-1.77)
married	-0.129	-0.178	0.000	0.073	0.122	0.000
	(-0.79)	(-0.83)	(.)	(0.96)	(1.33)	(.)
slfhvg	-0.044	-0.012	-0.054	-0.036*	-0.050*	-0.028
	(-1.39)	(-0.25)	(-1.31)	(-1.96)	(-1.82)	(-1.22)
slfhg	0.005	-0.008	0.030	-0.019	-0.050**	0.011
	(0.17)	(-0.18)	(0.81)	(-1.17)	(-2.04)	(0.55)
slfhp	0.065	0.033	0.075	0.004	-0.043	0.028
	(1.63)	(0.55)	(1.55)	(0.17)	(-1.25)	(1.17)
depression	-0.054	-0.038	-0.073	-0.086***	-0.106***	-0.072***

	(-1.51)	(-0.73)	(-1.58)	(-3.77)	(-2.97)	(-2.79)
cognitive	0.000	0.002	-0.000	-0.000	0.001	-0.000
	(0.17)	(1.52)	(-0.27)	(-0.11)	(1.03)	(-0.08)
workagri	-0.097	-0.025	-0.065	-0.030	0.039	-0.059
	(-1.46)	(-0.26)	(-0.77)	(-0.86)	(0.72)	(-1.19)
employed	-0.068**	0.002	-0.067*	-0.010	0.002	-0.006
	(-2.31)	(0.04)	(-1.94)	(-0.57)	(0.08)	(-0.31)
slfemploy	0.004	0.096	-0.025	-0.015	0.001	-0.015
	(0.11)	(1.57)	(-0.53)	(-0.71)	(0.02)	(-0.67)
wealthpc	0.016**	0.019*	0.010	0.002	0.009	-0.002
	(1.98)	(1.70)	(0.98)	(0.49)	(1.30)	(-0.36)
foodpc	-0.023	-0.033	-0.012	0.015	0.009	0.019
	(-1.29)	(-1.27)	(-0.53)	(1.42)	(0.68)	(1.43)
good_mob	0.129**	0.072	0.196**	0.003	0.055	0.000
	(2.06)	(0.83)	(2.50)	(0.07)	(0.74)	(0.00)
smoker	-0.012	-0.032	0.158**	0.004	-0.011	0.059
	(-0.41)	(-1.06)	(2.20)	(0.24)	(-0.62)	(1.63)
drinker	0.029	0.033	0.051	0.010	0.003	0.041
	(0.98)	(1.04)	(0.61)	(0.58)	(0.16)	(0.94)
phyact	0.000	0.001	-0.000	0.000	-0.000	0.001
	(0.11)	(0.53)	(-0.20)	(0.21)	(-0.08)	(1.03)
N	3890	1983	1904	3846	1959	1884
r2_p						

Note: Columns 1 to 3 refer to regressions for the probability of being overweight, whilst columns 4 to 6 refer to regressions for the probability of being obese. Columns 1 and 4 refer to the full sample; Columns 2 and 5 to males; and columns 3 and 6 to females. Marginal effects are reported, with *z* statistics in parentheses. *N* denotes the number of observations. Fair self-reported health status and unemployed are omitted categories. Precise definitions of all variables can be found in Appendix 4.1. Province dummies are included, but their marginal effects are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Table 4.4 Linear probability models for the probability of being overweight/obese

	(1) Tover	(2) Mover	(3) Fover	(4) Tobes	(5) Mobes	(6) Fobes
TP	0.068***	0.094***	0.051*	0.028*	0.033	0.027
	(2.94)	(2.90)	(1.65)	(1.95)	(1.54)	(1.44)
male	0.160***			0.033**		
	(5.80)			(1.99)		
age	0.017***	0.010	0.020**	-0.002	-0.004	-0.002
	(3.13)	(1.30)	(2.52)	(-0.51)	(-0.96)	(-0.47)
age2	-0.000***	-0.000	-0.000**	0.000	0.000	0.000
	(-3.12)	(-1.27)	(-2.23)	(0.07)	(0.44)	(0.44)
rural	0.002	0.024	-0.018	0.005	-0.002	0.010
	(0.04)	(0.35)	(-0.26)	(0.21)	(-0.07)	(0.32)
edu_y	-0.003	0.003	-0.007	-0.000	0.003	-0.003
	(-0.84)	(0.64)	(-1.31)	(-0.07)	(1.14)	(-1.11)
fedu_y	-0.001	-0.002	-0.000	-0.001	-0.003	-0.001
	(-0.20)	(-0.52)	(-0.13)	(-0.69)	(-1.02)	(-0.36)
medu_y	-0.002	0.003	-0.005	0.001	0.002	-0.000
	(-0.59)	(0.76)	(-1.14)	(0.28)	(0.85)	(-0.16)
sedu_y	0.001	0.005	-0.006	0.000	-0.000	-0.000
	(0.46)	(1.20)	(-1.41)	(0.29)	(-0.20)	(-0.04)
incomepc	0.030**	0.074***	-0.019	-0.003	0.013	-0.022
	(2.00)	(3.37)	(-0.98)	(-0.27)	(1.11)	(-1.50)
married	-0.130	-0.179	0.334***	0.056	0.119*	0.046*
	(-0.80)	(-0.83)	(4.83)	(1.32)	(1.81)	(1.79)
slfhvg	-0.043	-0.010	-0.050	-0.039*	-0.056	-0.020
	(-1.33)	(-0.21)	(-1.20)	(-1.95)	(-1.63)	(-0.91)
slfhg	0.005	-0.007	0.031	-0.023	-0.058*	0.015
	(0.17)	(-0.16)	(0.80)	(-1.24)	(-1.90)	(0.69)
slfhp	0.067*	0.033	0.084	0.001	-0.042	0.029
	(1.65)	(0.54)	(1.58)	(0.04)	(-1.17)	(0.97)
depression	-0.055	-0.035	-0.078	-0.064***	-0.073***	-0.059***

	(1)	(2)	(3)	(4)	(5)	(6)
	(-1.53)	(-0.68)	(-1.63)	(-5.02)	(-3.93)	(-3.18)
cognitive	0.000	0.002	-0.000	-0.000	0.001	-0.000
	(0.16)	(1.54)	(-0.31)	(-0.08)	(0.98)	(-0.06)
workagri	-0.100	-0.027	-0.074	-0.024	0.038	-0.053
	(-1.49)	(-0.28)	(-0.80)	(-0.89)	(0.81)	(-1.57)
employed	-0.069**	0.006	-0.072*	-0.009	0.008	-0.004
	(-2.30)	(0.11)	(-1.96)	(-0.55)	(0.31)	(-0.17)
slfemploy	0.005	0.097	-0.026	-0.011	0.010	-0.015
	(0.12)	(1.58)	(-0.54)	(-0.52)	(0.31)	(-0.61)
wealthpc	0.016**	0.019*	0.010	0.002	0.006	-0.002
	(1.99)	(1.70)	(0.94)	(0.45)	(1.12)	(-0.33)
foodpc	-0.023	-0.032	-0.012	0.014	0.007	0.022
	(-1.28)	(-1.25)	(-0.50)	(1.32)	(0.50)	(1.54)
good_mob	0.125**	0.061	0.197***	0.009	0.019	-0.001
	(2.07)	(0.72)	(2.63)	(0.28)	(0.47)	(-0.03)
smoker	-0.013	-0.032	0.176**	0.007	-0.007	0.109
	(-0.43)	(-1.05)	(2.20)	(0.35)	(-0.36)	(1.44)
drinker	0.030	0.035	0.058	0.011	0.005	0.060
	(0.99)	(1.08)	(0.64)	(0.56)	(0.27)	(0.81)
phyact	0.000	0.001	-0.000	0.000	-0.000	0.001
	(0.13)	(0.49)	(-0.21)	(0.23)	(-0.20)	(0.76)
_cons	-0.191	-0.514	-0.142	0.014	-0.177	0.231
	(-0.71)	(-1.37)	(-0.47)	(0.11)	(-0.98)	(1.24)
N	3890	1983	1907	3890	1983	1907
r2_p						

Note: Columns 1 to 3 refer to regressions for the probability of being overweight, whilst columns 4 to 6 refer to regressions for the probability of being obese. Columns 1 and 4 refer to the full sample; Columns 2 and 5 to males; and columns 3 and 6 to females. *t* statistics in parentheses. *N* denotes the number of observations. Fair self-reported health status and unemployed are omitted categories. Precise definitions of all variables can be found in Appendix 4.1. Province dummies are included, but their coefficients are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Table 4.5 OLS regressions for BMI

	(1) Total	(2) Males	(3) Females
TP	0.483***	0.651***	0.360*
	(3.38)	(3.41)	(1.84)
male	1.133***		
	(6.75)		
age	0.141***	0.060	0.180***
	(4.02)	(1.36)	(3.39)
age2	-0.001***	-0.001	-0.002***
	(-3.98)	(-1.45)	(-2.91)
rural	-0.064	-0.177	0.066
	(-0.21)	(-0.44)	(0.14)
edu_y	-0.013	0.044	-0.045
	(-0.59)	(1.44)	(-1.44)
fedu_y	-0.013	-0.036*	-0.000
	(-0.80)	(-1.67)	(-0.00)
medu_y	-0.017	-0.004	-0.013
	(-0.90)	(-0.15)	(-0.51)
sedu_y	0.017	0.043*	-0.047*
	(0.97)	(1.81)	(-1.93)
incomepc	0.059	0.245**	-0.181
	(0.60)	(2.00)	(-1.32)
married	-0.066	-0.054	2.563*
	(-0.09)	(-0.07)	(1.87)
slfhvg	-0.162	-0.026	-0.150
	(-0.77)	(-0.08)	(-0.53)
slfhg	0.086	-0.063	0.339
	(0.44)	(-0.22)	(1.31)
slfhp	0.320	0.221	0.321
	(1.22)	(0.62)	(0.90)
depression	-0.535***	-0.503*	-0.631**

	(-2.63)	(-1.88)	(-2.22)
cognitive	0.002	0.017**	-0.000
	(0.42)	(2.29)	(-0.02)
workagri	-0.792**	0.003	-0.937
	(-1.98)	(0.01)	(-1.62)
employed	-0.361*	0.180	-0.386
	(-1.87)	(0.60)	(-1.56)
slfemploy	0.097	0.601*	0.006
	(0.42)	(1.67)	(0.02)
wealthpc	0.065	0.118*	0.000
	(1.30)	(1.73)	(0.00)
foodpc	0.037	-0.071	0.149
	(0.33)	(-0.49)	(0.97)
good_mob	0.395	0.217	0.668
	(1.00)	(0.43)	(1.32)
smoker	-0.073	-0.222	1.016
	(-0.41)	(-1.24)	(1.54)
drinker	0.019	0.039	0.510
	(0.10)	(0.20)	(0.80)
phyact	0.007	0.007	0.012
	(0.82)	(0.62)	(0.92)
_cons	18.299***	17.769***	17.570***
	(11.84)	(9.25)	(7.23)
N	3890	1983	1907
r2	0.083	0.114	0.119

Note: Columns 1 to 3 refer to regressions for the BMI. Columns 1 refer to the full sample; Columns 2 to males; and columns 3 to females. *t* statistics in parentheses. *N* denotes the number of observations. Fair self-reported health status and unemployed are omitted categories. Precise definitions of all variables can be found in Appendix 4.1. Province dummies are included, but their coefficients are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Table 4.6 Quantile regressions for BMI

	(1) 10th	(2) 30th	(3) 50th	(4) 70th	(5) 90th
TP	0.312***	0.474***	0.563***	0.587***	0.672***
	(2.83)	(3.39)	(3.61)	(2.88)	(4.44)
male	1.285***	1.200***	1.129***	1.342***	1.067***
	(8.21)	(7.23)	(7.44)	(6.01)	(5.90)
age	0.180***	0.175***	0.188***	0.110***	0.050
	(4.54)	(5.08)	(5.13)	(2.60)	(1.01)
age2	-0.002***	-0.002***	-0.002***	-0.001***	-0.001*
	(-4.03)	(-4.65)	(-4.97)	(-3.14)	(-1.67)
rural	-0.629	-0.149	-0.156	-0.180	0.457
	(-1.47)	(-0.65)	(-0.54)	(-0.50)	(1.31)
edu_y	0.021	-0.014	-0.021	-0.037	-0.001
	(0.96)	(-0.62)	(-0.95)	(-1.25)	(-0.05)
fedu_y	0.023	0.000	-0.005	-0.012	-0.030
	(1.62)	(0.03)	(-0.31)	(-0.52)	(-1.63)
medu_y	-0.047***	-0.032*	-0.004	-0.000	0.026
	(-2.87)	(-1.85)	(-0.20)	(-0.00)	(1.10)
sedu_y	0.028*	0.025	0.011	0.006	-0.002
	(1.91)	(1.38)	(0.60)	(0.23)	(-0.11)
incomepc	0.113	0.084	0.270***	0.066	-0.060
	(1.53)	(1.20)	(3.00)	(0.50)	(-0.39)
married	-0.063	-1.734	-0.954**	0.562	2.584
	(-0.00)	(-1.35)	(-2.32)	(0.93)	(1.25)
slfhvg	0.348**	0.343*	-0.155	-0.571*	-0.794***
	(2.00)	(1.69)	(-0.71)	(-1.88)	(-3.04)
slfhg	0.409***	0.489***	0.017	-0.145	-0.244
	(2.81)	(2.73)	(0.08)	(-0.52)	(-1.08)

slfhp	0.256	0.771***	0.373	0.294	0.632**
	(0.89)	(3.02)	(1.34)	(0.83)	(2.37)
depression	-0.283**	-0.203	-0.360*	-0.888***	-0.807**
	(-2.17)	(-0.93)	(-1.74)	(-4.03)	(-2.55)
cognitive	0.007	0.003	-0.004	-0.002	-0.001
	(1.08)	(0.56)	(-0.56)	(-0.26)	(-0.13)
workagri	-0.536*	-1.163**	-0.857	-0.709	-0.391
	(-1.74)	(-2.48)	(-1.62)	(-1.37)	(-0.41)
employed	-0.111	-0.374*	-0.516***	-0.466*	-0.442**
	(-0.74)	(-1.85)	(-2.96)	(-1.72)	(-2.18)
slfemploy	0.395	0.114	0.117	-0.176	-0.682***
	(1.63)	(0.43)	(0.50)	(-0.56)	(-2.71)
wealthpc	-0.021	0.112***	0.070	0.142**	0.103
	(-0.62)	(2.91)	(1.39)	(2.04)	(1.43)
foodpc	0.049	0.046	-0.045	0.039	0.143
	(0.57)	(0.43)	(-0.35)	(0.26)	(1.00)
good_mob	0.157	0.503	0.506**	1.342***	0.891***
	(0.52)	(1.35)	(2.02)	(4.57)	(2.95)
smoker	-0.604***	-0.214	0.035	0.026	0.111
	(-3.76)	(-1.42)	(0.17)	(0.11)	(0.62)
drinker	0.297**	-0.244	0.193	0.082	0.038
	(1.97)	(-1.32)	(0.85)	(0.33)	(0.16)
phyact	0.016	0.011	-0.002	-0.011	0.006
	(0.94)	(1.62)	(-0.23)	(-1.29)	(0.59)
_cons	12.949	15.849***	16.399***	18.677***	22.671***
	(0.27)	(8.32)	(10.86)	(9.77)	(8.09)
N	3890	3890	3890	3890	3890
r2					

Note: The mean values of BMI among the total observations are 19.57, 21.83, 23.44, 25.07, and 27.60 in the 10th, 30th, 50th, 70th, and 90th quantiles, respectively. *t* statistics in parentheses. *N* denotes the number of observations. Fair self-reported health status and unemployed are omitted categories. Precise definitions of all variables can be found in Appendix 4.1. Province dummies are included, but their coefficients are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Table 4.7 Quantile regressions for BMI among males

	(1) 10th	(2) 30th	(3) 50th	(4) 70th	(5) 90th
TP	0.310*	0.615***	0.746***	0.654***	0.823***
	(1.92)	(3.96)	(4.05)	(4.02)	(5.77)
age	0.130***	0.109***	0.082	0.036	0.085
	(3.54)	(3.05)	(1.55)	(0.97)	(1.51)
age2	-0.001***	-0.001***	-0.001	-0.000	-0.001**
	(-3.41)	(-2.81)	(-1.63)	(-1.14)	(-1.98)
rural	-0.631*	0.066	-0.078	-0.032	0.219
	(-1.67)	(0.17)	(-0.19)	(-0.11)	(0.54)
edu_y	0.036	0.017	0.019	0.040	0.123***
	(1.17)	(0.77)	(0.59)	(1.55)	(5.16)
fedu_y	-0.027	-0.018	-0.021	-0.029	-0.043**
	(-1.47)	(-1.10)	(-0.92)	(-1.36)	(-2.24)
medu_y	-0.081***	-0.025	0.004	0.016	0.077***
	(-2.98)	(-1.22)	(0.16)	(0.64)	(3.40)
sedu_y	0.083***	0.061***	0.039	0.030	-0.008
	(3.53)	(3.05)	(1.51)	(1.28)	(-0.38)
incomepc	0.244**	0.234**	0.426***	0.279**	0.185*
	(2.38)	(2.03)	(4.39)	(2.53)	(1.73)
married	-1.797	-0.998	0.132	-0.621	2.235
	(-0.09)	(-0.80)	(0.06)	(-0.42)	(0.50)
slfhvg	-0.062	0.387	-0.042	-0.285	-0.632**
	(-0.19)	(1.55)	(-0.15)	(-0.83)	(-2.32)
slfhg	0.063	0.254	0.064	-0.379	-0.561**
	(0.22)	(1.11)	(0.23)	(-1.28)	(-1.98)
slfhp	0.777**	0.504	0.270	-0.008	-0.173
	(2.30)	(1.59)	(0.70)	(-0.02)	(-0.47)
depression	-0.060	0.112	-0.387	-1.076***	-1.171***
	(-0.22)	(0.38)	(-1.57)	(-5.91)	(-3.53)
cognitive	0.013**	0.008*	0.016**	0.018**	0.023***

	(1.99)	(1.85)	(2.09)	(2.56)	(3.24)
workagri	-0.421	-0.322	-0.853*	0.207	0.129
	(-0.59)	(-1.20)	(-1.66)	(0.24)	(0.17)
employed	0.122	-0.066	-0.239	0.475*	-0.351
	(0.47)	(-0.27)	(-1.22)	(1.67)	(-1.03)
slfemploy	0.492*	0.094	0.234	0.591*	-0.089
	(1.76)	(0.25)	(0.72)	(1.84)	(-0.24)
wealthpc	0.029	0.137***	0.222***	0.218***	0.026
	(0.43)	(2.94)	(3.81)	(3.83)	(0.39)
foodpc	-0.064	-0.064	-0.108	-0.133	0.269*
	(-0.56)	(-0.52)	(-0.75)	(-1.06)	(1.71)
good_mob	0.166	-0.098	0.153	-0.019	0.175
	(0.17)	(-0.24)	(0.35)	(-0.05)	(0.40)
smoker	-0.581***	-0.260*	-0.051	-0.131	-0.486***
	(-3.78)	(-1.78)	(-0.30)	(-0.86)	(-3.11)
drinker	0.085	-0.078	0.367*	0.259	-0.053
	(0.52)	(-0.53)	(1.70)	(1.55)	(-0.31)
phyact	0.003	0.000	0.002	-0.005	-0.003
	(0.17)	(0.01)	(0.15)	(-0.38)	(-0.12)
_cons	16.081	16.234***	14.450***	19.606***	18.017***
	(0.79)	(8.09)	(4.78)	(9.23)	(3.76)
N	1983	1983	1983	1983	1983
r2					

Note: The mean values of BMI among males are 20.21, 22.36, 23.94, 25.63, and 27.86 in the 10th, 30th, 50th, 70th, and 90th quantiles, respectively. *t* statistics in parentheses. *N* denotes the number of observations. Fair self-reported health status and unemployed are omitted categories. Precise definitions of all variables can be found in Appendix 4.1. Province dummies are included, but their coefficients are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Table 4.8 Quantile regressions for BMI among females

	(1) 10th	(2) 30th	(3) 50th	(4) 70th	(5) 90th
TP	0.102	0.299	0.387**	0.189	1.013***
	(0.55)	(1.55)	(2.58)	(0.93)	(4.74)
age	0.237***	0.199***	0.218***	0.214***	0.049
	(4.81)	(4.16)	(6.54)	(4.24)	(0.70)
age2	-0.002***	-0.002***	-0.002***	-0.002***	-0.000
	(-4.25)	(-3.67)	(-5.69)	(-3.76)	(-0.48)
rural	0.049	-0.760*	-0.790**	0.737	1.537***
	(0.08)	(-1.88)	(-2.16)	(0.85)	(4.89)
edu_y	-0.025	-0.032	-0.043*	-0.062**	-0.085*
	(-0.97)	(-1.01)	(-1.92)	(-1.99)	(-1.89)
fedu_y	0.047**	-0.003	-0.014	-0.022	-0.025
	(2.43)	(-0.14)	(-0.79)	(-1.18)	(-1.00)
medu_y	-0.016	-0.019	0.012	-0.010	0.001
	(-0.64)	(-0.81)	(0.60)	(-0.41)	(0.03)
sedu_y	-0.022	-0.037	-0.073***	-0.051**	-0.043
	(-1.01)	(-1.43)	(-3.52)	(-1.99)	(-1.26)
incomepc	-0.136	-0.152	-0.120	-0.071	-0.413**
	(-1.55)	(-1.36)	(-1.32)	(-0.46)	(-2.02)
married	0.469	1.837	4.087***	4.705	1.920
	(0.27)	(0.90)	(3.03)	(0.40)	(1.28)
slfhvg	0.614***	0.420*	-0.275	-0.186	-1.175***
	(2.85)	(1.70)	(-1.05)	(-0.61)	(-3.69)
slfhg	0.573***	0.645***	-0.018	0.485*	0.168
	(2.77)	(2.73)	(-0.07)	(1.72)	(0.65)
slfhp	-0.040	0.204	0.593*	1.068***	0.263
	(-0.12)	(0.46)	(1.65)	(3.04)	(0.62)
depression	0.082	-0.279	-0.807***	-0.996**	-1.090***
	(0.32)	(-1.28)	(-3.09)	(-2.35)	(-3.47)
cognitive	0.009	0.005	-0.005	-0.006	-0.010

	(1.12)	(0.63)	(-0.73)	(-0.76)	(-1.09)
workagri	-1.892*	-0.506	-0.617	-1.278**	-0.429
	(-1.77)	(-0.44)	(-1.25)	(-2.29)	(-0.49)
employed	-0.441*	-0.399**	-0.290	-0.539**	0.015
	(-1.76)	(-2.15)	(-1.50)	(-2.30)	(0.05)
slfemploy	-0.067	0.182	-0.056	-0.059	-0.421
	(-0.18)	(0.70)	(-0.25)	(-0.18)	(-1.04)
wealthpc	-0.007	-0.070	0.017	0.058	0.099
	(-0.09)	(-1.46)	(0.29)	(0.88)	(1.31)
foodpc	0.139	0.251	0.190	0.064	0.327*
	(1.14)	(1.58)	(1.64)	(0.53)	(1.73)
good_mob	-0.110	0.275	1.585***	2.229***	1.088
	(-0.26)	(0.51)	(6.10)	(7.25)	(0.48)
smoker	0.142	0.536	1.562***	2.157***	2.188**
	(0.37)	(0.55)	(2.60)	(4.68)	(2.16)
drinker	-0.020	0.059	0.865***	0.931	2.770
	(-0.06)	(0.04)	(2.77)	(1.23)	(1.51)
phyact	0.036***	0.012	-0.001	-0.020	-0.022
	(2.97)	(1.23)	(-0.10)	(-1.20)	(-1.09)
_cons	13.880	15.392***	13.204***	14.457	25.190***
	(.)	(5.55)	(7.00)	(1.20)	(6.28)
N	1907	1907	1907	1907	1907
r2					

Note: The mean values of BMI among females are 19.40, 21.31, 22.91, 24.49, and 27.30 in the 10th, 30th, 50th, 70th, and 90th, respectively. *t* statistics in parentheses. *N* denotes the number of observations. Fair self-reported health status and unemployed are omitted categories. Precise definitions of all variables can be found in Appendix 4.1. Province dummies are included, but their coefficients are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Table 4.9 Ordered Probit models for weight categories

	(1) underweight	(2) normal	(3) overweight	(4) obese
TP	-0.013***	-0.051***	0.039***	0.026***
	(-3.10)	(-3.22)	(3.22)	(3.17)
male	-0.031***	-0.118***	0.090***	0.059***
	(-5.40)	(-6.30)	(6.18)	(6.01)
age	-0.004***	-0.014***	0.011***	0.007***
	(-3.37)	(-3.63)	(3.57)	(3.58)
age2	0.000***	0.000***	-0.000***	-0.000***
	(3.45)	(3.71)	(-3.66)	(-3.65)
rural	-0.002	-0.009	0.007	0.004
	(-0.27)	(-0.27)	(0.27)	(0.27)
edu_y	0.000	0.002	-0.001	-0.001
	(0.62)	(0.62)	(-0.62)	(-0.62)
fedu_y	0.000	0.000	-0.000	-0.000
	(0.23)	(0.23)	(-0.23)	(-0.23)
medu_y	0.000	0.002	-0.001	-0.001
	(0.76)	(0.77)	(-0.77)	(-0.77)
sedu_y	-0.000	-0.001	0.001	0.001
	(-0.64)	(-0.64)	(0.64)	(0.64)
incomepc	-0.003	-0.011	0.009	0.006
	(-1.03)	(-1.03)	(1.03)	(1.03)
married	0.020	0.076	-0.057	-0.038
	(1.01)	(1.02)	(-1.02)	(-1.02)
slfhvg	0.011*	0.042*	-0.032*	-0.021*
	(1.78)	(1.80)	(-1.81)	(-1.79)
slfhg	0.001	0.004	-0.003	-0.002
	(0.19)	(0.19)	(-0.19)	(-0.19)
slfhp	-0.003	-0.011	0.008	0.005
	(-0.37)	(-0.37)	(0.37)	(0.37)
depression	0.012**	0.047**	-0.036**	-0.024**

	(2.11)	(2.10)	(-2.12)	(-2.07)
cognitive	-0.000	-0.000	0.000	0.000
	(-0.09)	(-0.09)	(0.09)	(0.09)
workagri	0.019	0.072	-0.054	-0.036
	(1.61)	(1.63)	(-1.63)	(-1.62)
employed	0.013**	0.051**	-0.039**	-0.026**
	(2.36)	(2.43)	(-2.42)	(-2.41)
slfemploy	0.002	0.006	-0.005	-0.003
	(0.23)	(0.23)	(-0.23)	(-0.23)
wealthpc	-0.002	-0.009	0.007	0.004
	(-1.61)	(-1.62)	(1.61)	(1.62)
foodpc	0.000	0.000	-0.000	-0.000
	(0.03)	(0.03)	(-0.03)	(-0.03)
good_mob	-0.014	-0.055	0.042	0.028
	(-1.32)	(-1.32)	(1.32)	(1.33)
smoker	0.002	0.008	-0.006	-0.004
	(0.40)	(0.40)	(-0.40)	(-0.40)
drinker	-0.003	-0.013	0.010	0.007
	(-0.62)	(-0.62)	(0.62)	(0.62)
phyact	-0.000	-0.001	0.001	0.000
	(-0.74)	(-0.74)	(0.74)	(0.74)
N	3890	3890	3890	3890
r2_p				

Note: Columns 1 to 4 refer to regressions for the underweight, normal weight, overweight, and obese. Marginal effects are reported, with *z* statistics in parentheses. *N* denotes the number of observations. Fair self-reported health status and unemployed are omitted categories. Precise definitions of all variables can be found in Appendix 4.1. Province dummies are included, but their marginal effects are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Table 4.10 Ordered Probit models for weight categories among males

	(1) underweight	(2) normal	(3) overweight	(4) obese
TP	-0.011***	-0.074***	0.048***	0.037***
	(-2.67)	(-3.07)	(3.03)	(3.01)
age	-0.001	-0.006	0.004	0.003
	(-0.98)	(-0.98)	(0.98)	(0.99)
age2	0.000	0.000	-0.000	-0.000
	(1.17)	(1.17)	(-1.17)	(-1.17)
rural	-0.002	-0.012	0.008	0.006
	(-0.23)	(-0.23)	(0.23)	(0.23)
edu_y	-0.001	-0.005	0.003	0.003
	(-1.22)	(-1.27)	(1.27)	(1.25)
fedu_y	0.000	0.003	-0.002	-0.001
	(0.96)	(0.97)	(-0.97)	(-0.96)
medu_y	-0.000	-0.000	0.000	0.000
	(-0.14)	(-0.14)	(0.14)	(0.14)
sedu_y	-0.001	-0.004	0.002	0.002
	(-1.08)	(-1.10)	(1.10)	(1.10)
incomepc	-0.007***	-0.044***	0.028***	0.022***
	(-2.71)	(-2.85)	(2.84)	(2.84)
married	0.011	0.071	-0.046	-0.036
	(0.56)	(0.57)	(-0.56)	(-0.57)
slfhvg	0.007	0.047	-0.030	-0.024
	(1.14)	(1.21)	(-1.20)	(-1.19)
slfhg	0.006	0.040	-0.026	-0.020
	(1.09)	(1.12)	(-1.12)	(-1.11)
slfhp	0.004	0.023	-0.015	-0.012
	(0.50)	(0.50)	(-0.50)	(-0.50)
depression	0.008	0.055	-0.036	-0.028
	(1.56)	(1.57)	(-1.59)	(-1.55)
cognitive	-0.000	-0.002	0.001	0.001

	(-1.59)	(-1.64)	(1.64)	(1.63)
workagri	-0.000	-0.003	0.002	0.001
	(-0.03)	(-0.03)	(0.03)	(0.03)
employed	-0.002	-0.016	0.011	0.008
	(-0.43)	(-0.44)	(0.44)	(0.44)
slfemploy	-0.012	-0.078*	0.050*	0.040*
	(-1.62)	(-1.77)	(1.74)	(1.75)
wealthpc	-0.002*	-0.015*	0.010*	0.008*
	(-1.77)	(-1.81)	(1.82)	(1.80)
foodpc	0.002	0.013	-0.009	-0.007
	(0.71)	(0.71)	(-0.71)	(-0.71)
good_mob	-0.006	-0.042	0.027	0.021
	(-0.71)	(-0.71)	(0.71)	(0.71)
smoker	0.005	0.032	-0.020	-0.016
	(1.33)	(1.39)	(-1.38)	(-1.38)
drinker	-0.002	-0.014	0.009	0.007
	(-0.57)	(-0.56)	(0.56)	(0.56)
phyact	-0.000	-0.001	0.001	0.000
	(-0.59)	(-0.60)	(0.59)	(0.60)
N	1983	1983	1983	1983
r2_p				

Note: Columns 1 to 4 refer to regressions for the underweight, normal weight, overweight, and obese. Marginal effects are reported, with *z* statistics in parentheses. *N* denotes the number of observations. Fair self-reported health status and unemployed are omitted categories. Precise definitions of all variables can be found in Appendix 4.1. Province dummies are included, but their marginal effects are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Table 4.11 Ordered Probit models for weight categories among females

	(1) underweight	(2) normal	(3) overweight	(4) obese
TP	-0.016**	-0.038**	0.034**	0.019**
	(-2.02)	(-2.05)	(2.06)	(2.02)
age	-0.007***	-0.016***	0.015***	0.008***
	(-2.97)	(-3.29)	(3.19)	(3.21)
age2	0.000***	0.000***	-0.000***	-0.000***
	(2.70)	(2.92)	(-2.85)	(-2.88)
rural	-0.003	-0.008	0.008	0.004
	(-0.21)	(-0.21)	(0.21)	(0.21)
edu_y	0.002	0.005	-0.004	-0.002
	(1.56)	(1.55)	(-1.56)	(-1.56)
fedu_y	-0.000	-0.000	0.000	0.000
	(-0.22)	(-0.22)	(0.22)	(0.22)
medu_y	0.001	0.002	-0.002	-0.001
	(0.75)	(0.75)	(-0.75)	(-0.75)
sedu_y	0.001	0.003	-0.003	-0.002
	(1.43)	(1.44)	(-1.44)	(-1.44)
incomepc	0.007	0.016	-0.015	-0.008
	(1.31)	(1.32)	(-1.32)	(-1.30)
married	-0.054***	-0.131***	0.119***	0.066***
	(-4.06)	(-3.58)	(3.73)	(3.72)
slfhvg	0.010	0.025	-0.022	-0.012
	(0.96)	(0.95)	(-0.95)	(-0.95)
slfhg	-0.013	-0.031	0.028	0.016
	(-1.29)	(-1.32)	(1.32)	(1.30)
slfhp	-0.010	-0.023	0.021	0.012
	(-0.69)	(-0.69)	(0.69)	(0.69)
depression	0.018	0.044	-0.040	-0.022
	(1.62)	(1.61)	(-1.63)	(-1.59)
cognitive	0.000	0.000	-0.000	-0.000

	(0.05)	(0.05)	(-0.05)	(-0.05)
workagri	0.024	0.057	-0.052	-0.029
	(1.17)	(1.18)	(-1.18)	(-1.17)
employed	0.020**	0.048**	-0.043**	-0.024**
	(2.00)	(2.08)	(-2.07)	(-2.05)
slfemploy	0.010	0.025	-0.023	-0.013
	(0.88)	(0.88)	(-0.88)	(-0.88)
wealthpc	-0.001	-0.003	0.003	0.001
	(-0.45)	(-0.44)	(0.44)	(0.44)
foodpc	-0.004	-0.010	0.009	0.005
	(-0.67)	(-0.67)	(0.67)	(0.67)
good_mob	-0.027	-0.066	0.060	0.033
	(-1.43)	(-1.41)	(1.42)	(1.42)
smoker	-0.041*	-0.099*	0.090*	0.050*
	(-1.82)	(-1.84)	(1.85)	(1.82)
drinker	-0.024	-0.059	0.054	0.030
	(-0.99)	(-0.99)	(0.99)	(0.98)
phyact	-0.000	-0.001	0.001	0.000
	(-0.64)	(-0.64)	(0.64)	(0.64)
N	1907	1907	1907	1907
r2_p				

Note: Columns 1 to 4 refer to regressions for the underweight, normal weight, overweight, and obese. Marginal effects are reported, with *z* statistics in parentheses. *N* denotes the number of observations. Fair self-reported health status and unemployed are omitted categories. Precise definitions of all variables can be found in Appendix 4.1. Province dummies are included, but their marginal effects are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Appendices

This section contains the definitions of variables and the results of robustness tests. Appendix 4.1 describes the definitions of all variables used in this chapter. Appendix 4.2 presents the empirical results of the robustness tests using an alternative measure of time preference with *TPc*.

Appendix 4.1 Definitions of variables used

Variables	Attribute	Definition
overweight	Dummy variable	Overweight according to WGOC [1]; Otherwise [0]
obesity	Dummy variable	Obese according to WGOC [1]; Otherwise [0]
BMI	Continuous variable	Mass (kg)/Height squared (m square)
order_bmi	Category variable	Underweight [1], Normal weight [2], Overweight [3], Obesity [4]
TP	Dummy variable	High time preference [1]; low time preference [0]
male	Dummy variable	Male [1]; Female [0]

age	Continuous variable	Age
age2	Continuous variable	Age squared
rural	Dummy variable	Living in rural area[1]; Living in urban area [0]
edu_y	Continuous variable	Years of education for the respondent
fedu_y	Continuous variable	Years of education for the respondent's father
medu_y	Continuous variable	Years of education for the respondent's mother
sedu_y	Continuous variable	Years of education for the respondent's spouse
incomepc	Continuous variable	Logarithm of the income per capita
rincomepc	Continuous variable	Amount of the income per capita by YUAN
married	Dummy variable	Married/ co-habiting [1]; Otherwise [0]
depression	Dummy variable	Moderate or high psychological distress [1]; Low psychological distress [0]
cognitive	Continuous variable	Sum of correct answers of word test and math test
slfhvg	Dummy variable	Very good health status[1]; Otherwise [0]
slfhg	Dummy variable	Current smoker [1]; Otherwise [0]
slfhf	Dummy variable	Fair health status[1]; Otherwise [0]

slfhp	Dummy variable	Poor health status[1]; Otherwise [0]
Unemployd	Dummy variable	Unemployed[1]; Otherwise [0]
workagri	Dummy variable	Works in agriculture[1]; Otherwise [0]
employed	Dummy variable	Employed[1]; Otherwise [0]
slfemploy	Dummy variable	Self-employed[1]; Otherwise [0]
wealthpc	Continuous variable	Log of family wealth per capita
rwealthpc	Continuous variable	Amount of family wealth per capita by YUAN
foodpc	Continuous variable	Log of consumption on food per capita
rfoodpc	Continuous variable	Amount of consumption on food per capita by YUAN
good_mob	Dummy variable	Having no problem with carrying out any of the daily activities [1]; Otherwise [0]
smoker	Dummy variable	Employed[1]; Otherwise [0]
drinker	Dummy variable	Drink alcohol at least 3 times a week in the past month [1]; Otherwise [0]
phyact	Continuous variable	Time spent doing physical exercises weekly (hour)

Appendix 4.2 Robustness tests using an alternative measure of time preferences

Table A. Probit models for the probability of being overweight/obese (alternative measure of time preferences with *TPc*)

	(1) Tover	(2) Mover	(3) Fover	(4) Tobes	(5) Mobes	(6) Fobes
TPc	0.021***	0.027**	0.017*	0.009**	0.008	0.010**
	(2.64)	(2.44)	(1.66)	(2.02)	(1.14)	(2.00)
male	0.158***			0.035**		
	(6.04)			(2.22)		
age	0.017***	0.011	0.020***	-0.002	-0.002	-0.002
	(3.13)	(1.40)	(2.61)	(-0.61)	(-0.41)	(-0.47)
age2	-0.000***	-0.000	-0.000**	0.000	-0.000	0.000
	(-3.13)	(-1.39)	(-2.38)	(0.14)	(-0.26)	(0.41)
rural	0.003	0.023	-0.012	0.004	0.008	0.014
	(0.06)	(0.33)	(-0.18)	(0.15)	(0.21)	(0.45)
edu_y	-0.003	0.003	-0.007	-0.000	0.003	-0.003
	(-0.90)	(0.57)	(-1.50)	(-0.21)	(0.90)	(-1.02)
fedu_y	-0.001	-0.002	-0.000	-0.001	-0.003	-0.000
	(-0.20)	(-0.48)	(-0.08)	(-0.68)	(-1.32)	(-0.25)
medu_y	-0.002	0.003	-0.005	0.000	0.001	-0.000
	(-0.62)	(0.69)	(-1.20)	(0.18)	(0.58)	(-0.01)
sedu_y	0.001	0.005	-0.006	0.000	-0.001	-0.001
	(0.36)	(1.19)	(-1.49)	(0.20)	(-0.44)	(-0.46)
incomepc	0.030**	0.077***	-0.017	-0.003	0.016	-0.016*
	(2.01)	(3.40)	(-0.95)	(-0.36)	(1.21)	(-1.76)
married	-0.117	-0.167	0.000	0.076	0.124	0.000
	(-0.72)	(-0.79)	(.)	(1.00)	(1.37)	(.)
slfhvg	-0.044	-0.014	-0.052	-0.036**	-0.051*	-0.028
	(-1.37)	(-0.29)	(-1.27)	(-1.98)	(-1.86)	(-1.18)

slfhg	0.005	-0.009	0.030	-0.019	-0.051**	0.011
	(0.18)	(-0.20)	(0.82)	(-1.18)	(-2.07)	(0.56)
slfhp	0.064	0.029	0.076	0.003	-0.045	0.029
	(1.62)	(0.48)	(1.57)	(0.13)	(-1.30)	(1.19)
depression	-0.056	-0.039	-0.076	-0.087***	-0.106***	-0.074***
	(-1.57)	(-0.75)	(-1.63)	(-3.81)	(-2.97)	(-2.85)
cognitive	0.000	0.002	-0.000	-0.000	0.001	-0.000
	(0.20)	(1.58)	(-0.27)	(-0.06)	(1.06)	(-0.04)
workagri	-0.096	-0.029	-0.063	-0.030	0.038	-0.059
	(-1.46)	(-0.30)	(-0.75)	(-0.86)	(0.70)	(-1.15)
employed	-0.068**	0.002	-0.067*	-0.010	0.002	-0.006
	(-2.31)	(0.03)	(-1.94)	(-0.55)	(0.08)	(-0.28)
slfemploy	0.006	0.096	-0.022	-0.014	0.002	-0.013
	(0.15)	(1.58)	(-0.48)	(-0.65)	(0.05)	(-0.59)
wealthpc	0.015*	0.019*	0.010	0.002	0.009	-0.002
	(1.94)	(1.65)	(0.94)	(0.47)	(1.30)	(-0.40)
foodpc	-0.024	-0.034	-0.012	0.015	0.009	0.018
	(-1.32)	(-1.34)	(-0.53)	(1.39)	(0.65)	(1.38)
good_mob	0.130**	0.073	0.196**	0.003	0.057	0.001
	(2.06)	(0.83)	(2.52)	(0.08)	(0.76)	(0.03)
smoker	-0.011	-0.031	0.157**	0.004	-0.011	0.058
	(-0.41)	(-1.04)	(2.17)	(0.24)	(-0.60)	(1.59)
drinker	0.027	0.031	0.050	0.008	0.001	0.040
	(0.91)	(0.96)	(0.60)	(0.52)	(0.08)	(0.93)
phyact	0.000	0.001	-0.000	0.000	-0.000	0.001
	(0.13)	(0.54)	(-0.19)	(0.22)	(-0.09)	(1.00)
N	3890	1983	1904	3846	1959	1884
r2_p						

Note: Columns 1 to 3 refer to regressions for the probability of being overweight, whilst columns 4 to 6 refer to regressions for the probability of being obese. Columns 1 and 4 refer to the full sample; Columns 2 and 5 to males; and columns 3 and 6 to females. Marginal effects are reported, with *z* statistics in parentheses. *N* denotes the number of observations. Fair self-reported health status and unemployed are omitted categories. Precise definitions of all variables can be found in Appendix 4.1. Province dummies are included, but their marginal effects are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Table B. Linear probability models for the probability of being overweight/obese (alternative measure of time preferences with *TPc*)

	(1) Tover	(2) Mover	(3) Fover	(4) Tobes	(5) Mobes	(6) Fobes
TPc	0.021***	0.027**	0.017	0.010*	0.009	0.011*
	(2.63)	(2.44)	(1.58)	(1.96)	(1.19)	(1.86)
male	0.160***			0.034**		
	(5.82)			(2.01)		
age	0.017***	0.010	0.020**	-0.002	-0.004	-0.002
	(3.18)	(1.31)	(2.58)	(-0.49)	(-0.95)	(-0.42)
age2	-0.000***	-0.000	-0.000**	0.000	0.000	0.000
	(-3.17)	(-1.29)	(-2.28)	(0.03)	(0.43)	(0.39)
rural	0.001	0.023	-0.018	0.004	-0.003	0.010
	(0.03)	(0.32)	(-0.26)	(0.19)	(-0.09)	(0.31)
edu_y	-0.003	0.003	-0.006	-0.000	0.003	-0.003
	(-0.85)	(0.52)	(-1.26)	(-0.09)	(1.07)	(-1.07)
fedu_y	-0.000	-0.002	-0.001	-0.001	-0.003	-0.001
	(-0.18)	(-0.44)	(-0.15)	(-0.67)	(-0.99)	(-0.38)
medu_y	-0.002	0.003	-0.004	0.001	0.002	-0.000
	(-0.60)	(0.70)	(-1.12)	(0.28)	(0.82)	(-0.13)
sedu_y	0.001	0.005	-0.006	0.000	-0.000	-0.000
	(0.42)	(1.23)	(-1.47)	(0.26)	(-0.18)	(-0.12)
incomepc	0.031**	0.075***	-0.018	-0.002	0.013	-0.021
	(2.05)	(3.44)	(-0.96)	(-0.24)	(1.14)	(-1.48)
married	-0.118	-0.165	0.355***	0.062	0.124*	0.058**
	(-0.73)	(-0.79)	(5.16)	(1.44)	(1.90)	(2.21)
slfhvg	-0.043	-0.012	-0.049	-0.039*	-0.057	-0.020
	(-1.33)	(-0.25)	(-1.17)	(-1.95)	(-1.64)	(-0.88)
slfhg	0.005	-0.008	0.031	-0.023	-0.058*	0.015
	(0.18)	(-0.18)	(0.81)	(-1.24)	(-1.90)	(0.69)
slfhp	0.066	0.028	0.084	0.000	-0.044	0.029
	(1.63)	(0.46)	(1.59)	(0.02)	(-1.21)	(0.98)

depression	-0.057	-0.036	-0.080*	-0.065***	-0.073***	-0.060***
	(-1.59)	(-0.70)	(-1.67)	(-5.09)	(-3.98)	(-3.23)
cognitive	0.000	0.002	-0.000	-0.000	0.001	-0.000
	(0.19)	(1.60)	(-0.31)	(-0.04)	(1.01)	(-0.04)
workagri	-0.100	-0.031	-0.073	-0.024	0.036	-0.052
	(-1.51)	(-0.33)	(-0.79)	(-0.88)	(0.78)	(-1.52)
employed	-0.069**	0.005	-0.072*	-0.009	0.007	-0.004
	(-2.30)	(0.10)	(-1.95)	(-0.54)	(0.30)	(-0.17)
slfemploy	0.006	0.098	-0.024	-0.010	0.011	-0.014
	(0.17)	(1.60)	(-0.49)	(-0.48)	(0.31)	(-0.57)
wealthpc	0.015*	0.019*	0.010	0.002	0.006	-0.002
	(1.95)	(1.67)	(0.91)	(0.42)	(1.10)	(-0.36)
foodpc	-0.024	-0.034	-0.012	0.014	0.006	0.022
	(-1.30)	(-1.31)	(-0.51)	(1.31)	(0.46)	(1.54)
good_mob	0.126**	0.062	0.197***	0.009	0.019	-0.001
	(2.07)	(0.72)	(2.64)	(0.29)	(0.48)	(-0.03)
smoker	-0.012	-0.031	0.174**	0.007	-0.007	0.108
	(-0.42)	(-1.01)	(2.18)	(0.35)	(-0.34)	(1.42)
drinker	0.029	0.032	0.058	0.010	0.004	0.061
	(0.93)	(1.00)	(0.65)	(0.52)	(0.22)	(0.81)
phyact	0.000	0.001	-0.000	0.000	-0.000	0.001
	(0.15)	(0.50)	(-0.20)	(0.24)	(-0.19)	(0.76)
_cons	-0.243	-0.561	-0.203	-0.011	-0.192	0.191
	(-0.90)	(-1.51)	(-0.66)	(-0.09)	(-1.07)	(1.01)
N	3890	1983	1907	3890	1983	1907
r2_p						

Note: Columns 1 to 3 refer to regressions for the probability of being overweight, whilst columns 4 to 6 refer to regressions for the probability of being obese. Columns 1 and 4 refer to the full sample; Columns 2 and 5 to males; and columns 3 and 6 to females. *t* statistics in parentheses. *N* denotes the number of observations. Fair self-reported health status and unemployed are omitted categories. Precise definitions of all variables can be found in Appendix 4.1. Province dummies are included, but their coefficients are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Table C. OLS regressions for BMI (alternative measure of time preferences with *TPc*)

	(1) Total	(2) Males	(3) Females
TPc	0.151***	0.185***	0.130*
	(3.07)	(2.76)	(1.96)
male	1.138***		
	(6.78)		
age	0.143***	0.061	0.183***
	(4.07)	(1.37)	(3.45)
age2	-0.001***	-0.001	-0.002***
	(-4.04)	(-1.47)	(-2.97)
rural	-0.069	-0.188	0.066
	(-0.23)	(-0.47)	(0.14)
edu_y	-0.014	0.040	-0.043
	(-0.60)	(1.30)	(-1.39)
fedu_y	-0.013	-0.034	-0.001
	(-0.78)	(-1.58)	(-0.02)
medu_y	-0.017	-0.006	-0.012
	(-0.92)	(-0.22)	(-0.48)
sedu_y	0.016	0.044*	-0.049**
	(0.92)	(1.84)	(-2.01)
incomepc	0.064	0.252**	-0.177
	(0.65)	(2.06)	(-1.30)
married	0.025	0.040	2.714**
	(0.03)	(0.05)	(1.98)
slfhvg	-0.160	-0.038	-0.141
	(-0.76)	(-0.12)	(-0.50)
slfhg	0.088	-0.069	0.341
	(0.45)	(-0.23)	(1.31)
slfhp	0.314	0.189	0.323
	(1.20)	(0.53)	(0.91)
depression	-0.549***	-0.514*	-0.645**

	(-2.71)	(-1.93)	(-2.28)
cognitive	0.003	0.018**	-0.000
	(0.46)	(2.36)	(-0.01)
workagri	-0.795**	-0.029	-0.925
	(-1.99)	(-0.05)	(-1.59)
employed	-0.360*	0.175	-0.385
	(-1.87)	(0.59)	(-1.55)
slfemploy	0.110	0.606*	0.022
	(0.47)	(1.69)	(0.07)
wealthpc	0.062	0.115*	-0.002
	(1.25)	(1.70)	(-0.04)
foodpc	0.034	-0.081	0.148
	(0.31)	(-0.56)	(0.96)
good_mob	0.397	0.225	0.667
	(1.01)	(0.43)	(1.33)
smoker	-0.073	-0.215	1.003
	(-0.40)	(-1.20)	(1.51)
drinker	0.007	0.021	0.513
	(0.04)	(0.11)	(0.80)
phyact	0.007	0.007	0.012
	(0.85)	(0.64)	(0.93)
_cons	17.917***	17.451***	17.097***
	(11.54)	(9.16)	(6.98)
N	3890	1983	1907
r2	0.082	0.111	0.119

Note: Columns 1 to 3 refer to regressions for the BMI. Columns 1 refer to the full sample; Columns 2 to males; and columns 3 to females. *t* statistics in parentheses. *N* denotes the number of observations. Fair self-reported health status and unemployed are omitted categories. Precise definitions of all variables can be found in Appendix 4.1. Province dummies are included, but their coefficients are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Table D. Quantile regressions for BMI (alternative measure of time preferences with *TPc*)

	(1) 10th	(2) 30th	(3) 50th	(4) 70th	(5) 90th
TPc	0.068*	0.161***	0.179***	0.153***	0.221***
	(1.75)	(3.56)	(3.41)	(2.65)	(4.61)
male	1.291***	1.251***	1.081***	1.331***	1.134***
	(9.16)	(7.94)	(6.54)	(6.73)	(6.33)
age	0.183***	0.188***	0.174***	0.116***	0.067
	(4.50)	(5.88)	(4.66)	(3.05)	(1.37)
age2	-0.002***	-0.002***	-0.002***	-0.001***	-0.001**
	(-3.94)	(-5.45)	(-4.65)	(-3.71)	(-2.11)
rural	-0.560	-0.148	-0.242	-0.222	0.449
	(-1.47)	(-0.61)	(-0.60)	(-0.78)	(1.25)
edu_y	0.019	-0.009	-0.035	-0.047**	-0.006
	(0.99)	(-0.43)	(-1.41)	(-2.24)	(-0.23)
fedu_y	0.023*	-0.005	-0.009	-0.008	-0.030*
	(1.81)	(-0.36)	(-0.52)	(-0.38)	(-1.96)
medu_y	-0.045***	-0.032*	0.001	-0.009	0.017
	(-2.79)	(-1.87)	(0.07)	(-0.35)	(0.80)
sedu_y	0.033**	0.019	0.008	0.010	0.001
	(2.28)	(1.04)	(0.44)	(0.47)	(0.05)
incomepc	0.120*	0.051	0.260***	0.012	-0.110
	(1.94)	(0.66)	(2.88)	(0.09)	(-0.76)
married	0.002	-1.658**	-0.891	0.675	3.051
	(.)	(-1.99)	(-0.80)	(0.51)	(1.38)
slfhvg	0.311*	0.321	-0.172	-0.510**	-0.825***
	(1.69)	(1.65)	(-0.70)	(-2.00)	(-3.09)
slfhg	0.427***	0.479***	-0.069	-0.111	-0.160
	(2.68)	(2.63)	(-0.29)	(-0.48)	(-0.68)
slfhp	0.334	0.707**	0.318	0.200	0.712***
	(1.31)	(2.57)	(1.12)	(0.72)	(2.65)

depression	-0.317*	-0.235	-0.324	-0.775***	-1.185***
	(-1.72)	(-1.24)	(-1.60)	(-4.04)	(-5.03)
cognitive	0.005	0.003	-0.004	0.001	-0.003
	(0.81)	(0.57)	(-0.57)	(0.09)	(-0.50)
workagri	-0.423	-1.201**	-0.923*	-0.732	-0.307
	(-1.05)	(-2.55)	(-1.96)	(-1.37)	(-0.31)
employed	-0.129	-0.410**	-0.387**	-0.467*	-0.363*
	(-0.88)	(-2.04)	(-2.01)	(-1.89)	(-1.70)
slfemploy	0.399*	0.143	0.161	-0.028	-0.567**
	(1.66)	(0.58)	(0.66)	(-0.10)	(-2.54)
wealthpc	-0.028	0.110***	0.065	0.112**	0.111*
	(-0.88)	(2.88)	(1.19)	(2.04)	(1.70)
foodpc	0.035	0.046	0.042	0.052	0.171
	(0.38)	(0.49)	(0.33)	(0.37)	(1.14)
good_mob	0.215	0.515	0.570***	1.363***	0.696**
	(1.02)	(1.23)	(2.64)	(5.69)	(2.18)
smoker	-0.576***	-0.252*	-0.006	0.040	0.110
	(-4.14)	(-1.73)	(-0.03)	(0.20)	(0.62)
drinker	0.220*	-0.326*	0.239	0.041	-0.041
	(1.69)	(-1.87)	(0.98)	(0.19)	(-0.21)
phyact	0.023	0.010**	-0.000	-0.010	0.013
	(1.28)	(2.09)	(-0.01)	(-0.89)	(0.83)
_cons	12.952***	15.608***	15.666***	19.059***	21.874***
	(9.27)	(10.05)	(8.22)	(9.42)	(7.40)
N	3890	3890	3890	3890	3890
r2					

Note: The mean values of BMI among the total observations are 19.58, 21.83, 23.44, 25.07, and 27.60 in the 10th, 30th, 50th, 70th, and 90th quantiles, respectively. *t* statistics in parentheses. *N* denotes the number of observations. Fair self-reported health status and unemployed are omitted categories. Precise definitions of all variables can be found in Appendix 4.1. Province dummies are included, but their coefficients are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Table E. Quantile regressions for BMI among males (alternative measure of time preferences with *TPc*)

	(1) 10th	(2) 30th	(3) 50th	(4) 70th	(5) 90th
TPc	0.125	0.187***	0.212***	0.158***	0.175**
	(1.61)	(3.02)	(3.09)	(2.71)	(2.52)
age	0.118***	0.100**	0.050	0.075*	0.072
	(2.74)	(2.39)	(0.86)	(1.77)	(1.42)
age2	-0.001***	-0.001**	-0.001	-0.001*	-0.001**
	(-2.59)	(-2.22)	(-1.03)	(-1.89)	(-2.03)
rural	-0.567	-0.170	-0.005	-0.124	0.488
	(-1.23)	(-0.46)	(-0.01)	(-0.42)	(1.19)
edu_y	0.047	0.014	0.016	0.022	0.119***
	(1.14)	(0.60)	(0.51)	(0.85)	(4.08)
fedu_y	-0.020	-0.007	-0.021	-0.033*	-0.064***
	(-0.99)	(-0.37)	(-0.94)	(-1.67)	(-2.60)
medu_y	-0.091***	-0.027	-0.013	0.026	0.067**
	(-2.74)	(-1.37)	(-0.46)	(1.06)	(2.39)
sedu_y	0.082***	0.071***	0.033	0.035	0.023
	(3.38)	(3.28)	(1.26)	(1.58)	(0.86)
incomepc	0.262**	0.218*	0.473***	0.258**	0.140
	(2.13)	(1.80)	(3.80)	(1.99)	(1.08)
married	-1.743	-1.000	0.307	-0.545	2.281*
	(-0.29)	(-0.64)	(0.16)	(-0.65)	(1.89)
slfhvg	0.044	0.415	-0.128	-0.195	-0.946***
	(0.12)	(1.46)	(-0.42)	(-0.53)	(-2.74)
slfhg	0.120	0.389	-0.019	-0.318	-0.834***
	(0.36)	(1.39)	(-0.06)	(-0.94)	(-2.65)
slfhp	0.840**	0.469	0.192	-0.055	-0.164
	(2.03)	(1.19)	(0.52)	(-0.13)	(-0.38)
depression	-0.005	0.245	-0.362	-1.058***	-1.171***
	(-0.01)	(0.82)	(-1.50)	(-5.68)	(-4.24)
cognitive	0.014**	0.009	0.014*	0.023***	0.019***

	(2.08)	(1.37)	(1.88)	(3.25)	(2.59)
workagri	-0.229	-0.182	-1.008*	-0.130	0.389
	(-0.29)	(-0.60)	(-1.72)	(-0.19)	(0.66)
employed	0.199	-0.146	-0.290	0.360	-0.052
	(0.75)	(-0.53)	(-1.16)	(1.16)	(-0.13)
slfemploy	0.680**	0.155	0.135	0.330	0.190
	(2.01)	(0.39)	(0.40)	(0.92)	(0.40)
wealthpc	0.042	0.107**	0.210***	0.226***	0.047
	(0.53)	(2.12)	(3.22)	(3.82)	(0.61)
foodpc	-0.024	-0.096	-0.121	-0.165	0.201
	(-0.17)	(-1.01)	(-0.77)	(-1.26)	(1.36)
good_mob	0.304	0.024	0.080	-0.233	0.105
	(0.28)	(0.07)	(0.20)	(-0.70)	(0.20)
smoker	-0.572***	-0.254*	-0.075	-0.078	-0.426**
	(-3.19)	(-1.65)	(-0.41)	(-0.50)	(-2.22)
drinker	0.076	-0.118	0.324	0.262	-0.100
	(0.42)	(-0.76)	(1.50)	(1.58)	(-0.49)
phyact	0.004	0.004	0.003	0.002	0.001
	(0.18)	(0.30)	(0.24)	(0.23)	(0.05)
_cons	14.837**	16.473***	15.159***	19.150***	19.082***
	(2.37)	(7.86)	(5.55)	(10.47)	(7.63)
N	1983	1983	1983	1983	1983
r2					

Note: The mean values of BMI among males are 20.21, 22.36, 23.94, 25.63, and 27.86 in the 10th, 30th, 50th, 70th, and 90th quantiles, respectively. *t* statistics in parentheses. *N* denotes the number of observations. Fair self-reported health status and unemployed are omitted categories. Precise definitions of all variables can be found in Appendix 4.1. Province dummies are included, but their coefficients are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Table F. Quantile regressions for BMI among females (alternative measure of time preferences with *TPc*)

	(1) 10th	(2) 30th	(3) 50th	(4) 70th	(5) 90th
TPc	0.025	0.088	0.137***	0.067	0.349***
	(0.33)	(1.20)	(2.94)	(0.97)	(3.61)
age	0.233***	0.213***	0.227***	0.211***	0.034
	(3.96)	(4.06)	(6.53)	(4.52)	(0.41)
age2	-0.002***	-0.002***	-0.002***	-0.002***	-0.000
	(-3.42)	(-3.55)	(-5.66)	(-4.10)	(-0.18)
rural	0.114	-0.657	-0.842***	0.755	1.315***
	(0.27)	(-1.43)	(-2.61)	(0.84)	(3.33)
edu_y	-0.015	-0.024	-0.043*	-0.055*	-0.082*
	(-0.46)	(-0.69)	(-1.84)	(-1.68)	(-1.93)
fedu_y	0.045**	-0.000	-0.021	-0.025	-0.027
	(2.13)	(-0.02)	(-1.24)	(-1.03)	(-1.09)
medu_y	-0.016	-0.021	0.016	-0.007	0.011
	(-0.57)	(-0.79)	(0.87)	(-0.26)	(0.28)
sedu_y	-0.026	-0.036	-0.066***	-0.055**	-0.052
	(-1.06)	(-1.34)	(-3.29)	(-2.50)	(-1.57)
incomepc	-0.094	-0.163	-0.126**	-0.039	-0.499***
	(-0.85)	(-1.42)	(-2.43)	(-0.26)	(-2.58)
married	0.556	1.971	4.285***	4.685	2.385
	(0.25)	(0.94)	(3.70)	(0.39)	(0.57)
slfhvg	0.604**	0.397	-0.256	-0.135	-1.249***
	(2.28)	(1.43)	(-0.97)	(-0.47)	(-3.40)
slfhg	0.570**	0.624**	-0.025	0.497*	0.044
	(2.49)	(2.43)	(-0.10)	(1.90)	(0.16)
slfhp	-0.001	0.156	0.511	1.091***	0.224
	(-0.00)	(0.34)	(1.39)	(2.77)	(0.58)
depression	0.013	-0.325	-0.807***	-0.958**	-1.213***
	(0.05)	(-0.97)	(-3.24)	(-2.37)	(-3.18)
cognitive	0.009	0.003	-0.004	-0.006	-0.007

	(1.09)	(0.33)	(-0.50)	(-0.80)	(-0.86)
workagri	-1.854**	-0.380	-0.914**	-1.265**	0.083
	(-2.04)	(-0.33)	(-2.10)	(-2.21)	(0.15)
employed	-0.405*	-0.336	-0.315**	-0.559**	-0.013
	(-1.81)	(-1.50)	(-2.03)	(-2.22)	(-0.04)
slfemploy	-0.002	0.140	-0.073	-0.061	-0.489
	(-0.01)	(0.50)	(-0.34)	(-0.17)	(-1.06)
wealthpc	-0.028	-0.070	-0.008	0.074	0.161*
	(-0.32)	(-1.19)	(-0.16)	(0.95)	(1.82)
foodpc	0.135	0.272*	0.211*	0.036	0.509***
	(0.87)	(1.65)	(1.94)	(0.23)	(2.87)
good_mob	-0.098	0.123	1.544***	2.149***	0.515
	(-0.22)	(0.25)	(4.83)	(4.99)	(0.23)
smoker	0.103	0.475	1.485**	2.132***	1.684***
	(0.19)	(0.53)	(2.32)	(5.04)	(3.26)
drinker	-0.057	0.088	0.811***	0.981***	2.257
	(-0.20)	(0.07)	(2.86)	(2.65)	(1.36)
phyact	0.036***	0.012	-0.005	-0.017	-0.033
	(2.97)	(1.07)	(-0.36)	(-0.86)	(-1.31)
_cons	13.661	14.774***	12.660***	14.210	23.577***
	(.)	(5.04)	(7.00)	(1.16)	(4.24)
N	1907	1907	1907	1907	1907
r2					

Note: The mean values of BMI among females are 19.40, 21.31, 22.91, 24.49, and 27.30 in the 10th, 30th, 50th, 70th, and 90th, respectively. *t* statistics in parentheses. *N* denotes the number of observations. Fair self-reported health status and unemployed are omitted categories. Precise definitions of all variables can be found in Appendix 4.1. Province dummies are included, but their coefficients are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Table G. Ordered Probit models for weight categories (alternative measure of time preferences with *TPc*)

	(1) underweight	(2) normal	(3) overweight	(4) obese
TPc	-0.004***	-0.015***	0.011***	0.008***
	(-2.69)	(-2.74)	(2.75)	(2.71)
male	-0.031***	-0.118***	0.090***	0.059***
	(-5.41)	(-6.32)	(6.20)	(6.03)
age	-0.004***	-0.014***	0.011***	0.007***
	(-3.41)	(-3.68)	(3.62)	(3.63)
age2	0.000***	0.000***	-0.000***	-0.000***
	(3.49)	(3.76)	(-3.71)	(-3.70)
rural	-0.002	-0.008	0.006	0.004
	(-0.25)	(-0.25)	(0.25)	(0.25)
edu_y	0.000	0.002	-0.001	-0.001
	(0.63)	(0.63)	(-0.63)	(-0.63)
fedu_y	0.000	0.000	-0.000	-0.000
	(0.21)	(0.21)	(-0.21)	(-0.21)
medu_y	0.000	0.002	-0.001	-0.001
	(0.78)	(0.79)	(-0.78)	(-0.79)
sedu_y	-0.000	-0.001	0.001	0.001
	(-0.59)	(-0.60)	(0.60)	(0.59)
incomepc	-0.003	-0.012	0.009	0.006
	(-1.07)	(-1.08)	(1.07)	(1.08)
married	0.017	0.066	-0.050	-0.033
	(0.90)	(0.90)	(-0.90)	(-0.90)
slfhvg	0.011*	0.041*	-0.031*	-0.021*
	(1.76)	(1.78)	(-1.79)	(-1.77)
slfhg	0.001	0.004	-0.003	-0.002
	(0.18)	(0.18)	(-0.18)	(-0.18)
slfhp	-0.003	-0.010	0.008	0.005
	(-0.35)	(-0.35)	(0.35)	(0.35)
depression	0.013**	0.049**	-0.037**	-0.024**

	(2.17)	(2.16)	(-2.19)	(-2.13)
cognitive	-0.000	-0.000	0.000	0.000
	(-0.12)	(-0.12)	(0.12)	(0.12)
workagri	0.019	0.072*	-0.055*	-0.036
	(1.63)	(1.65)	(-1.65)	(-1.64)
employed	0.013**	0.052**	-0.039**	-0.026**
	(2.36)	(2.43)	(-2.43)	(-2.41)
slfemploy	0.001	0.005	-0.004	-0.002
	(0.18)	(0.18)	(-0.18)	(-0.18)
wealthpc	-0.002	-0.009	0.007	0.004
	(-1.56)	(-1.57)	(1.57)	(1.57)
foodpc	0.000	0.001	-0.000	-0.000
	(0.05)	(0.05)	(-0.05)	(-0.05)
good_mob	-0.014	-0.055	0.042	0.028
	(-1.33)	(-1.32)	(1.32)	(1.33)
smoker	0.002	0.008	-0.006	-0.004
	(0.39)	(0.39)	(-0.39)	(-0.39)
drinker	-0.003	-0.012	0.009	0.006
	(-0.56)	(-0.56)	(0.56)	(0.56)
phyact	-0.000	-0.001	0.001	0.000
	(-0.76)	(-0.77)	(0.77)	(0.77)
N	3890	3890	3890	3890
r2_p				

Note: Columns 1 to 4 refer to regressions for the underweight, normal weight, overweight, and obese. Marginal effects are reported, with *z* statistics in parentheses. *N* denotes the number of observations. Fair self-reported health status and unemployed are omitted categories. Precise definitions of all variables can be found in Appendix 4.1. Province dummies are included, but their marginal effects are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Table H. Ordered Probit models for weight categories among males (alternative measure of time preferences with *TPc*)

	(1) underweight	(2) normal	(3) overweight	(4) obese
TPc	-0.003**	-0.022**	0.014**	0.011**
	(-2.31)	(-2.57)	(2.54)	(2.53)
age	-0.001	-0.006	0.004	0.003
	(-0.99)	(-0.99)	(0.99)	(1.00)
age2	0.000	0.000	-0.000	-0.000
	(1.18)	(1.19)	(-1.19)	(-1.18)
rural	-0.002	-0.011	0.007	0.005
	(-0.21)	(-0.21)	(0.21)	(0.21)
edu_y	-0.001	-0.005	0.003	0.002
	(-1.10)	(-1.14)	(1.14)	(1.13)
fedu_y	0.000	0.002	-0.002	-0.001
	(0.88)	(0.89)	(-0.89)	(-0.89)
medu_y	-0.000	-0.000	0.000	0.000
	(-0.08)	(-0.08)	(0.08)	(0.08)
sedu_y	-0.001	-0.004	0.002	0.002
	(-1.11)	(-1.14)	(1.14)	(1.14)
incomepc	-0.007***	-0.045***	0.029***	0.023***
	(-2.76)	(-2.92)	(2.90)	(2.90)
married	0.009	0.060	-0.038	-0.030
	(0.49)	(0.49)	(-0.49)	(-0.50)
slfhvg	0.007	0.049	-0.031	-0.025
	(1.17)	(1.24)	(-1.23)	(-1.22)
slfhg	0.006	0.040	-0.026	-0.021
	(1.11)	(1.14)	(-1.14)	(-1.13)
slfhp	0.004	0.027	-0.017	-0.014
	(0.57)	(0.58)	(-0.58)	(-0.58)
depression	0.009	0.056	-0.036	-0.028
	(1.59)	(1.61)	(-1.63)	(-1.58)
cognitive	-0.000*	-0.002*	0.001*	0.001*

	(-1.65)	(-1.70)	(1.70)	(1.69)
workagri	0.000	0.001	-0.001	-0.001
	(0.01)	(0.01)	(-0.01)	(-0.01)
employed	-0.002	-0.016	0.010	0.008
	(-0.42)	(-0.43)	(0.42)	(0.42)
slfemploy	-0.012	-0.079*	0.051*	0.040*
	(-1.64)	(-1.78)	(1.76)	(1.77)
wealthpc	-0.002*	-0.015*	0.009*	0.007*
	(-1.74)	(-1.77)	(1.78)	(1.76)
foodpc	0.002	0.015	-0.009	-0.007
	(0.77)	(0.78)	(-0.78)	(-0.78)
good_mob	-0.006	-0.043	0.028	0.022
	(-0.71)	(-0.71)	(0.71)	(0.71)
smoker	0.005	0.031	-0.020	-0.016
	(1.30)	(1.35)	(-1.35)	(-1.34)
drinker	-0.002	-0.012	0.008	0.006
	(-0.48)	(-0.48)	(0.48)	(0.48)
phyact	-0.000	-0.001	0.001	0.000
	(-0.60)	(-0.61)	(0.61)	(0.61)
N	1983	1983	1983	1983
r2_p				

Note: Columns 1 to 4 refer to regressions for the underweight, normal weight, overweight, and obese. Marginal effects are reported, with z statistics in parentheses. N denotes the number of observations. Fair self-reported health status and unemployed are omitted categories. Precise definitions of all variables can be found in Appendix 4.1. Province dummies are included, but their marginal effects are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Table I. Ordered Probit models for weight categories among females (alternative measure of time preferences with *TPc*)

	(1) underweight	(2) normal	(3) overweight	(4) obese
TPc	-0.004*	-0.011*	0.010*	0.005*
	(-1.73)	(-1.73)	(1.75)	(1.70)
age	-0.007***	-0.016***	0.015***	0.008***
	(-3.02)	(-3.35)	(3.25)	(3.27)
age2	0.000***	0.000***	-0.000***	-0.000***
	(2.75)	(2.98)	(-2.91)	(-2.93)
rural	-0.003	-0.008	0.008	0.004
	(-0.21)	(-0.21)	(0.21)	(0.21)
edu_y	0.002	0.004	-0.004	-0.002
	(1.49)	(1.48)	(-1.49)	(-1.49)
fedu_y	-0.000	-0.000	0.000	0.000
	(-0.20)	(-0.20)	(0.20)	(0.20)
medu_y	0.001	0.002	-0.002	-0.001
	(0.73)	(0.73)	(-0.73)	(-0.73)
sedu_y	0.001	0.004	-0.003	-0.002
	(1.51)	(1.52)	(-1.52)	(-1.52)
incomepc	0.007	0.016	-0.015	-0.008
	(1.28)	(1.29)	(-1.29)	(-1.28)
married	-0.060***	-0.145***	0.132***	0.073***
	(-4.57)	(-3.97)	(4.17)	(4.11)
slfhvg	0.010	0.024	-0.021	-0.012
	(0.92)	(0.91)	(-0.91)	(-0.91)
slfhg	-0.013	-0.031	0.029	0.016
	(-1.31)	(-1.33)	(1.33)	(1.32)
slfhp	-0.010	-0.023	0.021	0.012
	(-0.71)	(-0.70)	(0.70)	(0.70)
depression	0.019*	0.045*	-0.041*	-0.023
	(1.67)	(1.66)	(-1.68)	(-1.64)
cognitive	0.000	0.000	-0.000	-0.000

	(0.07)	(0.07)	(-0.07)	(-0.07)
workagri	0.024	0.057	-0.052	-0.029
	(1.15)	(1.16)	(-1.16)	(-1.15)
employed	0.020**	0.048**	-0.043**	-0.024**
	(2.00)	(2.07)	(-2.06)	(-2.04)
slfemploy	0.010	0.024	-0.022	-0.012
	(0.82)	(0.82)	(-0.83)	(-0.82)
wealthpc	-0.001	-0.003	0.002	0.001
	(-0.40)	(-0.40)	(0.40)	(0.40)
foodpc	-0.004	-0.010	0.009	0.005
	(-0.67)	(-0.68)	(0.68)	(0.67)
good_mob	-0.027	-0.065	0.060	0.033
	(-1.44)	(-1.42)	(1.43)	(1.43)
smoker	-0.041*	-0.098*	0.089*	0.050*
	(-1.80)	(-1.82)	(1.83)	(1.80)
drinker	-0.025	-0.059	0.054	0.030
	(-0.99)	(-0.99)	(1.00)	(0.98)
phyact	-0.000	-0.001	0.001	0.000
	(-0.66)	(-0.67)	(0.67)	(0.67)
N	1907	1907	1907	1907
r2_p				

Note: Columns 1 to 4 refer to regressions for the underweight, normal weight, overweight, and obese. Marginal effects are reported, with *z* statistics in parentheses. *N* denotes the number of observations. Fair self-reported health status and unemployed are omitted categories. Precise definitions of all variables can be found in Appendix 4.1. Province dummies are included, but their marginal effects are not reported for brevity. * indicates statistical significance at the 10% level; **, at the 5% level; ***, at the 1% level.

Chapter Five: Conclusion

5.1 Summary of the findings

This thesis consists of three empirical studies on unhealthy behaviours in China. We firstly examine how cognitive abilities are associated with smoking, drinking, and physical inactivity among the middle-aged and elderly in China. Then, we investigate the association between the time to the first cigarette (TTFC) after waking up and smoking cessation among the middle-aged and elderly people in China. Moreover, using a national representative dataset in China, we investigate whether time preferences are a determinant of the probability of being overweight, being obese and BMI. This thesis contributes to the literature by exploring the determinants of unhealthy behaviours in China from the aspects of cognitive abilities, time to first cigarette after waking up, and time preferences.

In Chapter Two, we use the 2011, 2013, and 2015 waves of the China and Retirement Longitudinal Study, and examine the association between cognitive abilities and unhealthy behaviours, including smoking, drinking, and physical inactivity, among the middle-aged and elderly in China. To the best of our knowledge, we are the first to examine this association in the Chinese context. Controlling for demographic, health and financial variables, we find that individuals with higher cognitive ability scores are less likely to be current smokers, heavy alcohol drinkers, and to be physically inactive. However, we do not find significant associations between cognition and the number of cigarettes smoked per day, as well as the unit of alcohol drinks consumed per week. Among the married respondents, the probabilities of smoking, heavy drinking, and physical inactivity are significantly positively influenced by that of their spouses'. These findings indicate that individuals with higher cognitive abilities are prone to adopt healthier behaviours, and individuals' behaviours are concordant with their spouses'.

In Chapter Three, using the 2011, 2013, and 2015 waves of the CHARLS, we investigate the impact of the TTFC on smoking cessation among middle-aged and elderly people in China. It is worth mentioning that little research has been conducted on this topic in China. We contribute to the literature by examining the relationship between the TTFC and cessation. The results show that TTFC is positively associated with smoking cessation among middle-aged and elderly people in China. The positive association is higher among light smokers than heavy smokers.

In Chapter Four, we examine whether time preference affects the probability of being overweight or obese, and BMI, using the 2014 wave of the China Family Panel Studies. Currently, the relationship between time preferences and obesity in China has not been investigated with national representative datasets. The results show that time preferences have a positive relationship with the probability of being overweight, being obese, and BMI. Also, time preferences have a stronger association with the probability of being overweight than being obese.

5.2 Policy implications

Due to the high prevalence of unhealthy behaviours in China, this thesis provides a better understanding of the determinants smoking, drinking, physical inactivity, and obesity. We provide the following policy implications.

The findings in Chapter Two show that individuals with higher cognitive abilities are less likely to be current smokers and heavy drinkers, and to be physically inactive. Thus, we suggest that an enhancement of cognitive abilities, targeted at middle-aged and elderly people in China, may be effective in reducing the prevalence of unhealthy behaviours. Also, it is worth mentioning that those with lower cognitive abilities have a higher likelihood of adopting risky behaviours, so policy makers need to pay more attentions to these people, and make recommendations that are easy for them to understand and implement. Moreover, this study finds that the spousal effect of unhealthy behaviours exists among husbands and wives in China. It is beneficial for the implication of health behaviour interventions among all Chinese. When any couple member is affected by the health behaviours interventions, his/her spouse is more likely to have the same habit.

In Chapter Three, we find that delaying the time to the first cigarette after waking up contributes to smoking cessation for middle-aged and elderly people in China. Our results suggest that it is necessary for policy makers to advise people to smoke their first cigarette later after waking up in the morning, because this method is effective in reducing their dependence on nicotine. Smoking control bans, such as no cigarette within residential buildings and office buildings before noon, may be helpful for cigarette quitting. Also, promoting health knowledge in primary and secondary schools increases children's awareness of the benefits of smoking cessation, and, in turn, children's world of mouth may play a heavier role on middle-aged and elderly people in China.

In Chapter Four, we find that time preferences are positively associated with body weight in the Chinese context, which means that impatient people are prone to be overweight, be obese, or have higher BMI. These findings suggest that policies aimed at educating people to be more far-sighted are needed, as they may contribute to reducing the prevalence of overweight and obesity in China. Policies, such as those emphasising the importance of nutrition balance and improving the awareness of calories content in food, are much needed in China. Good nutritional knowledge helps people make long-term oriented choices, such as adopting healthier dietary habits, and ultimately decreases the prevalence of overweight and obesity. At the same time, policies aimed at improving physical activities may contribute to lowering the prevalence of overweight and obesity, because doing physical exercise is an investment on future health.

References

Abdullah, A.S.M., Ho, L.M., Kwan, Y.H., Cheung, W.L., McGhee, S.M. and Chan, W.H., 2006. Promoting smoking cessation among the elderly: what are the predictors of intention to quit and successful quitting?. Journal of Aging and Health, 18(4), pp.552-564.

Abdullah, A.S.M., Lam, T.H., Chan, S.K., Leung, G.M., Chi, I., Ho, W.W. and Chan, S.S., 2008. Effectiveness of a mobile smoking cessation service in reaching elderly smokers and predictors of quitting. BMC Geriatrics, 8(1), p.25.

Allen, S.C., 2008. What determines the ability to stop smoking in old age?. Age and Ageing, 37, pp.490-491.

Andrews, J.O., Heath, J. and Graham-Garcia, J., 2004. Management of tobacco dependence in older adults: using evidence-based strategies. Journal of Gerontological Nursing, 30(12), pp.13-24.

Anstey, K.J., Low, L.F., Christensen, H. and Sachdev, P., 2009. Level of cognitive performance as a correlate and predictor of health behaviors that protect against cognitive decline in late life: The path through life study. Intelligence, 37(6), pp.600-606.

Appel, D.W. and Aldrich, T.K., 2003. Smoking cessation in the elderly. Clinics in Geriatric Medicine, 19(1), pp.77-100.

Artaud, F., Dugravot, A., Sabia, S., Singh-Manoux, A., Tzourio, C. and Elbaz, A., 2013. Unhealthy behaviours and disability in older adults: Three-City Dijon cohort study. British Medical Journal, 347, p.f4240.

Babb S, Malarcher A, Schauer G, Asman K, Jamal A. 2017. Quitting Smoking Among Adults — United States, 2000–2015. MMWR Morbidity and Mortality Weekly Report, 65:1457–1464. DOI: http://dx.doi.org/10.15585/mmwr.mm6552a1

Baker, T.B., Piper, M.E., McCarthy, D.E., Bolt, D.M., Smith, S.S., Kim, S.Y., Colby, S., Conti, D., Giovino, G.A., Hatsukami, D. and Hyland, A., 2007. Time to first cigarette in the morning as an index of ability to quit smoking: implications for nicotine dependence. Nicotine & Tobacco Research, 9(Suppl_4), pp.S555-S570.

Baltagi, B. (1995). Econometric Analysis of Panel Data, Chichester: John Wiley and Sons.

Banks, J., o'Dea, C. and Oldfield, Z., 2010. Cognitive function, numeracy and retirement saving trajectories. The Economic Journal, 120(548), pp.F381-F410.

Batty, G.D., Deary, I.J. and Macintyre, S., 2006. Childhood IQ and life course socioeconomic position in relation to alcohol induced hangovers in adulthood: the Aberdeen children of the 1950s study. Journal of Epidemiology & Community Health, 60(10), pp.872-874.

Batty, G.D., Deary, I.J. and Macintyre, S., 2007a. Childhood IQ in relation to risk factors for premature mortality in middle-aged persons: the Aberdeen Children of the 1950s study. Journal of Epidemiology & Community Health, 61(3), pp.241-247.

Batty, G.D., Deary, I.J., Schoon, I. and Gale, C.R., 2007b. Childhood mental ability in relation to food intake and physical activity in adulthood: the 1970 British Cohort Study. Pediatrics, 119(1), pp.e38-e45.

Baum II, C.L. and Ruhm, C.J., 2009. Age, socioeconomic status and obesity growth. Journal of Health Economics, 28(3), pp.635-648.

Becker, G.S. and Mulligan, C.B., 1997. The endogenous determination of time preference. The Quarterly Journal of Economics, 112(3), pp.729-758.

Bellavia, A., Bottai, M., Wolk, A. and Orsini, N., 2014. Alcohol consumption and mortality: a dose-response analysis in terms of time. Annals of Epidemiology, 24(4), pp.291-296.

Benowitz, N.L., Jacob, P.I.I.I., Jones, R.T. and Rosenberg, J., 1982. Interindividual variability in the metabolism and cardiovascular effects of nicotine in man. Journal of Pharmacology and Experimental Therapeutics, 221(2), pp.368-372.

Borghans, L. and Golsteyn, B.H., 2006. Time discounting and the body mass index: Evidence from the Netherlands. Economics & Human Biology, 4(1), pp.39-61.

Brega, A.G., Grigsby, J., Kooken, R., Hamman, R.F. and Baxter, J., 2008. The impact of executive cognitive functioning on rates of smoking cessation in the San Luis Valley Health and Aging Study. Age and Ageing, 37(5), pp.521-525.

Brown, H. and Biosca, O., 2016. Exploring the relationship between time preference, body fatness, and educational attainment. Social Science & Medicine, 158, pp.75-85.

Burkhauser, R.V. and Cawley, J., 2008. Beyond BMI: the value of more accurate measures of fatness and obesity in social science research. Journal of Health Economics, 27(2), pp.519-529.

Cappellari, L. and Jenkins, S.P., 2003. Multivariate probit regression using simulated maximum likelihood. The Stata Journal, 3(3), pp.278-294.

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

Chaaya, M., Mehio-Sibai, A. and El-Chemaly, S., 2006. Smoking patterns and predictors of smoking cessation in elderly populations in Lebanon. The International Journal of Tuberculosis and Lung Disease, 10(8), pp.917-923.

Chabris, C.F., Laibson, D., Morris, C.L., Schuldt, J.P. and Taubinsky, D., 2008. Individual laboratory-measured discount rates predict field behavior. Journal of Risk and Uncertainty, 37(2-3), p.237.

Chan, C.W. and Leung, S.F., 2015. Lifestyle health behaviors of Hong Kong Chinese: results of a cluster analysis. Asia Pacific Journal of Public Health, 27(3), pp.293-302.

Chang, L.C., Loh, E.W., Tsai, Y.W., Chiou, S.T. and Chen, L.K., 2014. Clinical benefits of smoking cessation in reducing all-cause and disease-specific mortality among older people in Taiwan: A 10-year nationwide retrospective cohort study. European Geriatric Medicine, 5(3), pp.149-154.

Chatterji, S., Kowal, P., Mathers, C., Naidoo, N., Verdes, E., Smith, J.P. and Suzman, R., 2008. The health of aging populations in China and India. Health Affairs, 27(4), pp.1052-1063.

Choi, N.G. and DiNitto, D.M., 2014. Role of new diagnosis, social isolation, and depression in older adults' smoking cessation. The Gerontologist, 55(5), pp.793-801.

Chou, K.L., 2008. The prevalence and clustering of four major lifestyle risk factors in Hong Kong Chinese older adults. Journal of Aging and Health, 20(7), pp.788-803.

Collins, N., Sachs-Ericsson, N., Preacher, K.J., Sheffield, K.M. and Markides, K., 2009. Smoking increases risk for cognitive decline among community-dwelling older Mexican Americans. The American Journal of Geriatric Psychiatry, 17(11), pp.934-942.

Compton, W.M., Gfroerer, J., Conway, K.P. and Finger, M.S., 2014. Unemployment and substance outcomes in the United States 2002–2010. Drug and Alcohol Dependence, 142, pp.350-353.

Cook, P.J. and Moore, M.J., 2002. The economics of alcohol abuse and alcohol-control policies. Health Affairs, 21(2), pp.120-133.

Courtemanche, C., Heutel, G. and McAlvanah, P., 2015. Impatience, incentives and obesity. The Economic Journal, 125(582), pp.1-31.

Cui, X., Rockett, I.R., Yang, T. and Cao, R., 2012. Work stress, life stress, and smoking among rural–urban migrant workers in China. BMC Public Health, 12(1), p.979.

Cutler, D.M., Glaeser, E.L. and Rosen, A.B., 2009. Is the US population behaving healthier?. In Social security policy in a changing environment. University of Chicago Press, pp. 423-442.

Cutler, D.M. and Lleras-Muney, A., 2010. Understanding differences in health behaviors by education. Journal of Health Economics, 29(1), pp.1-28.

Deary, I.J., Whalley, L.J., Lemmon, H., Crawford, J.R. and Starr, J.M., 2000. The stability of individual differences in mental ability from childhood to old age: follow-up of the 1932 Scottish Mental Survey. Intelligence, 28(1), pp.49-55.

Deary, I.J., 2014. The stability of intelligence from childhood to old age. Current Directions in Psychological Science, 23(4), pp.239-245.

Dioikitopoulos, E.V., Minos, D. and Vandoros, S., 2019. The (Agri-) Cultural origins of obesity. Social Science & Medicine, p.112523.

Dodd, M.C., 2014. Intertemporal discounting as a risk factor for high BMI: Evidence from Australia, 2008. Economics & Human Biology, 12, pp.83-97.

Doll, R., Peto, R., Boreham, J. and Sutherland, I., 2004. Mortality in relation to smoking: 50 years' observations on male British doctors. British Medical Journal., 328(7455), p.1519.

Du, S., Mroz, T.A., Zhai, F. and Popkin, B.M., 2004. Rapid income growth adversely affects diet quality in China—particularly for the poor!. Social Science & Medicine, 59(7), pp.1505-1515.

Du, H., Bennett, D., Li, L., Whitlock, G., Guo, Y., Collins, R., Chen, J., Bian, Z., Feng, S. and Chen, X. 2013. Physical activity and sedentary leisure time and their associations with Bmi, waist circumference, and percentage body fat in 0.5 million adults: the China Kadoorie Biobank study. The American Journal of Clinical Nutrition, 97, 487-496.

Ezzati, M. and Lopez, A.D., 2003. Estimates of global mortality attributable to smoking in 2000. The Lancet, 362(9387), pp.847-852.

Fagerstrom, K., 2003. Time to first cigarette; the best single indicator of tobacco dependence?. Monaldi Archives for Chest Disease, 59(1), pp.91-94.

Feng, X.L., Pang, M. and Beard, J., 2013. Health system strengthening and hypertension awareness, treatment and control: data from the China Health and Retirement Longitudinal Study. Bulletin of the World Health Organisation, 92, pp.29-41.

Fogelholm, M.I.K.A.E.L., Malmberg, J.A.R.M.O., Suni, J., Santtila, M.A.T.T.I., Kyröläinen, H., Mäntysaari, M. and Oja, P., 2006. International physical activity questionnaire: Validity against fitness. Medicine & Science in Sports & Exercise, 38(4), pp.753-760.

Ford, N.D., Patel, S.A. and Narayan, K.V., 2017. Obesity in low-and middle-income countries: burden, drivers, and emerging challenges. Annual Review of Public Health, 38, pp.145-164.

Fu, M., Martínez-Sánchez, J.M., Agudo, A., Pascual, J.A., Borràs, J.M., Samet, J.M., Fernández, E. and DCOT Study Investigators, 2011. Association between time to first cigarette after waking up and salivary cotinine concentration. Nicotine & Tobacco Research, 13(3), pp.168-172.

Fu, L., 2016. 'The effects of cognitive abilities on middle aged and elderly people's health behaviours: Is China different from the UK?', *07 24404: Advanced Research methods*. University of Birmingham. Unpublished research proposal.

Fuchs, V.R., 1982. Time preference and health: An exploratory study. In Economic Aspects of Health. University of Chicago Press, pp. 93-120.

Gao, J., Nehl, E.J., Fu, H., Jia, Y., Liu, X. and Zheng, P., 2013. Workplace social capital and smoking among Chinese male employees: A multi-level, cross-sectional study. Preventive Medicine, 57(6), pp.831-836.

Gao, P. and Shen, J., 2017. An empirical analysis on the determinants of overweight and obesity in China. Applied Economics, 49(20), pp.1923-1936.

García, J., Lera-López, F. and Suárez, M.J., 2011. Estimation of a structural model of the determinants of the time spent on physical activity and sport: Evidence for Spain. Journal of Sports Economics, 12(5), pp.515-537.

Garrett, N.A., Brasure, M., Schmitz, K.H., Schultz, M.M. and Huber, M.R., 2004. Physical inactivity: direct cost to a health plan. American Journal of Preventive Medicine, 27(4), pp.304-309.

Greene, W.H., 2011. Econometric Analysis, 7th ed., Prentice Hall

Guthold, R., Ono, T., Strong, K.L., Chatterji, S. and Morabia, A., 2008. Worldwide variability in physical inactivity: a 51-country survey. American Journal of Preventive Medicine, 34(6), pp.486-494.

Hallal, P.C., Andersen, L.B., Bull, F.C., Guthold, R., Haskell, W., Ekelund, U. and Lancet Physical Activity Series Working Group, 2012. Global physical activity levels: surveillance progress, pitfalls, and prospects. The Lancet, 380(9838), pp.247-257.

Halme, J.T., Seppä, K., Alho, H., Poikolainen, K., Pirkola, S. and Aalto, M., 2010. Alcohol consumption and all-cause mortality among elderly in Finland. Drug and Alcohol Dependence, 106(2-3), pp.212-218.

Harden, K.P., Turkheimer, E. and Loehlin, J.C., 2007. Genotype by environment interaction in adolescents' cognitive aptitude. Behavior Genetics, 37(2), pp.273-283.

Haseli-Mashhadi, N., Pan, A., Ye, X., Wang, J., Qi, Q., Liu, Y., Li, H., Yu, Z., Lin, X. and Franco, O.H., 2009. Self-Rated Health in middle-aged and elderly Chinese: distribution, determinants and associations with cardio-metabolic risk factors. BMC Public Health, 9(1), p.368.

Holahan, C.J., North, R.J., Holahan, C.K., Hayes, R.B., Powers, D.A. and Ockene, J.K., 2012. Social influences on smoking in middle-aged and older women. Psychology of Addictive Behaviors, 26(3), p.519.

Huang, C.Q., Dong, B.R., Zhang, Y.L., Wu, H.M. and Liu, Q.X., 2009. Association of cognitive impairment with smoking, alcohol consumption, tea consumption, and exercise among Chinese nonagenarians/centenarians. Cognitive and Behavioral Neurology, 22(3), pp.190-196.

Huffman, D., Maurer, R. and Mitchell, O.S., 2017. Time discounting and economic decision-making in the older population. The Journal of the Economics of Ageing. Available at: http://dx.doi.org/10.1016/j.jeoa.2017.05.001

Hunter, J.E., 1986. Cognitive ability, cognitive aptitudes, job knowledge, and job performance. Journal of Vocational Behavior, 29(3), pp.340-362.

Hyland, A., Li, Q., Bauer, J.E., Giovino, G.A., Steger, C. and Cummings, K.M., 2004. Predictors of cessation in a cohort of current and former smokers followed over 13 years. Nicotine & Tobacco Research, 6(Suppl_3), pp.S363-S369.

Ikeda, S., Kang, M.I. and Ohtake, F., 2010. Hyperbolic discounting, the sign effect, and the body mass index. Journal of Health Economics, 29(2), pp.268-284.

Iparraguirre, J., 2015. Socioeconomic determinants of risk of harmful alcohol drinking among people aged 50 or over in England. British Medical Journal Open, 5(7), p.e007684.
Jha, P., Ramasundarahettige, C., Landsman, V., Rostron, B., Thun, M., Anderson, R.N., McAfee, T. and Peto, R., 2013. 21st-century hazards of smoking and benefits of cessation in the United States. New England Journal of Medicine, 368(4), pp.341-350.
Jiang, Y., Elton-Marshall, T., Fong, G.T. and Li, Q., 2010. Quitting smoking in China: findings from the ITC China Survey. Tobacco Control, 19(Suppl 2), pp.i12-i17.
Jin, L., Huang, Y., Bi, Y., Zhao, L., Xu, M., Xu, Y., Chen, Y., Gu, L., Dai, M., Wu, Y. and Hou, J., 2011. Association between alcohol consumption and metabolic syndrome in 19,215 middle-aged and elderly Chinese. Diabetes Research and Clinical Practice, 92(3), pp.386-392.
Jones-Smith, J.C., Gordon-Larsen, P., Siddiqi, A. and Popkin, B.M., 2012. Emerging disparities in overweight by educational attainment in Chinese adults (1989–2006). International Journal of Obesity, 36(6), p.866.
Jurj, A.L., Wen, W., Li, H.L., Zheng, W., Yang, G., Xiang, Y.B., Gao, Y.T. and Shu, X.O., 2006. Spousal correlations for lifestyle factors and selected diseases in Chinese couples. Annals of Epidemiology, 16(4), pp.285-291.
Kang, Y.K., Guo, W.J., Xu, H., Chen, Y.H., Li, X.J., Tan, Z.P., Li, N., Gesang, Z.R., Wang, Y.M., Liu, C.B. and Luo, Y., 2015. The 6-item Kessler psychological distress scale to survey serious mental illness among Chinese undergraduates: Psychometric properties and prevalence estimate. Comprehensive Psychiatry, 63, pp.105-112.
Kang, M.I. and Ikeda, S., 2016. Time discounting, present biases, and health-related behaviors: Evidence from Japan. Economics & Human Biology, 21, pp.122-136.

Keenan, P.S., 2009. Smoking and weight change after new health diagnoses in older adults. Archives of Internal Medicine, 169(3), pp.237-242.

Kenney, B.A., Holahan, C.J., Holahan, C.K., Brennan, P.L., Schutte, K.K. and Moos, R.H., 2009. Depressive symptoms, drinking problems, and smoking cessation in older smokers. Addictive Behaviors, 34(6-7), pp.548-553.

Kerr, S., Watson, H., Tolson, D., Lough, M. and Brown, M., 2006. Smoking after the age of 65 years: a qualitative exploration of older current and former smokers' views on smoking, stopping smoking, and smoking cessation resources and services. Health & Social Care in the Community, 14(6), pp.572-582.

Kim, S.K., Park, J.H., Lee, J.J., Lee, S.B., Kim, T.H., Han, J.W., Youn, J.C., Jhoo, J.H., Lee, D.Y. and Kim, K.W., 2013. Smoking in elderly Koreans: prevalence and factors associated with smoking cessation. Archives of Gerontology and Geriatrics, 56(1), pp.214-219.

Kinsella, K. and He, W., 2009. US Census Bureau, international population reports. Washington, DC: US Census Bureau.

Kirchner, J.E., Zubritsky, C., Cody, M., Coakley, E., Chen, H., Ware, J.H., Oslin, D.W., Sanchez, H.A., Durai, U.N.B., Miles, K.M. and Llorente, M.D., 2007. Alcohol consumption among older adults in primary care. Journal of General Internal Medicine, 22(1), pp.92-97.

Komlos, J., Smith, P.K. and Bogin, B., 2004. Obesity and the rate of time preference: is there a connection?. Journal of Biosocial Science, 36(2), pp.209-219.

Kubička, L., Matějček, Z., Dytrych, Z. and Roth, Z., 2001. IQ and personality traits assessed in childhood as predictors of drinking and smoking behaviour in middle-aged adults: a 24-year follow-up study. Addiction, 96(11), pp.1615-1628.

Laaksonen, M., Prättälä, R. and Lahelma, E., 2003. Sociodemographic determinants of multiple unhealthy behaviours. Scandinavian Journal of Public Health, 31(1), pp.37-43.

Lam, T.H., Li, Z.B., Ho, S.Y., Chan, W.M., Ho, K.S., Tham, M.K., Cowling, B.J., Schooling, C.M. and Leung, G.M., 2007. Smoking, quitting and mortality in an elderly cohort of 56 000 Hong Kong Chinese. Tobacco Control, 16(3), pp.182-189.

Latif, E., 2014. The impact of recession on drinking and smoking behaviours in Canada. Economic Modelling, 42, pp.43-56.

Lawrence, E.M., Pampel, F.C. and Mollborn, S., 2014. Life course transitions and racial and ethnic differences in smoking prevalence. Advances in Life Course Research, 22, pp.27-40.

Lee, S., Tsai, Y.W. and Sung, H.Y., 2016. Smoking cessation and receipt of cessation advice from health professionals among older smokers in Taiwan. Preventive Medicine, 91, pp.89-95.

Lei, X., Hu, Y., McArdle, J.J., Smith, J.P. and Zhao, Y., 2012. Gender differences in cognition among older adults in China. Journal of Human Resources, 47(4), pp.951-971.

Lei, X., Smith, J.P., Sun, X. and Zhao, Y., 2014. Gender differences in cognition in China and reasons for change over time: evidence from CHARLS. The Journal of the Economics of Ageing, 4, pp.46-55.

Lei, X. and Liu, H., 2018. Gender difference in the impact of retirement on cognitive abilities: Evidence from urban China. Journal of Comparative Economics, 46(4), pp.1425-1446.

Levy, D.T., Romano, E. and Mumford, E., 2005. The relationship of smoking cessation to sociodemographic characteristics, smoking intensity, and tobacco control policies. Nicotine & Tobacco Research, 7(3), pp.387-396.

Li, Q., Hsia, J. and Yang, G., 2011a. Prevalence of smoking in China in 2010. New England Journal of Medicine, 364(25), pp.2469-2470.

Li, Y., Jiang, Y., Zhang, M., Yin, P., Wu, F. and Zhao, W., 2011b. Drinking behaviour among men and women in China: the 2007 China Chronic Disease and Risk Factor Surveillance. Addiction, 106(11), pp.1946-1956.

Li, X.Y., Jiang, Y., Hu, N., Li, Y.C., Zhang, M., Huang, Z.J. and Zhao, W.H., 2012. Prevalence and characteristic of overweight and obesity among adults in China, 2010. Zhonghua Yu Fang Yi Xue Za Zhi [Chinese Journal of Preventive Medicine], 46(8), pp.683-686.

Li, K.K., Cardinal, B.J. and Acock, A.C., 2013. Concordance of physical activity trajectories among middle-aged and older married couples: Impact of diseases and functional difficulties. Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 68(5), pp.794-806.

Li, J., Wu, B., Tevik, K., Krokstad, S. and Helvik, A.S., 2019. Factors associated with elevated consumption of alcohol in older adults—comparison between China and Norway: the CLHLS and the HUNT Study. British Medical Journal Open, 9(8), p.e028646.

Lim, S.H., Tai, B.C., Yuan, J.M., Mimi, C.Y. and Koh, W.P., 2013. Smoking cessation and mortality among middle-aged and elderly Chinese in Singapore: the Singapore Chinese Health Study. Tobacco Control, 22(4), pp.235-240.

Lin, J.C., Guerrieri, J.G. and Moore, A.A., 2011. Drinking patterns and the development of functional limitations in older adults: longitudinal analyses of the health and retirement survey. Journal of Aging and Health, 23(5), pp.806-821.

Mai, Y., Peng, X. and Chen, W., 2013. How fast is the population ageing in China?. Asian Population Studies, 9(2), pp.216-239.

Mata, J. and Machado, J.A., 1996. Firm start-up size: A conditional quantile approach. European Economic Review, 40(6), pp.1305-1323.

McArdle, J.J., Fisher, G.G. and Kadlec, K.M., 2007. Latent variable analyses of age trends of cognition in the Health and Retirement Study, 1992-2004. Psychology and Aging, 22(3), p.525.

McLaren, L., 2007. Socioeconomic status and obesity. Epidemiologic Reviews, 29(1), pp.29-48.

Mercincavage, M., Branstetter, S.A., Muscat, J.E. and Horn, K.A., 2013. Time to first cigarette predicts cessation outcomes in adolescent smokers. Nicotine & Tobacco Research, 15(12), pp.1996-2004.

Messer, K., Trinidad, D.R., Al-Delaimy, W.K. and Pierce, J.P., 2008. Smoking cessation rates in the United States: a comparison of young adult and older smokers. American Journal of Public Health, 98(2), pp.317-322.

Mi, Y.J., Zhang, B., Wang, H.J., Yan, J., Han, W., Zhao, J., Liu, D.W. and Tian, Q.B., 2015. Prevalence and secular trends in obesity among Chinese adults, 1991− 2011. American Journal of Preventive Medicine, 49(5), pp.661-669.

Miles, R., 2006. Neighborhood disorder and smoking: findings of a European urban survey. Social Science & Medicine, 63(9), pp.2464-2475.

Molander, R.C., Yonker, J.A. and Krahn, D.D., 2010. Age‐related changes in drinking patterns from mid‐to older age: results from the Wisconsin longitudinal study. Alcoholism: Clinical and Experimental Research, 34(7), pp.1182-1192.

Mortensen, L.H., Sørensen, T.I. and Grønbæk, M., 2005. Intelligence in relation to later beverage preference and alcohol intake. Addiction, 100(10), pp.1445-1452.

Muntner, P., Gu, D., Wildman, R.P., Chen, J., Qan, W., Whelton, P.K. and He, J., 2005. Prevalence of physical activity among Chinese adults: results from the International Collaborative Study of Cardiovascular Disease in Asia. American Journal of Public Health, 95(9), pp.1631-1636.

Muscat, J.E., Stellman, S.D., Caraballo, R.S. and Richie, J.P., 2009. Time to first cigarette after waking predicts cotinine levels. Cancer Epidemiology and Prevention Biomarkers, 18(12), pp.3415-3420.

Ng, S.W., Norton, E.C., Guilkey, D.K. and Popkin, B.M., 2012. Estimation of a dynamic model of weight. Empirical Economics, 42(2), pp.413-443.

Ning, M., Zhang, Q. and Yang, M., 2016. Comparison of self-reported and biomedical data on hypertension and diabetes: findings from the China Health and Retirement Longitudinal Study (CHARLS). British Medical Journal Open, 6(1), p.e009836.

Pabayo, R., Belsky, J., Gauvin, L. and Curtis, S., 2011. Do area characteristics predict change in moderate-to-vigorous physical activity from ages 11 to 15 years?. Social Science & Medicine, 72(3), pp.430-438.

Palipudi, K.M., Gupta, P.C., Sinha, D.N., Andes, L.J., Asma, S., McAfee, T. and GATS Collaborative Group, 2012. Social determinants of health and tobacco use in thirteen low and middle income countries: evidence from Global Adult Tobacco Survey. PloS One, 7(3), p.e33466.

Pickering, K., Monahan, M., Guariglia, A. and Roberts, T.E., 2017. Time to tighten the belts? Exploring the relationship between savings and obesity. PloS One, 12(6), p.e0179921.

Pisinger, C., Aadahl, M., Toft, U. and Jørgensen, T., 2011. Smoking reduction did not promote future smoking cessation in a general population. Open Journal of Preventive Medicine, 1(03), p.73.

Plurphanswat, N. and Rodu, B., 2014. The association of smoking and demographic characteristics on body mass index and obesity among adults in the US, 1999–2012. BMC Obesity, 1(1), p.18.

Pouliou, T. and Elliott, S.J., 2010. Individual and socio-environmental determinants of overweight and obesity in Urban Canada. Health & Place, 16(2), pp.389-398.

Qin, X. and Pan, J., 2016. The medical cost attributable to obesity and overweight in China: Estimation based on longitudinal surveys. Health Economics, 25(10), pp.1291-1311.

Reed, M.B., Anderson, C.M., Vaughn, J.W. and Burns, D.M., 2008. The effect of cigarette price increases on smoking cessation in California. Prevention Science, 9(1), pp.47-54.

Reynolds, K., Gu, D., Whelton, P.K., Wu, X., Duan, X., Mo, J., He, J. and InterASIA Collaborative Group, 2007. Prevalence and risk factors of overweight and obesity in China. Obesity, 15(1), pp.10-18.

Rieger, M., 2015. Risk aversion, time preference and health production: Theory and empirical evidence from Cambodia. Economics & Human Biology, 17, pp.1-15.

Roodman, D., 2011. Fitting fully observed recursive mixed-process models with cmp. The Stata Journal, 11(2), pp.159-206.

Ruhm, C.J., 2005. Healthy living in hard times. Journal of Health Economics, 24(2), pp.341-363.

Sabia, S., Elbaz, A., Dugravot, A., Head, J., Shipley, M., Hagger-Johnson, G., Kivimaki, M. and Singh-Manoux, A., 2012. Impact of smoking on cognitive decline in early old age: the Whitehall II cohort study. Archives of General Psychiatry, 69(6), pp.627-635.

Samuelson, P.A., 1937. A note on measurement of utility. The Review of Economic Studies, 4(2), pp.155-161.

Schneider, U. and Schneider, B.S., 2006. The effects of education and working hours on health: A multivariate probit approach. Available at SSRN 897727.

Schuit, A.J., van Loon, A.J.M., Tijhuis, M. and Ocké, M.C., 2002. Clustering of lifestyle risk factors in a general adult population. Preventive Medicine, 35(3), pp.219-224.

Shankar, A., McMunn, A., Banks, J. and Steptoe, A., 2011. Loneliness, social isolation, and behavioral and biological health indicators in older adults. Health Psychology, 30(4), p.377.

Shaw, B.A. and Agahi, N., 2012. A prospective cohort study of health behavior profiles after age 50 and mortality risk. BMC Public Health, 12(1), p.803.

Skowron, M.A., Stodolska, M. and Shinew, K.J., 2008. Determinants of leisure time physical activity participation among Latina women. Leisure Sciences, 30(5), pp.429-447.

Smith, P.K., Bogin, B. and Bishai, D., 2005. Are time preference and body mass index associated?: Evidence from the National Longitudinal Survey of Youth. Economics & Human Biology, 3(2), pp.259-270.

Sobal, J. and Stunkard, A.J., 1989. Socioeconomic status and obesity: a review of the literature. Psychological Bulletin, 105(2), p.260.

Stenbacka, M., Leifman, A. and ROMELSJÖ, A., 2010. Mortality and cause of death among 1705 illicit drug users: a 37 year follow up. Drug and Alcohol Review, 29(1), pp.21-27.

Stimpson, J.P., Masel, M.C., Rudkin, L. and Peek, M.K., 2006. Shared health behaviors among older Mexican American spouses. American Journal of Health Behavior, 30(5), pp.495-502.

Strauss, J., Lei, X., Park, A., Shen, Y., Smith, J.P., Yang, Z. and Zhao, Y., 2010. Health outcomes and socio-economic status among the elderly in China: Evidence from the CHARLS Pilot. Journal of Population Ageing, 3(3-4), pp.111-142.

Swinburn, B.A., Sacks, G., Hall, K.D., McPherson, K., Finegood, D.T., Moodie, M.L. and Gortmaker, S.L., 2011. The global obesity pandemic: shaped by global drivers and local environments. The Lancet, 378(9793), pp.804-814.

Tabuchi, T., Nakamura, M., Nakayama, T., Miyashiro, I., Mori, J.I. and Tsukuma, H., 2016. Tobacco price increase and smoking cessation in Japan, a developed country with affordable tobacco: a national population-based observational study. Journal of Epidemiology, 26(1), pp.14-21.

Taylor, M.D., Hart, C.L., Smith, G.D., Starr, J.M., Hole, D.J., Whalley, L.J., Wilson, V. and Deary, I.J., 2005. Childhood IQ and social factors on smoking behaviour, lung function and smoking-related outcomes in adulthood: Linking the Scottish Mental Survey 1932 and the Midspan studies. British Journal of Health Psychology, 10(3), pp.399-410.

Trzaskowski, M., Harlaar, N., Arden, R., Krapohl, E., Rimfeld, K., McMillan, A., Dale, P.S. and Plomin, R., 2014. Genetic influence on family socioeconomic status and children's intelligence. Intelligence, 42, pp.83-88.

Tsai, A.C.H., Lin, Y.A. and Tsai, H.J., 2012. Predictors of smoking cessation in 50–66-year-old male Taiwanese smokers: A 7-year national cohort study. Archives of Gerontology and Geriatrics, 55(2), pp.295-300.

Wang, Q., 2015. Smoking and body weight: evidence from China health and nutrition survey. BMC Public Health, 15(1), p.1238.

Wang, S., Ungvari, G.S., Forester, B.P., Chiu, H.F., Wu, Y., Kou, C., Fu, Y., Qi, Y., Liu, Y., Tao, Y. and Yu, Y., 2017. Gender differences in general mental health, smoking, drinking and chronic diseases in older adults in Jilin province, China. Psychiatry Research, 251, pp.58-62.

Wang, Q., Shen, J.J., Sotero, M., Li, C.A. and Hou, Z., 2018. Income, occupation and education: Are they related to smoking behaviors in China?. PloS One, 13(2).

Weiser, M., Zarka, S., Werbeloff, N., Kravitz, E. and Lubin, G., 2010. Cognitive test scores in male adolescent cigarette smokers compared to non-smokers: a population-based study. Addiction, 105(2), pp.358-363.

West, R., McEwen, A., Bolling, K. and Owen, L., 2001. Smoking cessation and smoking patterns in the general population: a 1-year follow-up. Addiction, 96(6), pp.891-902.

WHO, 2009. Global health risks: mortality and burden of disease attributable to selected major risks. Geneva: World Health Organisation.

WHO, 2010. Global recommendations on physical activity for health. Geneva: World Health Organisation.

WHO, 2014. Global status report on alcohol and health. Management of Substance Abuse Unit. World Health Organisation.

WHO, 2018. Overweight and obesity. Fact sheet. Available at: https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight (Accessed: 16 February 2018)

Wu, X., 2011. The household registration system and rural-urban educational inequality in contemporary China. Chinese Sociological Review, 44(2), pp.31-51.

Xiao, Y., Zhao, N., Wang, H., Zhang, J., He, Q., Su, D., Zhao, M., Wang, L., Zhang, X., Gong, W. and Hu, R., 2013. Association between socioeconomic status and obesity in a Chinese adult population. BMC Public Health, 13(1), p.355.

Xie, Y. and Hu, J., 2014. An introduction to the China family panel studies (CFPS). Chinese Sociological Review, 47(1), pp.3-29.

Xie, Y. and Lu, P., 2015. The sampling design of the China family panel studies (CFPS). Chinese Journal of Sociology, 1(4), pp.471-484.

Yang, T., Abdullah, A.S.M., Mustafa, J., Chen, B. and Feng, X., 2009. Factors associated with smoking cessation among Chinese adults in rural China. American Journal of Health Behavior, 33(2), pp.125-134.

Yang, Y., Zheng, A., Li, M., Duan, W., Mu, X. and Wang, X., 2016. Medical economic burden of the ageing population: a multistage sampling analysis of 3 532 517 cases. The Lancet, 388, p.S79.

Yu, Z.B., Han, S.P., Cao, X.G. and Guo, X.R., 2010. Intelligence in relation to obesity: a systematic review and meta-analysis. Obesity Reviews, 11(9), pp.656-670.

Yu, N.W., Chen, C.Y., Liu, C.Y., Chau, Y.L. and Chang, C.M., 2011. Association of body mass index and depressive symptoms in a Chinese community population: results from the Health Promotion Knowledge, Attitudes, and Performance Survey in Taiwan. Chang Gung Medical Journal, 34(6), pp.620-627.

Yuan, J.M., Ross, R.K., Gao, Y.T., Henderson, B.E. and Mimi, C.Y., 1997. Follow up study of moderate alcohol intake and mortality among middle aged men in Shanghai, China. British Medical Journal, 314(7073), p.18.

Zhang, L.E.I. and Rashad, I., 2008. Obesity and time preference: the health consequences of discounting the future. Journal of Biosocial Science, 40(1), pp.97-113.

Zhang, X., Dagevos, H., He, Y., Van der Lans, I. and Zhai, F., 2008. Consumption and corpulence in China: a consumer segmentation study based on the food perspective. Food Policy, 33(1), pp.37-47.

Zhang, L., Liu, K., Li, H., Li, D., Chen, Z., Zhang, L.L. and Guo, L.L., 2016. Relationship between body mass index and depressive symptoms: the “fat and jolly” hypothesis for the middle-aged and elderly in China. BMC Public Health, 16(1), p.1201.

Zhang, Y., 2018. Understanding saving, consumption, and healthcare systems in China (Doctoral dissertation, University of Birmingham).

Zhao, Y., Hu, Y., Smith, J. P., Strauss, J. and Yang, G. 2012. Cohort profile: The China health and retirement longitudinal study (CHARLS). International Journal of Epidemiology, 43, 61-68.

Zhou, B.F., 2002. Predictive values of body mass index and waist circumference for risk factors of certain related diseases in Chinese adults--study on optimal cut-off points of body mass index and waist circumference in Chinese adults. Biomedical and Environmental Sciences: BES, 15(1), pp.83-96.

Zwar, N.A., Mendelsohn, C.P. and Richmond, R.L., 2014. Supporting smoking cessation. British Medical Journal, 348, p.f7535.