Netspar

Network for Studies on Pensions, Aging and Retirement

NETSPAR ACADEMIC SERIES

Lifestyle ex ante moral hazard

Insights from the Dutch voluntary deductible

Annelotte Zwemstra

MSc 01/2019-016

Lifestyle *ex ante* moral hazard: insights from the Dutch voluntary deductible

Annelotte Zwemstra*, supervised by Prof. Dr. Rob Alessie.
Faculty of Economics and Business, Rijksuniversiteit Groningen

Abstract This research identifies the existence of *ex ante* moral hazard in Dutch health insurance. Four of the main lifestyle variables associated with the development of non-communicable disease are investigated empirically. It concerns obesity, physical inactivity, heavy smoking and heavy alcohol consumption. The analysis points out that indeed the propensity of being obese, being a heavy smoker and being a heavy drinker decreases with the uptake of a voluntary deductible. Respectively, the average likeliness of adopting these lifestyle variables decreases with 29.7; 12.16 and 13.25 percentage points when opting for a voluntary deductible, compared to the reference group that goes with the default option of a zero voluntary deductible. A deductible exceeding the obligatory amount indeed suffices as a financial incentive to decrease unhealthy lifestyle adoption. Identification of these effects happens by employing two models on a dataset composed from the Dutch LISS panel. The dataset covers the years 2009-2017, with the exception of 2014. The models employed are a univariate probit and a bivariate probit with the inclusion of instrumental variables. Both models control for individual sources of selection, namely socio-demographic status, risk type and risk preferences. The second model includes the possibility of additional endogeneity from unobserved variables. The two instrumental variables introduced in this second model are the uptake of complementary health insurance and a measure for extreme pessimism. Comparison of the models allows me to conclude that there indeed exists additional endogeneity when considering obesity, heavy smoking and heavy drinking. This research contributes to the scarce empirical evidence on the existence of *ex ante* moral hazard in health insurance markets and provides an advice for Dutch health insurers and policy makers, concerning an income bounded increase in the supposedly ideal level of consumer deductible.

DATE:	January 10th 2019
JEL:	I12, I13, I14, I15, I16, I17, I18
FIELD:	Health economics, micro/behavioural economics.
KEY WORDS:	*Ex ante* moral hazard; health insurance; lifestyle choices; deductibles; chronic disease crisis, non-communicable disease (NCD) global crisis.
COURSE CODE:	EBM877A20, Rijksuniversiteit Groningen.

* Corresponding author at: Faculty of Economics and Business, student number s2555697,
Nettelbosje 2, 9747 AE Groningen, NL.
E-mail address: a.l.s.zwemstra@student.rug.nl (A.L.S. Zwemstra)
Combined with an internship at EY Financial Services Advisory

I. Introduction

As the World Health Organization emphasizes, the spread of non-communicable diseases (NCD's) presents a global crisis, putting all groups from different socio-demographic backgrounds in danger of developing this type of disease (World Health Organisation (after: WHO), 2005). The main NCD's are heart diseases, stroke, diabetes, cancers and chronic respiratory diseases. Out of these cases, most are chronic in nature. Besides concern from the WHO, the World Economic Forum now acknowledges NCD's as one of the top global threats to economic development (Bloom et al., 2012). On the individual level, this implies adverse outcomes on utility in the form of decreased quality of life, foregone income and increased costs for medical care, possibly leading to debt (WHO, 2005). On the level of societies, consequences are negative in the sense of a productivity loss of the population leading to foregone tax income and limiting economic growth, and a continuously increasing financial burden from financing the health budget (Bloom et al., 2012). This impacts both policy makers and health insurers. Concluding, the increasing prevalence of NCD's is a problem of high concern for policy makers worldwide.

It is established that over the past years, changes in the social and economic environment have resulted in a quick spread of the risk factors for NCD's over all countries (Geneau et al., 2010). The well-known and common main causes of NCD's are as follows: tobacco use; unhealthy food consumption, in particular excessive amounts of salt and sugar; physical inactivity; and the damaging excessive consumption of alcohol (Beaglehole et al., 2011). Where the intake of unhealthy food in combination with lack of physical activity can cause obesity. These lifestyle variables are responsible for over two third of new NCD cases and alongside increase the likeliness of additional obstacles in individuals that are already diagnosed with an NCD (Beaglehole et al., 2011). To introduce policy that effectively limits the prevalence of chronic diseases it is necessary to understand individual dynamics in lifestyle decision-making, as lifestyle is a phenomenon that is mostly impossible to influence through enforcement.

Basic economic theory suggests that the use of financial incentives could successfully influence individual lifestyle behaviour. Often, individuals value far-in-the-future or highly uncertain outcomes less than would be rational, whilst overvaluing short-term and more certain outcomes (Mazur, 1987). Financial incentives can be used to increase the short-term value – or decrease the short-term *dis*value – individuals attach to making healthy decisions

(Loewenstein et al., 2007). Even though health insurers and policy makers incorporate different kinds of financial incentives when designing health insurance systems, too little is known about the impact of such incentives on individual lifestyle behaviour.

There exists a broad stream of literature concerning the impact of offering different types of health insurance contracts to individuals. Theory is split in the concepts of *ex ante* and *ex post* moral hazard. Economists that developed the first of these two concepts, reason that the motivation to prevent a hazardous event from occurring decreases once insured against the financial consequences of such an event (Arrow, 1959). Prevention is a broad concept, economic theory refers both to use of preventive healthcare and to adoption of a healthy lifestyle. Unfortunately, there exists little empirical evidence on the relation between lifestyle and different levels or structures of health insurance. Some of the research on *ex ante* moral hazard focuses mostly on the relation between preventive healthcare use and insurance coverage, without zooming in on lifestyle. The few empirical researches that do focus on lifestyle behaviour find little – and sometimes contradicting – evidence on the occurrence of *ex ante* moral hazard on health insurance markets.

In this thesis, I extend the scarce empirical evidence on the occurrence of *ex ante* moral hazard in health insurance by sketching a clear picture on the relationship between four of the main lifestyle causes of NCD's and different levels of deductible in the Netherlands. In my research four different lifestyle variables are considered; obesity, total lack of physical exercise (inactivity), heavy smoking behaviour and heavy alcohol consumption. Not only are these highly correlated with the development of NCD's, also these are the variables mostly investigated in previous research on *ex ante* moral hazard. This provides me with the opportunity to compare my findings with previous evidence. The methodology I apply in some aspects is similar to the methodology of Stanciole (2008) and of Alessie et al. (2018). Stanciole (2008) investigates the existence of *ex ante* moral hazard using the same lifestyle variables with two similar models, I include several additional and improved (instrumental) variables in the analysis and I make use of a dataset covering a longer period. The models I apply are partially similar to models applied in the study of Alessie et al. (2018). The main difference is that whilst the study of these authors subjects *ex post* moral hazard, my study captures *ex ante* moral hazard. Consequently, the key-, control- and instrumental variables used in my study are largely different. The analysis in this study builds on data from the Longitudinal Internet Studies for Social sciences, better known as the LISS panel. This survey database contains yearly observations on a large sample representative of Dutch society.

The existence of *ex ante* moral hazard has never been researched within the Netherlands. As previous evidence on this topic is scarce and sometimes contradicting, it is very interesting to gain insight on the Dutch health insurance market. Besides contributing to empirical evidence on *ex ante* moral hazard in health insurance, I formulate an advice for Dutch health insurers and policy makers. My advice to Dutch policy makers and health insurers concerns an income bounded increase in the supposedly ideal level of consumer deductible.

I proceed as follows. Within the next chapter, I elaborate on the theoretical background and previous research on *ex ante* moral hazard. Highlighting both the foundation of theory on this topic, and previous empirical evidence. Next, I provide background information by discussing the unique nature of the Dutch health insurance system. Understanding the dynamics of this system is important, as in this research I use data on Dutch residents with Dutch health insurance contracts. Next I explain the research framework used within this thesis, proposing the main research question and the related hypotheses. After this, I discuss the data that this empirical research builds on. Providing the reader with a thorough understanding of the construction of, and motivation for, the different variables that are used. After this, I present a clear overview of the econometric model I use. Then finally, the results of the analysis are discussed thoroughly. Then I close with a discussion and conclusion.

II. Literature Review

II.I. Theoretical framework

During the past century, insurance and in particular health insurance – the insurance against financial consequences of needing medical care – has been a topic of interest. Arrow (1963) states that there exist multiple aspects evolving in particular around uncertainty, in which the medical care market differs from any standard competitive market. Later on, theory on the unique nature of health insurance is extended by Pauly (1968); Rothschild and Stiglitz (1976); Hemenway (1990) and Finkelstein and McGarry (2006). Construing this theory allows me to explain the principle of moral hazard in health insurance thoroughly.

Foundation of health insurance theory

In his paper, Arrow (1963) demonstrates the dynamics of an ideal health insurance market. In this market a pooled insurance contract is offered, with premiums slightly

exceeding the actuarially fair level due to a combination of administrative costs and a small fee percentage for insurers. All individuals will demand this insurance contract and there exists an economically (Pareto) efficient competitive equilibrium, where a transfer of purchasing power from the healthy to the ill takes place, representing both an individual welfare gain in the form of utility maximization and a social welfare gain. This social welfare gain can be explained with the 'law of large numbers'. When assuming that medical risks are largely independent (excluding epidemical risk), combining these risks for a large pool of individuals leads to a significantly lower total risk bore by the insurer (Arrow, 1963).

Arrow (1963) has introduced two Welfare theorems evolving around risk, to which market dynamics must comply in order to be reviewed as appropriate for insuring. Firstly, individuals are required to be utility maximizing (such as in the Bernoulli theorem, 1738) and risk averse. Secondly, the occurrence of an adverse event must be random, ensuring that spreading of risk indeed reduces overall risk significantly. Further elaborating on this topic, Pauly (1968) adds a condition to these assumptions. For an insurance market to be sustainable, a fairly small or zero price elasticity in quantity demanded of the covered good/service is crucial. Even if the incidence of illness would be completely random, whether or not the presence of insurance will alter the randomness of medical expenses is dependent on the elasticity of demand for medical care (Pauly, 1968).

In reality, these conditions are not nearly satisfied. Individuals are bounded in their rationality, inconsistent choices are made over time and within populations there exists large variation in risk preferences, health risks and elasticity of demand (Arrow, 1963; Pauly, 1968). Besides this, consumers do not necessarily possess the information and knowledge necessary to decide what is best for them and physicians can lack the incentive to reasonably ration care based on a results to cost ratio (Arrow, 1963). Even though the incidence of disease is random to some extent, the probability of falling ill does depend on human behaviour and can therefore be influenced by insurance coverage. Arrow (1963) describes this using an example of increased carelessness when insured. Besides this, the amount of care used when falling sick does in fact depend on the price of consumption (Pauly, 1968). To conclude, the nature of healthcare demand is highly unpredictable and previously mentioned factors make it (nearly) impossible to construct a competitive equilibrium in a health insurance market, offering a pooled contract with a hundred-percentage coverage.

Asymmetric information

Rothschild and Stiglitz (1976) introduce a model evolving around the theory of asymmetric information in insurance, where *risk type* is the key variable. Individuals hold private information on the actual probability of an adverse event occurring to them specifically. In other words, it is unknown to the insurer what risk type an individual is. The authors use a simplified model where individuals are either high risk, or low risk types. 'High risk' means: high risk of incurring costs due to illness. They state that individuals can use their private information to their advantage. Meaning that when one pooled insurance contract is offered, especially the high risk types will be interested – driving the price of the insurance contract upward – while the low risk types are likely to start opting out – driving the price even further upward. The authors name this principle *adverse selection*. What can be concluded from this theory is that high-risk types are associated with a larger demand for insurance coverage and a higher willingness to pay for insurance. A logical proposition following from this model is that a high insurance level might decrease the incentive to minimize the probability of actually needing to use the insurance coverage.

Related to this principle, is a counterpart theory proposed by Hemenway (1990), which associates the purchase of insurance with lower risk types. This is called *propitious selection.* The occurrence of *propitious selection* is underpinned with the argument that knowledge and awareness of the risks one carries increases prevention, thus signing an insurance contract might increase awareness of any risks and can therefore work as a trigger to minimize the chance of actually needing to use the insurance coverage.

Finkelstein and McGarry (2006) propose a relatively recent extension of models involving asymmetric information in insurance contracts. This paper provides an explanation for the principle Hemenway (1990) introduced. In their paper an empirical framework including asymmetric information on multiple dimensions, is built. Following the theory of Pauly (1968), the authors state that besides the risk type, consumers also possess private information on their *risk preferences*. Finkelstein and McGarry (2006) state that it is important to include this in any analysis, as risk type and risk preference can work in an opposite direction when determining whether or not to buy a certain insurance coverage level and when determining how much of insurance to actually use. High risk types are more likely to buy a high insurance coverage contract (following Rothschild and Stiglitz, 1976), whilst people with strong risk aversion are also more likely to buy such a contract. The first group consists of individuals bearing more risk for the insurer, driving costs and premiums upward. The latter does not necessarily consist of such a group, as highly risk averse individuals are

assumed to be less likely to take risks that can lead to sickness and are more likely to buy insurance coverage to buy of *any* level of risk. *Advantageous selection* can occur when the effect of highly risk averse people buying the high coverage contract, dominates the effect of high risk types buying this same contract.

Moral hazard

Moral hazard is an important phenomenon in insurance theory. Within health insurance, this phenomenon can be split into provider moral hazard and consumer moral hazard. Arrow (1963) defines provider moral hazard as an excess demand for healthcare due to supplier induced demand. Rice (1983) further develops theory around this concept, whose roots can be found in information asymmetry between supplier (physician) and patient. Consumer moral hazard can be defined as unconscionable demand for health investments due to health insurance coverage (Pauly, 1968). Basic economic theory suggests the existence of consumer moral hazard due to information asymmetries between individuals and insurers, as these make it impossible to set actuarially fair priced insurance premiums. Consequently, consumers are able to increase risky behaviour and increase the use of insurance when an adverse event has happened. The second of these consequences is usually referred to as *ex post* moral hazard. *Ex post* moral hazard focuses on the actual demand for health care once a hazardous event happened. In other words, *ex post* moral hazard implies all forms of healthcare overconsumption given an individual's health state. Historically, attention has been mostly focused on *ex post* moral hazard. *Ex ante* moral hazard can be formally defined as the mechanism where insurance reduces incentives to prevent the loss from occurring (Shavell, 1979), or where market insurance has a negative effect on self-protection, increasing the probability of a hazardous event occurring (Arrow, 1959). The economic reason behind consumer moral hazard is rather rational. The costs of excess usage are spread over all purchasers of the insurance contract (Pauly, 1968). This presents a prisoner's dilemma, where none of the insured individuals face an incentive to actually do what they know is best: minimize their demand for healthcare.

Extending the theory

For simplicity, so far only pooled insurance contracts with full coverage have been discussed. Health insurance theory can be extended with a variety of types of health insurance contracts, with different levels of price rationing (including differing levels of coverage) and with different levels of government involvement. Motivation for applying different types of

price rationing and government involvement is the imperfect nature of the health insurance market as emphasized in previous sub-sections.

As Pauly (1968) concludes, for an efficient market outcome in health insurance, at least some price rationing at the point of service is required. In general there are three main forms of customer involvement that insurers use in their contracts: the *deductible*, a form of *co-payment* and a form of *co-insurance*. The first of these, a contract including a deductible, is characterized by 100% coverage of incurred costs above some fixed set amount (Arrow, 1963). A contract with co-insurance has coverage at some level beneath 100%, whereas a contract with co-payments requires the insured individual to pay a flat fee on the costs of each treatment or drug prescription.

In the pooled contract with 100% coverage discussed previously, it is assumed that all individuals are automatically accepted into the contract. In the real world this only happens if strict regulations exist, forbidding insurers from risk rating. Risk rating can be defined as setting premiums according to individual health risks, determined by evaluating individual observable characteristics such as health state, gender and age (Light, 1992). A possible consequence of risk rating is that different premiums are offered to different groups, or in the extreme case that some individuals will not be accepted into any insurance contract. There exist different types of health insurance systems. These systems vary from maintaining policies that obligate all individuals to buy (or receive) public health insurance, to systems where the government does not interfere in the health insurance market at all. Consequently, in some markets indeed a (public) pooled insurance contract is offered to all individuals where community rating is applied, whereas in other markets separating contracts are offered by private parties.

II.II. Empirical evidence

Ex post moral hazard in health insurance

There exists a large variety of evidence on the existence of *ex post* moral hazard. As this thesis subjects *ex ante* moral hazard rather than *ex post* moral hazard, I mention only a few empirical investigations confirming the existence of *ex post* moral hazard.

Firstly, it is worthwhile to mention the largest health insurance experiment ever executed worldwide: the US RAND experiment. From 1974 until 1982, individuals were randomly assigned to different types of health insurance plans, containing different shares of cost coverage for insured individuals. The behaviour of these insured individuals was

followed thoroughly during this period. Changes in healthcare usage were clearly observable, this way amongst others Newhouse (1993) and later Manning and Marquis (1996) were able to identify the existence of *ex post* moral hazard. These authors were able to conclude that health insurance did not necessarily increase physical wellbeing amongst insured individuals compared to uninsured individuals, whilst it did increase the demand for healthcare. When performing this analysis, Newhouse (1993) does not include the demand for health directly into the utility function, whereas Manning and Marquis (1996) do include this into the individual utility function. Both investigations confirm the existence of *ex post* moral hazard, confirming a trade off between welfare gains of risk spreading and welfare loss due to increased demand for healthcare as proposed by Arrow (1963).

Less well-known but worth mentioning recent Dutch research on *ex post* moral hazard is executed by Alessie et al (2018). After confirmation of the existence of asymmetric information between insurer and customers, the effect of the voluntary deductible on the probability of having any hospital stays and doctors visits is investigated empirically. The authors make use of a regular probit model and a bivariate probit estimation using an instrument for voluntary deductible choice. The authors use the choice for supplementary health insurance as instrument. Furthermore, the effect on the amount of visits and stays is estimated through a Hurdle-Cragg model (1971). A strong feature of this research is that the authors control for asymmetric information on multiple dimensions as proposed by Rothschild and Stiglitz (1976), both individual risk preferences and risk type is taken into account, and there is controlled for socio-economic status.

Ex ante moral hazard in health insurance

Even though theoretically *ex ante* moral hazard is a well-known concept (Ehrlich and Becker, 1972; Shavell, 1979), there is little knowledge on the net effect of *ex ante* moral hazard in health insurance. Some of the available literature discusses that the negative effect of *ex ante* moral hazard might be small or non-existent in the health insurance sector, as health insurance does not insure against all loss since only the financial costs of getting sick are covered. Even if the financial loss in the form of foregone earnings and medical expenditures is covered, an individual is not insured against utility loss from being sick (Kenkel, 2000). An important counterargument for this – proposed by Dave and Kaestner (2009) – is the well reported occurrence of *ex ante* moral hazard in other insurance sectors such as the automobile and workers' compensation insurance sector, even though the argument induced by Kenkel (2000) also applies to these types of insurance.

Empirical evidence on *ex ante* moral hazard in health insurance is scarce and mixed. As a first, papers exist investigating the impact of health insurance on the demand for preventive healthcare. An example of this type of research is the paper of Kenkel (1994). Critique on this method of identification (Courbage and De Coulon, 2004) is that it fails to make a distinction between a shift in demand for preventive care due to reduced financial consequences of the health risk (which would measure *ex ante* moral hazard), and a shift in the demand for preventive care due to (partial) coverage of the costs of this care. Cherkin et al. (1990) indeed observe a decrease in demand for preventive care when small co-payments are introduced for this type of care. As these two effects are offsetting, identification of the *ex ante* moral hazard effect by analysing demand for preventive care is complicated and inconvenient.

A second group of empirical papers exists. This group of papers follows the reasoning of Kenkel (2000), which implies that the true *ex ante* moral hazard effect can be identified by researching the effect of health insurance on various lifestyle choices at the individual level, as these activities of prevention are not insured. Authors of these papers all take into consideration the effect of endogeneity that might be a problem when identifying the relationship between lifestyle choices and insurance. They use different techniques to capture the true *ex ante* moral hazard effect. I will now discuss the main papers investigating lifestyle *ex ante* moral hazard empirically.

Courbage and de Coulon (2004) fail to confirm the existence of *ex ante* moral hazard in the UK. The two behavioural preventive variables researched are exercising and smoking. As a benchmark to compare with previous research on *ex ante* moral hazard that mostly focuses on the demand for preventive care, the authors also investigate two regular preventive care variables for females. Besides using a simple probit model, to cope with the endogeneity problem discussed previously, this research uses the instrumental variable (IV) estimation technique. The authors use the following instruments: either conservative or labour political party support and being employed. Unfortunately, the results on smoking behaviour found by Courbage and de Coulon (2004) turn out to be invalid as the instruments used are not exogenous. The results they find on the other behavioural variable, exercising, are partly significant and do not point towards the existence of *ex ante* moral hazard. In my opinion it is unsurprising that the combination of instruments used in this paper turn out to be invalid, in particular for the lifestyle variable smoking. The authors assume the combination of instruments they use to be uncorrelated with risk aversion, as they do not correct for this. Though in reality, the choice of being employed is likely to be related to risk aversion.

Accepting a permanent contract over a *ZZP* position or over starting a business alone is a more logical choice for a risk averse individual than for a risk seeking person. Smoking is not unlikely to be related to risk aversion as well. Therefore it is not surprising that smoking behaviour and employment choice are related through the error term, which includes unobserved risk aversion. Courbage and de Coulon (2004) find that the opposite of *ex ante* moral hazard might be the case for the variable exercising, their results point out that being privately insured increases the probability for individuals to exercise.

Stanciole (2008) executed research on the American health insurance market, and was able to identify the existence of *ex ante* moral hazard. He investigates four lifestyle variables, namely heavy drinking, heavy smoking, lack of exercise and obesity. To allow for endogeneity from unobserved variables, he uses both a bivariate probit and a multivariate probit model. The author finds that except for heavy drinking, all variables are sensitive to *ex ante* moral hazard. Results from the bivariate and multivariate model are similar. Additionally, the latter model identifies correlation between the different lifestyle variables. Stanciole (2008) uses the variable *score* to suffice as instrument for the uptake of insurance. This variable consists of a measure for arthritis and rheumatism, one for mental health inventory and one for depression. The instrument does not seem exogenous to lifestyle. Firstly, lifestyle behaviour – and in particular smoking – is assumed to be highly correlated with mental health. Secondly, the lifestyle variable physical inactivity is likely to be correlated with whether an individual suffers from rheumatism/arthritis. Suffering this type of disease might complicate one's ability to perform any form of exercise. Similarly, this instrument might be indirectly correlated with obesity. In my opinion this is a weak aspect of this paper. By bundling the three measures into one instrumental variable, Stanciole (2008) makes it impossible to perform a test for over-identifying restrictions such as the Sargan test or to individually evaluate the effect of the different elements of which this instrument consists. This seems like an illogical choice.

The third large research executed within this second group uses an exogenous shock in the insurance level to identify the true *ex ante* moral hazard effect. Dave and Kaestner (2009), focus on the US Medicare program that covers all American residents aged 65 and over. The following lifestyle variables are taken into consideration; alcohol consumption, smoking behaviour and physical exercise. The authors control for the previously discussed indirect effect of obtaining health insurance, and find that men aged 65 and older indeed reduce prevention activities since unhealthy lifestyle behaviours are increased once insured. In other words, *ex ante* moral hazard is observed, but only amongst men. This *ex ante* moral hazard

effect is largely offset by the indirect effect from a larger number of doctors' visits creating awareness of lifestyle impact on health. A less strong aspect of this paper is the lack of control variables, there is not corrected for state of health in the form of the most commonly observed NCD's or self-reflected state of health. An individual's health state is likely to both affect how someone behaves and whether someone has already signed up for a private health insurance contract prior to entering the Medicare program. The control group used in this research consists of uninsured individuals aged 60 to 64. I suspect this group to be a selection of the more healthy individuals. For example someone who has been diagnosed with lung cancer two years ago is likely to have quit smoking and is also likely to have signed a health insurance contract last year. The same principle applies to someone suffering another chronic condition. I do not believe that the dummy variables measuring doctor- and hospital visits capture the effect that individual health state – risk type – has, both on lifestyle and choice to be insured prior to the age of 65. Besides this it is noteworthy that the demographic characteristics of the control group are quite different from the group of insured individuals prior to age 65. Especially when considering assets: the control group possesses on average around 50% of the assets that the group that is insured prior to the age of 65 owns.

III. The Dutch health insurance system

In the Netherlands, an elaborate health insurance system with some unique features is maintained. It consists of obligatory basic health insurance and non-obligatory supplementary insurance. Basic health insurance is offered on a private market, within a unique set of elaborate regulations to manage competition and protect clients. On the other hand, supplementary insurance is completely private in nature. The market for supplementary insurance is relatively small and is more likely to suffer adverse selection problems. This thesis focuses on Dutch basic health insurance, there are some features that make this system interesting and that make it appropriate for this type of research in particular.

Mandatory basic health insurance

As previously mentioned, in the Netherlands, basic health insurance is offered on a strongly regulated private market. All Dutch residents aged over 17 are obligated to have a basic health insurance contract with one of the private health insurers. The government decides on the benefits that are covered in the contracts offered by these insurers, consequently, insurers compete on prices to attract customers. To be able to offer the best

price and quality possible, insurers are allowed to discuss with care providers on the nature and price of the services they offer. This way, supplier induced demand is limited in the Netherlands. *Risk rating* is strongly forbidden in the Dutch health insurance system, individuals must be accepted into an insurance contract no matter what. To compensate health insurers for holding on to a large share of risky customers and possibly incurring a large amount of costs, there exists a *risk adjustment fund*. Through this fund, a distribution of money takes place to compensate those insurers with a large share of high-risk individuals in their portfolio. The risk adjustment fund functions as a reasonably useful tool preventing insurers from feeling a strong urge to attract low risk customers (Nederlandse Zorgautoriteit, 2017). The Dutch health insurance system is financed for a large part through insurance premiums that Dutch individuals pay directly to their insurer, low-income individuals receive subsidy to be able to afford this. The remaining part of the budget comes from government- and employer contributions. These contributions are distributed amongst the health insurers via the risk adjustment fund.

Limiting moral hazard

In the Netherlands, a deductible is used to tackle the effects of both *ex ante* and *ex post* moral hazard. The Dutch deductible is unique in the sense that it is built up from both an obligatory and a voluntary deductible amount. Every year in November, individuals aged over 17 are asked to choose an insurance contract and pick a voluntary deductible level for the coming year. Individuals are allowed to choose between a voluntary deductible of respectively 0, 100, 200, 300, 400 or 500 euros on top of the obligatory deductible. A higher level of voluntary deductible is compensated for with a lower level of insurance premium. In Appendix A, the development of the Dutch obligatory deductible in euros can be quickly observed. Noteworthy is the growth over the years in the obligatory deductible set by the government. Simultaneously, a growth in the share of individuals opting for a positive voluntary deductible is observed (Trouw, 2017; CBS, 2017). Whether the Dutch voluntary deductible actually tackles moral hazard remains a question of interest. Each year, individuals might adjust their deductible level for the year to come with help of an estimate of the amount of healthcare they might use the coming year. This way, the Dutch deductible provides individuals with the opportunity to minimize their costs by using their private information on, amongst others, risk type and lifestyle behaviour when setting their deductible for the year to come. According to the expected utility theory, low risk/healthy individuals would profit from picking a maximum voluntary deductible to minimize their premium payments, as they don't

expect to incur any healthcare costs. High-risk/less healthy individuals on the other hand, might pick a zero voluntary deductible to minimize the financial consequences from the costs they expect to make. These mechanisms lead to foregone income for health insurers and an upward pressure on the budget necessary to finance the health budget (Kapteyn and Teppa, 2011). On the other hand, according to empirical research by *De Nederlandsche Bank* Dutch residents tend to over-insure. This indicates that in reality risk aversion might play a larger part in individual Dutch deductible choice than risk type does (Gorter and Schilp, 2012).

Within the Netherlands, visiting the general practitioner is free for all residents. The general practitioner functions as a gatekeeper to the rest of the healthcare system; a referral from the general practitioner is generally necessary in order to gain access to other forms of care. This mechanism is supposed to decrease the occurrence of *ex post* moral hazard. A comparable additional measure to tackle *ex ante* moral hazard does not exist in the Netherlands.

IV. Research Framework and Hypotheses

Summing up, previous empirical evidence on lifestyle *ex ante* moral hazard is scarce and rather mixed. Even though theoretically the existence of *ex ante* moral hazard is a well-known phenomenon, the question remains whether it occurs in a real world setting. Consequently, it is unclear whether and to what extent policy makers and insurers should take this into consideration when developing strategy. Several implications can be derived from previous empirical research, but there is still a lot of work left to do. The stream of papers that investigate *ex ante* moral hazard considering lifestyle behaviour is strong in the sense that these empirical researches try to identify the true *ex ante* moral hazard effect whilst correcting for endogeneity from either risk aversion, risk type or observable characteristics. Though the different methods used to do so are valuable in essence, there are some points of critique on each of these papers. In this thesis I will therefore try to create clarity on the occurrence of *ex ante* moral hazard in a real world setting, making use of a well thought-through method consisting of a combination of aspects used in some of previously mentioned papers. My method consists mostly of aspects from the Stanciole (2008) paper, identifying *ex ante* moral hazard using lifestyle behaviour. Besides this, I employ some aspects of the methodology of Alessie et al. (2018).

The Dutch health insurance market is unique in multiple aspects. Through multiple checks and balances, the Dutch government makes sure that community rating is applied by

private insurance parties, covering all Dutch residents. Perhaps the most unique selling point of the Dutch health insurance system is the option for individuals to choose a certain voluntary deductible on top of an obligatory deductible. This makes the Dutch health insurance system extremely interesting to investigate when considering *ex ante* moral hazard, as different levels of insurance coverage can be considered. This way, new insights on the importance of financial (dis-)incentives for the occurrence of *ex ante* moral hazard can be deduced. Therefore the main research question addressed in this thesis is as follows.

To what extent does ex ante moral hazard occur on the Dutch health insurance market?

By answering this question, I hope to fulfil my two main research objectives. As a first, this way I want to contribute to the scarce empirical evidence on the existence and nature of *ex ante* moral hazard. Second, I hope to provide insight on whether the Dutch deductible can serve as a useful tool to stimulate self-prevention in the form of a healthy lifestyle. So far, it has been unclear whether the Dutch deductible actually reduces any form of moral hazard or is used mostly as a tool by low risk consumers to minimize their own costs. With the results of my thesis, I hope to develop an advice for Dutch policy makers and health insurers, providing direction in the choice for obligatory and voluntary consumer deductibles in health insurance during the years to come.

Hypothesis 1: There exists ex ante moral hazard in the Dutch health insurance market.

The main hypothesis follows directly from the research question. To be able to find an answer to this question, I identify the relation between the height of the Dutch deductible and the main lifestyle variables that increase an individual's chance of developing an NCD (Beaglehole et al., 2011; WHO, 2003, 2005). These lifestyle variables consist of obesity; total lack of physical exercise (inactivity); heavy smoking behaviour and excessive alcohol consumption. Lifestyle variables I include are similar to the research Stanciole (2008) executes on the American health insurance market and overlap at least partly with all other papers relating lifestyle to insurance coverage. This provides me with the opportunity to compare my results to previous work. Within the section on data and in *table 1*, I provide elaborate definitions and background information on the construction of the different variables used. The four sub-hypotheses that need to be confirmed or rejected to provide an answer to the main research question concern the lifestyle variables. Respectively, L^1 corresponds to obesity, L^2 corresponds to inactivity, L^3 corresponds to heavy smoking and L^4 corresponds to heavy drinking.

Sub-hypothesis 1^{I}: In the Netherlands, individual choices on lifestyle variable L^{I} are influenced by the choice to either opt for a zero or non-zero voluntary deductible.

To establish whether *ex ante moral* hazard happens on the Dutch health insurance market, it is necessary to isolate the causal relation between lifestyle and the Dutch deductible. This means that variables that possibly are related both to the voluntary deductible an individual chooses and the lifestyle that this same person adopts, should be included in the analysis. More formally, a possible effect of endogeneity originating from omitted variables influencing both variables of interest should be excluded. Theory and previous empirical research on this topic has proven that this can be extremely challenging. There are a few main categories of variables that should be corrected for, as was pointed out within the literature and previous empirical evidence. As emphasized by Rothschild and Stiglitz (1976), it is necessary to capture private information on individual risk type. Therefore, I control for lagged health state, measuring the risk type of an individual at the moment of voluntary deductible choice. Secondly, a measure of risk aversion should be included to capture the effect of preference for insurance (Finkelstein and McGarry, 2006). Thirdly, I correct for observable characteristics indicating socio-economic status. Again I use lagged variables to capture the impact of these socio-economic variables at the moment of voluntary deductible choice. To sum up, similarly to Alessie et al. (2018) I control for three selection sources: individual risk type, individual risk preferences and socio-economic status.

After controlling for these sources of selection, there is still a possibility that endogeneity from omitted variables exists. For example, both a particular lifestyle variable and the voluntary deductible choice might be related to another lifestyle variable. Stanciole (2008) finds that this is indeed the case in the US. To avoid the possible effect of unobserved variables that both correlate with the individual's choice on a level of voluntary deductible and with lifestyle behaviour, I estimate the relation between the variables of interest using a model that corrects for *any* endogeneity from omitted variables. A second concern to take into consideration is the possibility of reverse causality or simultaneity. I aim to estimate the effect of the voluntary deductible choice on lifestyle. To make sure the relation I identify is causal and does not suffer from endogeneity, I use variables to instrument for the voluntary deductible choice. These instrumental variables should be strongly correlated with the voluntary deductible choice whilst being exogenous to each of the four lifestyle variables.

Employing a method that takes these concerns into consideration, I will use the four sub-hypotheses to be able to confirm or reject the main hypothesis. My main aim is to identify

the *ex ante* moral hazard effect in the market for Dutch health insurance. Simultaneously sketching an elaborate picture on what aspects of lifestyle are sensitive to the financial incentive that the deductible represents, and allowing me to draw conclusions on how the voluntary deductible is related to lifestyle.

V. Data and Methodology

V.I. Data

This study makes use of panel data from the Dutch LISS panel (the Longitudinal Internet Studies for Social sciences), which is a survey database containing data on a large range of topics gathered yearly from a large sample representative of Dutch society. Around 8000 individuals are asked yearly to submit their answers on surveys considering multiple relevant topics. These individuals are selected as a realistic representation of the total Dutch population, based on their observable characteristics. Both the core topics addressed by the LISS panel, and the sample included in the LISS panel remain relatively constant over the years, making the data appropriate to do econometric analysis (balanced dataset), capturing both within- and across time trends.

Both the key variables and control variables necessary for this study can be found in different studies of the LISS panel. Beneath I explain which data can be found in what LISS study, and how I construct the variables from this data. More generally speaking, similarly to Alessie et al. (2018), this study combines data from the yearly health study, the regularly updated background statistics and a one-time study named “Measuring Higher Order Risk Attitudes of the General Population” (further: “Risk Attitudes”). Furthermore I include extra control variables originating from the core study on work and education and I include an instrumental variable from the core study concerning personality.

The key variables considered in this research are the choice to opt for a voluntary deductible and the following previously mentioned lifestyle variables: obesity, total lack of physical exercise (further: inactivity), heavy smoking and heavy drinking. In subsection V.I.I., I explain the construction of and the economic rational behind the key variables in more detail, furthermore the definitions that are used in this study can be found in *table 1*. These key variables can all be deduced from the health study, which is one of the core topics of the LISS panel. Appropriate data on these key variables is available for the years 2009-2013 and 2015-2017, the LISS panel addresses this data as wave 3-10. Unfortunately, the questionnaire before the year 2009 was formulated as such that the individual choice for voluntary

deductible could not be deduced, making these waves inappropriate for this research. Also, in 2014 the health study was not administered. Therefore, this study is based on the years 2009-2017 with the exception of 2014.

Numerous control variables, capturing three types of individual sources of selection are included in this study as well. It is of main concern to correct for the two types of asymmetric information in deductible choice: risk type and risk aversion. These are the first two individual sources of selection. To capture the individual risk type, several control variables representing individual health risk are constructed. These variables are labelled under the category *lagged health status*. Than, to capture risk aversion, a control variable measuring individual risk attitudes is constructed. Unlike this first control variable, this measure of risk aversion is assumed to be constant over time. Than the third individual source of selection is socio-demographic state. To capture this, multiple observable characteristics are included in the analysis. These characteristics are socio-economic and demographic in nature and are measured yearly at the time the voluntary deductible choice takes place. The economic rationale behind and the construction of the control variables is further discussed in subsection V.I.II, besides this, the definitions per variable can be found in *table 1*.

The merged dataset contains 98.768 observations, distributed fairly equally over the years. I exclude all respondents aged beneath 18, these individual do not manage their own insurance choices yet and do not bear any financial responsibility, making them inappropriate for this analysis. After dropping all observations on respondents aged beneath 18, I am left with 80.805 observations.

Key variables

To identify whether the uptake of voluntary deductible affects individual lifestyle behaviour, it is important to monitor different levels of voluntary deductible. Each year, people can choose between a voluntary deductible of 0, 100, 200, 300, 400 or 500 euros. For the convenience of this study, I construct a binary variable on voluntary deductible with the following definition. A zero indicates a voluntary deductible equal to zero, whereas a one indicates a positive voluntary deductible. The main reason to use this definition to measure voluntary deductible choice is that opting for a high voluntary deductible has only since recently gained popularity amongst Dutch citizens (Trouw, 2017; FD, 2017). For this reason I am afraid that results concerning opting for a maximum voluntary deductible versus opting for a lower voluntary deductible will be sensitive to bias since it concerns a relatively small group of individuals mainly concentrated in the last years of my sample.

The first lifestyle variable included in this research is being obese. Currently, over 50% of Dutch adult citizens is overweight and over 14% is obese (CBS, 2017), this is a shockingly large number considering the health risks that being overweight impose (WHO, 2003), both on long and short term. From the health study I calculated the body mass index (BMI) per individual. A healthy BMI is supposed to be no less than 18.5 and no more than 25. Obesity is defined as having a BMI exceeding 30. Therefore, I categorized being obese as having a BMI larger than 30.

To capture physical inactivity I measure a total lack of light, moderate and heavy physical exercise. The Dutch Institute of Sport and Movement (in Dutch the NISB), states that the minimum movement for adults is exercising light to moderate movements for 30 minutes on five days per week. This adds up to two and a half hour of movement per week. The LISS panel only measures these types of exercise when executed more than 10 minutes in a row. This complicates constructing a variable measuring lack of movement, as *shortly* walking the stairs; walking your dog; walking to a close by shop; etcetera, can add up to 30 minutes at least on some days. I therefore define total lack of exercise as never spending more than 10 minutes in a row on light, moderate or heavy physical exercise. Since it is quite safe to say that if someone never performs any form of exercise more than 10 minutes in a row, this person does not comply with the NISB measure for minimum movement per week. The CBS (2017) uses a similar definition for inactivity.

The third lifestyle variable, heavy smoking, currently is defined as smoking at least 20 cigarettes per day on average. This definition is used in recent editions from medical journals; also the Dutch CBS (2017) applies this definition when measuring heavy smoking statistics. Besides this it is convenient that in previous research concerning lifestyle *ex ante* moral hazard, this definition is used as well (Stanciole, 2008). In this study, I include smoking of cigars, cigarettes, e-cigarettes and pipes.

Heavy drinking is the last lifestyle variable considered. Note that this is not necessarily the same as alcoholism, which is defined as an addiction causing either physical or mental dependence on the drug (Morse and Flavin, 1992). Heavy drinking can be defined as consuming large amounts of alcohol, either through drinking quite regular amounts very frequently or drinking excessive amounts incidentally (WHO). Unfortunately, the LISS panel does not allow for a measure on average number of alcoholic beverages per day. Therefore in this study, I use the following definition when measuring heavy drinking: consuming at least one alcoholic beverage at least five days per week on average *and* consuming more than five alcoholic beverages on the day of this week that most alcohol was consumed. This way, both

very frequent drinking (in small to large amounts) and drinking excessive amounts at least once a week is included in the criterion to measure heavy drinking.

Control variables

Additively to this data from the key health study, I use control variables concerning the lagged health state of individuals. The lagged health state is an indicator of the risk type of an individual. As previously discussed, this is unobservable to health insurers, or at least it is forbidden to act on this for insurers (Rothschild & Stiglitz 1976), but it is known to individuals when opting for a voluntary deductible. These variables are two self-reflective variables on physical and mental health and five objective variables concerning different types of NCD's. The two self-reflective variables are *good self assessed health* and *emotional problems*. The five types of physical condition I consider are *chronic condition, diabetes, cancer* (including only serious forms), *lung disease, heart problems*. Where this last variable includes cases of heart attack, stroke, dangerously high blood pressure and other types of heart and vascular disease. All lagged health state variables are dummy variables taking on either the value of zero – no such condition – or one – indeed suffering from this type of condition. For an overview and brief description of these variables I refer to *table 1*. Because the risk type is of relevance when choosing a voluntary deductible for the year to come, I measure the information on these variables November/December the year before lifestyle and deductible are analysed. Otherwise, it would be unclear whether individuals have included this in their voluntary deductible choice.

Besides controlling for *risk type* in the form of lagged health state, I control for individual *risk preferences* with the variable risk aversion. This variable is constructed from the Risk Attitudes survey from 2009. This survey contains details on the number of safe and risky choices individuals make when facing several lottery choices. Following Noussair et al. (2013), and Alessie et al. (2018), I model a constant measure of risk aversion as the number of safe choices made in five lotteries. Thus assuming that risk aversion is constant over time, similarly to Alessie et al. (2018). The main reason for this is the lack of availability of data on risk preferences in other years than 2009. This is a disadvantage of the data that I use. In the risk preferences questionnaire, choosing to receive a fixed amount with certainty represents the safe choice. Through the different lotteries presented in the questionnaire, this amount varies from 20 (first question) to 40 (last question) euros. On the contrary, in each question the risky option is choosing to be unsure between receiving either 5 or 65 euros. The structure of the lottery questionnaire makes it irrational to switch more than once from risky to safe, if

switching at all. Within this questionnaire, it also would be irrational to switch from safe to risky, as the sure outcome increases throughout the questionnaire whilst the risky outcomes stay equal. To correct for individual inconsistencies in the safe choices made, I exclude all observations that show either one or both of these inconsistency types. This reduces my observations by 571, leaving me with 2883 consistent observations in total. Now I construct a dummy variable capturing risk aversion. Within this risk aversion parameter, individuals performing more than two safe choices, with no inconsistencies, are classified as risk averse. On the other hand, individuals performing less than or exactly two out of five safe choices without inconsistencies, are classified as risk seeking to risk neutral. This leaves me with a proportion of 73% risk averse individuals. Given previous research performed by *De Nederlandsche Bank*, finding that 74% of Dutch inhabitants is risk averse (Kapteyn and Teppa, 2011), I conclude that this is a realistic proportion and that my measure for risk aversion is an appropriate control variable. Unfortunately, important to mention is that Kapteyn and Teppa (2011) find that a lottery based risk aversion control measure might not function as well as a more intuitive measure based on factor analysis. Due to data unavailability I am forced to make use of such a measure, though this is a possible weakness of the research I conduct.

Besides these two types of control variables, I take into consideration several observable demographic and social characteristics. These are derived from the background variables in the LISS panel, which are monthly updated by the head of each household of which members participate in the LISS panel. The variable indicating whether an individual is a student in the year considered is constructed from the work and education core study. Again I make use of lagged control variables, by each time using the information from November the year before the analysis of lifestyle and voluntary deductible takes place. An overview and clear definitions of these variables can be found in *table 1*.

Instrumental variables

To cope with potential endogeneity after controlling for previously mentioned sources of selection I employ instrumental variables for the uptake of a voluntary deductible. Making use of instruments to measure individual insurance level is a common approach, but as I explained in the literature section, instruments used are often questionable. I have developed three instruments that have not yet been used to capture the *ex ante* moral hazard effect before. The instruments are derived from different questionnaires. Testing of these three instruments indicates that the third instrument – financial literacy – is non-relevant. Therefore

I use the first two instruments within the analysis. As I do believe that *financial literacy* is an important possible instrument in this research area, I have reported the construction of this instrument in this section. Within the discussion (section VI.III) I elaborate on points of improvement for such an instrument in future research.

The uptake of *complementary health insurance* is the first and trickiest instrument considered. Complementary insurance has been used to instrument for insurance level before, but only in research trying to measure *ex post* moral hazard (Alessie et al., 2018). Rationally speaking, the uptake of complementary insurance is highly likely to be a relevant instrument for the uptake of a voluntary deductible. An individual with consistent decision-making would not be likely to buy of financial risk of needing a certain type of supplementary care by accepting a complementary health insurance contract, whilst simultaneously opting for extra financial risk in the basic health insurance by opting for a positive voluntary deductible. Therefore, I expect the uptake of complementary health insurance to be a sufficient predictor for the uptake of a voluntary deductible in health insurance. Within the Netherlands, complementary insurance does capture a completely different package than basis insurance does. For this reason, when investigating the relationship between basic package healthcare use and voluntary deductible, the uptake of complementary insurance is highly likely to be independent of basic package healthcare use. Meaning that this can be considered a valid, exogenous instrument. When considering the relationship between lifestyle and voluntary deductible however, this is not necessarily the case. In the Netherlands, the most common forms of complementary insurance are dental insurance and physiotherapy coverage. This last form of complementary insurance is likely to be correlated to at least two out of four considered lifestyle variables: obesity and inactivity. Healthy living individuals that exercise frequently or play a form of sports might be more likely for the uptake of physiotherapy insurance, because of awareness of higher risks for injury or a past of injuries and physiotherapy use. On the other hand, people with fragile physical features such as less serious forms of back injuries (not necessarily covered by the control variables) are more likely for the uptake of physiotherapy insurance as well, while at the same time their chance of being obese and inactive might be higher. Therefore, I have included an extra control variable when using this instrument, namely needing *physiotherapy* in the year that the voluntary deductible choice is made. This way, I expect the need for complementary insurance to be exogenous to lifestyle in this estimation. Making this instrument both valid and relevant. Both *complementary insurance* and *physiotherapy* are derived from the health survey from the LISS panel and constructed as a dummy variable.

The second instrument that I use is *extreme pessimism*. The rationale behind the choice for this instrument is straightforward: an individual expecting mostly bad things to happen to him/her would be unlikely to opt for additional financial risk in any field. Therefore I expect extreme pessimism to be an appropriate predictor (relevant instrument) of voluntary deductible choice. At the same time, there is doubt on the relationship between pessimism and lifestyle. On the one hand, pessimist individuals are more likely to be unhappy. Unhappy people are more likely to be depressed or experience other forms of emotional problems and might therefore be more likely to adopt unhealthy lifestyle features. This is not unthinkable as association between unhealthy lifestyle and mental health problems is established by among others Walsh (2011). On the other hand, in the extreme case the avoidance of risk taking in the financial field might be applicable to the field of lifestyle behaviour as well. If a pessimist individual is risk averse enough, this individual might avoid the risk of an unhealthy lifestyle. As I control both for emotional problems and for risk aversion in my analysis, I expect extreme pessimism to be able to suffice as an exogenous instrument, being uncorrelated with lifestyle via the error term. I construct this instrument with help of the LISS survey on personality traits, the survey is constructed as such that the years 2012, 2013, 2015 and 2017 are usable for construction of this variable. Individuals are asked to state how likely it is on a categorical scale of 1 to 5 that generally bad or good things are going to happen to them in the future. I define an individual as being extremely pessimist if on at least six out of seven questions the negative answer is chosen.

The third and last instrument that I investigate is *financial literacy*. As previously discussed, Dutch residents tend to over-insure (Gorter and Schilp, 2012). Besides being an indicator of risk aversion, this might also be an indicator of lack of financial knowledge: there might be a sufficient proportion of Dutch residents that is unaware of the individual possible gains or losses of taking up a non-zero voluntary deductible. People that possess a very small stock of financial knowledge and are insecure about this knowledge stock are unlikely to engage actively in the financial stock market (van Rooij et al., 2011). For this reason, these low financially literate individuals might also be unlikely to actively take up a voluntary deductible. This is simply something they are not interested in and it might even scare them to engage in any form of financial risk taking, as they do not understand the concept. Therefore, I expect the combination of financial literacy and confidence in one's own financial literacy to be a good predictor and therefore a relevant instrument for the uptake of a voluntary deductible. The relationship of financial knowledge and confidence in this area with lifestyle behaviour might be non-zero because both variables are highly likely to be correlated with

educational level. As I control both for high educational level and low educational level, I expect this to be no problem. This makes financial literacy an appropriate instrument to test for the uptake of voluntary deductible. I construct this variable by using the single wave LISS study on financial literacy, conducted in 2011. In this study, four questions measuring financial literacy are imposed, and one question on confidence of ones own financial literacy is stated. The variable on financial literacy and confidence that I construct takes on the value of one if a person rates his/her financial knowledge at least a five on a scale of seven whilst simultaneously being able to answer at least three out of four financial literacy questions correct. If these conditions are not met, the variable takes on zero.

V.II. Methodology

Univariate probit model

The first model I apply incorporates the possibility for individuals to use private information on multiple dimensions when determining on a level of voluntary deductible. This means that within this *univariate probit model* I estimate *ex ante* moral hazard while controlling for the defined sources of selection: risk type (represented by lagged health state), risk preferences (represented by the number of safe choices made) and socio-economic status (represented by lagged background variables). Here, the uptake of voluntary deductible is treated as exogenous to lifestyle after controlling for these three sources of selection. Besides correcting for the previously mentioned sources of selection via vector $\boldsymbol{x}'_{\boldsymbol{it}}$, I allow for intragroup correlation by computing standard errors which are clustered at the individual level. Taking into account that I use repeated observations on the same individuals, clustering is necessary to obtain robust variance estimates.

$$L_{it}^{I} = \begin{cases} 1 \text{ if } \lambda^{I} d_{it} + \boldsymbol{x}'_{\boldsymbol{it}} \boldsymbol{\kappa}^{\boldsymbol{I}} + \varepsilon_{it}^{I} > 0 \\ 0 \text{ otherwise} \end{cases} \quad (1)$$

Within equation (1), I = 1, 2, 3, 4. Where L_{it}^{I} indicates whether person i has opted for an unhealthy lifestyle on lifestyle variable I in year t. d_{it} represents a binary variable indicating whether this same person has opted for a non-zero voluntary deductible in year t. For each lifestyle variable I the *ex ante* moral hazard effect is represented by λ^{I}. $\boldsymbol{x}'_{\boldsymbol{it}}$ represents a vector of explanatory variables, this includes all the sources of selection previously discussed, $\boldsymbol{\kappa}^{\boldsymbol{I}}$ is a vector of coefficients. Conditional upon $\boldsymbol{x}'_{\boldsymbol{it}}$ and d_{it}, the error term ε_{it}^{I} is standard normally

distributed. This implies that the voluntary deductible choice, d_{it}, is treated as an exogenous variable.

Bivariate probit model with instrumental variables

Now I introduce a model that corrects for possible endogeneity after controlling for the three sources of selection and clustering on the individual level. The model I use is a bivariate probit with instrumental variables. There are two important advantages of applying this model. Firstly, this provides me with the opportunity to prove that there indeed exists endogeneity after controlling for the selection sources. Second, this enables me to identify the causal relation between d_{it} and L_{it}^{I}.

Bivariate probit models are built both on a structural form equation for the voluntary deductible chosen, d_{it}^{*}, and a reduced form equation for each of the lifestyle choices considered, L_{it}^{I}. When applying this model each time one of the lifestyle variables is estimated simultaneously with the voluntary deductible, the errors are allowed to correlate freely. Now the bivariate probit model can be extended by the introduction of instrumental variables for the voluntary deductible. Thus to be valid, these instruments should be strongly correlated with the voluntary deductible (relevance), whilst being uncorrelated with lifestyle behaviour through the error term (exogenous) (Stock and Watson, 2012). As explained in the data section, I use two out of three tested instrumental variables. Instrumental variables *complementary insurance* and *extreme pessimism* are included in the structural equation for the voluntary deductible uptake. Given the observed covariates $\boldsymbol{x_{it}'}$ and the instrumental variables z_{it}, the bivariate probit model with instrumental variables can be specified as equation (2).

$$\begin{cases} d_{it}^{*} = \pi z_{it} + \boldsymbol{x_{it}'\gamma_1} + v_{it} \\ L_{it}^{I*} = \delta^{I}\, d_{it} + \boldsymbol{x_{it}'\gamma_2} + u_{it}^{I} \end{cases} \tag{2}$$

Here π represents a vector of the coefficients of the exogenous and relevant variables z_{it} that are used to instrument the choice to either opt for a zero or non-zero voluntary deductible, γ_1 and γ_2 are vectors of coefficients for the covariates $\boldsymbol{x_{it}'}$**.** If the application of valid instruments in this model leads to a significant non-zero value of ρ, the *ex ante* moral hazard effect can be found in the coefficient for the effect of the voluntary deductible choice, δ^{I} . The error terms are represented by v_{it} and u_{it}^{I}. The observations in this extended model range according to equation (3).

$$d_{it} = \begin{cases} 1 \; if \; d_{it}^{*} > 0 \\ 0 \; otherwise \end{cases} \qquad (3)$$

$$L_{it}^{I} = \begin{cases} 1 \; if \; L_{it}^{I*} > 0 \\ 0 \; otherwise \end{cases}$$

The bivariate probit model with instrumental variables consists of a bi-dimensional integral with a closed form solution over the distributions of the residuals. Assuming the error terms are drawn from a standard bivariate normal distribution with zero means, unit variances and correlation coefficient ρ. Where a significant value of $\rho \neq 0$ implies that the error terms of the equations under equation (2) are significantly correlated. Meaning that the voluntary deductible choice indeed is endogenous in lifestyle equation *I*, and that there thus is sufficient reason to use a bivariate probit model to estimate the *ex ante* moral hazard effect. In equation (4) a clear representation of the error term distribution in the bivariate probit model can be found. Similarly as in the previously formulated univariate probit model, results are clustered on the individual level.

$$(v_{it}, u_{it}^{I}) \sim N_2 \left(\binom{0}{0}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right) \qquad (4)$$

The method of maximum likelihood is used to consistently estimate the unknown parameters in this model. With a correctly specified log likelihood function, this method leads to consistent asymptotically normal and efficient estimators for the parameters.

VI. Analysis

Within this section a thorough interpretation of the results from the analysis can be found. First, I provide the descriptive characteristics of the data that I use. Alongside giving an interpretation of these statistics, highlighting some of the more and less realistic features I encounter in the dataset. Then secondly I provide both an overview and a detailed interpretation of the outcomes the models I apply. The results allow me to either accept or reject the four sub-hypotheses concerning the relation between the choice for a zero or non-zero voluntary deductible and the choice for a certain lifestyle.

Table 1 Variable definitions and descriptive statistics

Variable name	*Definition*	*Sample*		
		Full	*Vol ded > 0*	*Vol ded = 0*
Lifestyle behaviour				
Obese	1 if body mass index exceeds 30, 0 otherwise	0.150	0.107	0.158
		(0.009)	(0.015)	(0.010)
Inactive	1 if never engages in any physical activity, 0 otherwise	0.319	0.329	0.317
		(0.009)	(0.020)	(0.009)
Heavy smoker	1 if smokes 20+ cigarettes/cigars/pipes per day, 0 otherwise	0.046	0.043	0.046
		(0.005)	(0.011)	(0.006)
Heavy drinker	1 if drinks daily and has over nine alcoholic drinks at least once a week, 0 otherwise	0.059	0.064	0.059
		(0.005	(0.012)	(0.006)
Deductible				
Voluntary deductible	1 if an individual has a voluntary deductible, 0 otherwise	0.161	1.000	0.000
		(0.008)	(0.000)	(0.000)
Voluntary deductible size	Size of voluntary deductible, ranging from 0, 100, 200, 300, 400, 500.	57.327	355.937	0.000
		(3.250)	(6.967)	(0.000)
Mandatory deductible size	Size of mandatory deductible	327.523	326.717	327.678
		(0.506)	(2.223)	(0.639)
Total deductible	Size of total deductible	384.850	682.654	327.678
		(3.310)	(8.131)	(0.639)
Lagged socio-demographic variables				
Age	Age in number of years	57.553	51.841	58.649
		(0.381)	(0.751)	(0.392)
Age^2	Squared age in number of years	3,515.179	2,891.947	3,634.827
		(43.112)	(78.109)	(45.104)
Male	1 if male, 0 if female	0.518	0.597	0.503
		(0.014)	(0.027)	(0.015)
Kids	1 if the individual has kids, 0 otherwise	0.297	0.386	0.280
		(0.012)	(0.027)	(0.013)
Married	1 if married, 0 otherwise	0.678	0.621	0.689
		(0.013)	(0.027)	(0.013)
Education low	1 if up to intermediate secondary education, 0 if higher	0.298	0.213	0.314
		(0.013)	(0.022)	(0.014)
Education high	1 if up to higher vocational education or university, 0 if lower	0.382	0.512	0.358
		(0.014)	(0.028)	(0.014)
Student	1 if goes to University this year, 0 otherwise	0.003	0.003	0.002
		(0.001)	(0.002)	(0.001)
Self-employment	1 if has ZZP function or is self-employed, 0 otherwise	0.045	0.063	0.042
		(0.005)	(0.013)	(0.005)
Unemployed	1 if unemployed, 0 otherwise	0.100	0.076	0.104
		(0.007)	(0.013)	(0.008)
Retired	1 if retired, 0 otherwise	0.310	0.200	0.331

		(0.012)	(0.020)	(0.013)
Log income	Log of netto household income	7.878	7.969	7.860
		(0.012)	(0.026)	(0.013)
Countryside	1 if lives at country side, 0 if lives at more rural area	0.147	0.120	0.152
		(0.010)	(0.018)	(0.011)
Lagged health variables				
Good self assessed health	1 if very good or excellent self assessed health, 0 otherwise	0.829	0.910	0.814
		(0.009)	(0.013)	(0.010)
Emotional problems	1 if seriously suffers from emotional problems, 0 otherwise	0.084	0.066	0.088
		(0.006)	(0.011)	(0.006)
Chronic condition	1 if has a chronic condition (from disease or accident), 0 otherwise	0.369	0.243	0.393
		(0.013)	(0.022)	(0.014)
Diabetes	1 if ever diagnosed diabetes, 0 otherwise	0.065	0.037	0.071
		(0.007)	(0.008)	(0.007)
Cancer	1 if ever diagnosed cancer, 0 otherwise	0.028	0.009	0.032
		(0.004)	(0.004)	(0.004)
Lung disease	1 if ever diagnosed lung disease, 0 otherwise	0.034	0.023	0.036
		(0.005)	(0.006)	(0.005)
Heart problems	1 if ever diagnosed heart attack/disease, or vascular disease, 0 otherwise	0.238	0.146	0.255
		(0.011)	(0.018)	(0.012)
Physiotherapy	1 if needed physiotherapy at least twice that year, 0 otherwise	0.243	0.182	0.254
		(0.009)	(0.016)	(0.010)
Risk aversion				
Risk averse	1 if out of five lottery choices, at least 4 consistent safe choices, 0 otherwise	0.730	0.687	0.738
		(0.012)	(0.026)	(0.013)
Instruments				
Complementary insurance	1 if complementary health insurance contract, 0 otherwise	0.785	0.648	0.812
		(0.010)	(0.025)	(0.010)
Extreme pessimism	1 if on 5 out of 7 questions an extremely pessimist answer, 0 otherwise	0.009	0.001	0.010
		(0.002)	(0.001)	(0.002)
Financial literacy	1 if financial literacy above 40% and confident about own knowledge, 0 otherwise	0.632	0.644	0.630
		(0.013)	(0.026)	(0.014)
Number of observations at the individual level		4340	699	3641

Note. The means are tabulated without parentheses; the standard errors are clustered on the individual level and presented in parentheses.

VI.I. Descriptive Statistics

In *table 1* the summarizing statistics can be found. The table shows results for the complete set of observations as well as separately for the sub-sample with a voluntary deductible and the sub-sample without a voluntary deductible. When focusing on the lifestyle variables, in terms of heavy drinking and inactivity, individuals with a positive amount of

voluntary deductible tend to lead an unhealthier lifestyle. As can be seen, the group of individuals with a voluntary deductible consists of 1.1% of heavy drinkers versus 0.8% in the group without a voluntary deductible, similarly 32.9% of inactive individuals with a voluntary deductible versus 31.7% without a voluntary deductible. When focusing on the other two lifestyle variables, heavy smoking and being obese, the group without a voluntary deductible tends to lead a less health conscious life. 4.6% of the zero voluntary deductible group is a heavy smoker, versus 4.3% of the group with a voluntary deductible. Similarly, 15.8% of the zero voluntary deductible group is obese, whereas only 10.7% of the people with a voluntary deductible is obese. Overall, the numbers between the two subsamples lie fairly close to each other. Only the last lifestyle variable being obese shows a large difference between the two subgroups that might be seen as a first indicator of the existence of ex ante moral hazard in Dutch health insurance. In *table 2* additional information on the percentage of individuals adopting an unhealthy lifestyle on the four aspects considered can be found. For comparison in *appendix A* the proportion of unhealthy lifestyle adoption in the Netherlands can be found, as established by the CBS (2017). Respectively considering the sample used in this thesis, and the proportions of Dutch inhabitants that the CBS (2017) established. Except for one variable, the definitions used per lifestyle factor are equal in both tables. The definition used for heavy drinking differs slightly; the CBS considers heavy drinkers as individuals drinking over 5 alcoholic beverages at least once a week, I add the condition that these drinkers consume at least one alcoholic beverage at least five days per week. As discussed in the data section, reason for me to do so is that heavy drinkers are qualified as both frequent and excessive drinkers. Consequently the proportion of heavy drinkers in my sample is far below the proportion the CBS reports. The proportion of heavy smokers and obese individuals in my sample is very realistic, as these proportions are fairly close to the proportions the CBS reports. Striking is that the proportion of individuals that never perform any form of physical exercise, in my dataset is twice as high as it is in the Netherlands according to CBS (2017). Therefore, I conclude that there is bias in my dataset when it comes to inactivity. The remaining part of the lifestyle data that I use can be considered representative of Dutch society.

When considering the voluntary deductible, over the observations in the years 2009-2017 in *table 1*, 16.1% of individuals has a positive voluntary deductible. On average, this group sets a voluntary deductible of 356 euros. In *table 3*, additional information on the voluntary deductible that is chosen each year can be found. As can be seen from the observations in this study, the proportion of individuals opting for any positive voluntary

deductible increases over the years, especially the number of individuals that choose a maximum voluntary deductible has increased. This development is realistic as over the past years an increasing percentage of Dutch inhabitants opt for a maximum or high voluntary deductible (CBS, 2017; Trouw, 2017).

Considering lagged socio-demographic variables, it can be seen that the average age in the group with a voluntary deductible is lower than in the group without a voluntary deductible, as well is the proportion of females in this first group. The average level of education, household income and self-employed individuals in this first group is higher and this group contains a larger proportion of individuals with kids than the group of individuals with a zero voluntary deductible. On the other hand, the group of individuals without a voluntary deductible contains a higher proportion of married individuals, unemployed individuals, retired individuals and individuals that live at the countryside.

When focusing on lagged health state variables that determine the individual risk type, it is rather obvious that the proportion of healthy individuals in the group with a positive voluntary deductible exceeds the second group without a voluntary deductible on all health variables considered. The proportion of individuals in this study with good to excellent self-assessed health is 83%, this is very comparable to the 80% that Leu et al. (2009) observe when evaluating the Dutch and Swiss markets for health insurance. It can be observed that this proportion is 10% higher in the researched group with a voluntary deductible than in the group without a voluntary deductible. At the same time, the proportion of individuals with any of the NCD's considered is higher in the group without a voluntary deductible. These observations are an indication that individuals indeed use their private information on risk type when choosing an insurance contract.

The last control variable is risk aversion, similarly to previous results and what is pointed out by the literature, the proportion of risk averse individuals in the group without a voluntary deductible is over 5% higher than in the group with a voluntary deductible. The total proportion of risk averse individuals in my dataset is 73%, which is close to the 74% of risk averse individuals Kapteyn and Teppa (2011) observe in the Netherlands.

When considering the instrumental variables in *table 1*, it can be noted that the proportion of individuals with complementary health insurance is over 16% higher in the group without a voluntary deductible than in the group with a positive amount of voluntary deductible. Similarly, the percentage of individuals with extremely pessimistic personal traits is ten times as high in the group with a zero voluntary deductible. Lastly, the percentage of financial literate individuals is 1.4% higher in the group with a voluntary deductible.

Table 2 Proportion of unhealthy lifestyle occurrence per year, in % of total sample

Year	*Category*			
	Obese	*Inactive*	*Heavy smoker*	*Heavy drinker*
2009	13.62	34.06	6.94	4.18
2010	14.62	33.67	7.11	4.19
2011	14.3	34.25	5.76	3.74
2012	14.12	34.43	5.59	3.3
2013	14.4	34.75	4.46	3.26
2015	14.56	36.58	4.04	3.74
2016	15.29	35.41	3.97	3.11
2017	15.97	35.14	3.31	2.97
Total	14.61	34.8	5.12	3.56

In this table, the proportion of the sample that adopts unhealthy lifestyle variable(s) can be overviewed per year.

Table 3 Proportion of individuals per level of voluntary deductible per year, in % of sample

Year	*Voluntary deductible in euro*						*Total N*
	0	*100*	*200*	*300*	*400*	*500*	
2009	86.08	5.61	5.21	0.96	0.24	1.89	5811
2010	85.23	4.51	6.43	1.23	0.29	2.30	5429
2011	82.52	3.78	9.28	1.23	0.43	2.76	4862
2012	81.22	2.40	8.70	4.02	0.54	3.13	5553
2013	82.41	1.86	2.33	6.28	0.93	6.20	5161
2015	80.64	1.00	2.55	4.95	2.05	8.80	4386
2016	81.62	1.46	1.83	4.14	1.70	9.25	5190
2017	81.35	1.72	1.62	3.84	2.07	9.39	5748
Total %	82.71	2.85	4.76	3.28	1.01	5.38	
Total N	34853	1203	2006	1383	426	2269	42140

Percentage of observed individuals opting for a certain level of voluntary deductible in a certain year. As can be seen, in total, 82.71% (= 34,853) of all observations have a zero voluntary deductible, whereas the other 17.29% (= 7,287 observations) has a positive voluntary deductible. An increasing percentage of population opts for a maximum voluntary deductible, and a decreasing percentage opts for a zero voluntary deductible. This fits current trends in Dutch society as report by Trouw (4-01-2017) and CBS (17-02-2017).

VI.II Results

In this section, I discuss the outcomes of the analysis conducted. Most attention is paid to the marginal effects identified when employing the two different models. This way I provide a description and interpretation of the relation that can be observed between the voluntary deductible choice and each of the investigated lifestyle variables. Additionally, in the second model the possible identification of endogeneity between voluntary deductible choice and lifestyle choices after controlling for sources of selection is discussed. This allows me to determine which model to use when drawing conclusions from the dataset used within this research. Enabling me to individually accept or reject the sub-hypotheses, after which I can provide an answer to the main research question. This question can be answered either by accepting or rejecting the hypothesis on the occurrence of lifestyle *ex ante* moral hazard on the Dutch health insurance market. Besides this, I describe and interpret the relation between the different sources of selection and each of the four lifestyle choice variables in sub-section VI.II.II.

VI.II.I. Relating voluntary deductible uptake to lifestyle choices

Within Appendix B, *table 7* and *table 8* describe respectively the application of the univariate probit and of the bivariate probit model with instrumental variables. *Table 4* displays the key results concerning *ex ante* moral hazard effects observed in both models. The coefficients, the related standard errors, significance and the marginal effects are displayed for the two models. Besides this, the values of ρ and the related standard errors and significance are displayed for the bivariate model.

Table 4 Summary of effects voluntary deductible choice on lifestyle variables

Variable	*Model applied*					*Hypothesis*
	I. Univariate probit		*II. Bivariate instrumental probit*			
	Effect	*Marg. effect*	*Effect*	*Marg. effect*	ρ	
Obese	-0.068	Non identified	-1.215***	-29.7%	0.661***	Accept
	(0.062)		(0.240)		(0.157)	
Inactive	0.031	Non identified	-0.230	Non identified	0.178	Reject
	(0.040)		(0.366)		(0.206)	
Heavy smoker	-0.013	Non identified	-1.041***	-12.16%	0.635**	Accept
	(0.099)		(0.307)		(0.185)	
Heavy drinker	0.064	Non identified	-1.141***	-13.25%	0.727**	Accept
	(0.085)		(0.333)		(0.217)	

Notes: This table summarizes the main findings concerning the key variables. The full tables of the applied models can be found in Appendix B. Coefficients for the two probit models per lifestyle variable are displayed within this table under the *effect* columns without parentheses, the significance and sign of these values but not the magnitude can be interpreted. The stars indicate significance levels of either * $p<0.1$; ** $p<0.05$; *** $p<0.01$. Cluster-robust standard errors can be found in parentheses displayed under the coefficients. The magnitude of the *ex ante* moral hazard effect in percentage points, if observed any, can be found under the *marginal effect* columns. Here, the marginal effects can be interpreted as a certain percentage lower chance for people with a vol. deductible than the reference group of people without a vol. deductible to be unhealthy when considering a certain lifestyle variable. Besides this, the Rho value and significance of the bivariate instrumental model are displayed under ρ. The significance levels indicated by stars are similar in definition to the significance levels in the *effects* tables.

The univariate probit model

Starting off, I discuss the first model that builds on the assumption of exogeneity in the voluntary deductible choice after controlling for the selection sources. In *table 4* and *table 7* the (key) results of the univariate probit model on the relationship between the uptake of a positive voluntary deductible and the four lifestyle variables can be found. As can be derived from the *effects* column under the first model displayed in *table 4*, the univariate probit model identifies no significant relationship between the lifestyle variables and the uptake of a voluntary deductible. Therefore, no significant marginal effect can be derived, implying that all four sub-hypotheses should be rejected. According to this model building on an assumption of exogeneity, there does not exist *ex ante* lifestyle moral hazard on the Dutch health insurance market.

The instrumental bivariate probit model

Within the second model, the assumption involving exogeneity in voluntary deductible choice after controlling for various selection sources is dropped. The results of the bivariate

probit model I apply to the relationship between the lifestyle choices and the uptake of a voluntary deductible can be found in *table 8*. A summary of the key results concerning whether or not and to what extent *ex ante* lifestyle moral hazard is identified can be found in *table 4*.

As displayed in *table 8* the instrumental variables complementary insurance and over pessimism are both significantly negatively related to the voluntary deductible choice. The choice for complementary health insurance is accompanied with around a ten percentage point lower chance of taking up a voluntary deductible. This result is significant at the one percent level. Similarly, being over pessimistic decreases your chance of taking up a voluntary deductible with around 15 percentage points. This result is significant at the ten percent level.

When considering the column labeled effect under the second model displayed in *table 4*, it can be observed that this model provides highly significant results on the lifestyle variables being obese, being a heavy smoker and being a heavy drinker. The coefficients on the relation between voluntary deductible uptake and these three lifestyle variables are significant at the 1% level. The sign but not the magnitude of these coefficients can be interpreted. In the second column under the bivariate probit model in *table 4*, the marginal effects identified within this model are displayed. The interpretation is as follows. On average, an individual with a positive voluntary deductible has a 29.7 percentage point lower chance of being obese than an individual from the reference group that does not hold a voluntary deductible. This is a surprisingly large percentage. Similarly, an individual that holds a voluntary deductible on average is 12.16 percentage points less likely to be a heavy smoker than an individual that does not have a positive voluntary deductible. Finally, on average an individual that holds on to a voluntary deductible is 13.25 percentage points less likely to be a heavy drinker than an individual that does not hold on to a voluntary deductible. Only for the variable concerning inactivity no marginal effect can be identified, as the coefficient of the regression between this variable and the voluntary deductible uptake is insignificant. Summarizing, according to this second model the sub-hypotheses concerning obesity, heavy smoking and heavy drinking should be accepted. Leading me to confirm that when applying this model indeed *ex ante* lifestyle moral hazard can be identified within my dataset.

Comparing the models, interpretation of ρ

To be able to draw conclusions from my analysis, I check whether voluntary deductible choice is exogenous to lifestyle variables after controlling for selection sources. To test this, the estimated ρ parameter should be interpreted. If this value is significantly different from zero, the voluntary deductible choice is endogenous to the lifestyle variable considered even after controlling for selection sources. If this is the case, the bivariate instrumental probit model should indeed be used to draw conclusions.

The ρ of lifestyle variables being obese, being a heavy smoker and being a heavy drinker is non-zero and significant either at the 1% or 5% level. Meaning that indeed the error terms when estimating the relation between the uptake of a voluntary deductible and the choice for these lifestyle variables are correlated. There exists endogeneity from omitted variables even after correcting for lagged health state, risk preferences and socio-demographic background (the selection sources). It therefore makes sense to estimate and interpret the instrumental bivariate probit model for these three lifestyle variables. The fourth lifestyle variable considered, being inactive, yields no significantly different from zero value of ρ. This implies that in this case there is no reason to estimate an instrumental bivariate probit model, as the uptake of voluntary deductible can be treated as exogenous after controlling for the three selection sources. To conclude, for lifestyle variables being obese, being a heavy smoker and being a heavy drinker, the instrumental bivariate probit model should be used when drawing conclusions, whereas for lifestyle variable being inactive the univariate probit model should be sufficient.

Concluding remarks per sub-hypothesis

Using the instrumental bivariate probit model, I accept the sub-hypotheses concerning obesity, heavy smoking and heavy drinking. The chance of behaving unhealthy on each of these lifestyle variables indeed decreases with the uptake of a voluntary deductible. Respectively, the chance of behaving unhealthy on these lifestyle traits is 29.7, 12.16 and 13.25 percentage points less likely when a positive voluntary deductible value is held compared to the reference group that does not hold a voluntary deductible. Interpretation of the univariate probit model leads me to reject the sub-hypothesis concerning physical inactivity. No marginal effect of voluntary deductible uptake on physical inactivity is identified within my research.

The acceptance of my first, third and fourth sub-hypotheses enables me to confirm that there indeed exists *ex ante* lifestyle moral hazard on the Dutch health insurance market. In

other words, I accept my main hypothesis. The voluntary deductible works as an incentive to decrease unhealthy lifestyle behaviour concerning being obese, being a heavy smoker and being a heavy drinker. These lifestyle variables are among the main causes of the development of NCD's. The findings of my research are therefore extremely interesting, not only for economic theory concerning *ex ante* lifestyle moral hazard, but also for policy makers.

VI.II.II. Relating sources of selection to lifestyle and voluntary deductible choice

When considering the effect of the sources of selection I control for, some patterns emerge. After briefly discussing some socio-demographic variables, I interpret the influence of both *risk type* and *risk preferences* on lifestyle choices and voluntary deductible choice. The parameter estimates on all control variables can be found within Appendix B, respectively in *table 7* and *table 8*. Similarly to the interpretation of my key results, for variables voluntary deductible uptake, obesity, heavy smoking and heavy drinking I interpret the instrumental bivariate probit model in *table 8*. For the regression concerning lifestyle variable physical inactivity, I interpret the univariate probit model in *table 7*.

Considering socio-demographic variables, the patterns that emerge are mostly in line with expectations and previous research. Matching previous research (Pronk et al., 2004), I find that higher education is on average related with a lower likeliness of adopting an unhealthy lifestyle. Besides increasing with educational level, on average the likeliness to opt for a voluntary deductible also increases with log netto household income. Implying that high-income households fear the possibility of a financial hit of high health costs less than low-income households do.

When considering lagged health state, it is unsurprising that good to excellent self assessed health is related to on average a 3.51 percentage point higher likeliness to opt for a voluntary deductible than the reference group that does not assess their health well. At the same time on average the good assessment of ones health is related to a lower chance of being a heavy smoker and associated with a lower chance to be inactive than the reference group. Suffering from emotional problems is associated with a higher chance of being inactive. A remark that I want to make on these first two variables is that it is quite tricky, as self-assessed health is a subjective matter. Perhaps people aged over 60 or 65 positively assess their health relatively more easily than individuals of younger age. For this reason, it is of main importance to also incorporate the more objective lagged health state variables. These

objective variables result in no unexpected results. It is noteworthy that individuals suffering from a chronic condition on average have a lower likeliness of opting for a voluntary deductible, and also have a lower propensity of adopting two of the unhealthy lifestyle variables. Individuals that suffer from a chronic condition might be more aware of the urge to live healthy to avoid their complaints to worsen. Besides this, these people might have more contact with physical specialists advising them on the importance of a healthy lifestyle. As discussed within the literature section, this type of relation is sometimes referred to as the indirect effect of health insurance (Dave and Kaestner, 2009). When suffering a condition, health costs to be made are usually expected to be higher than when being completely healthy, therefore decreasing the motivation to take up a voluntary deductible.

No significant relation is found between being risk averse and the propensity to lead an unhealthy lifestyle or to opt for a voluntary deductible. This suggests that the risk aversion measure does not capture all risk aversion, which is not unlikely considering previous findings of Kapteyn and Teppa (2011) as discussed within the data section. Unfortunately I have to conclude that the control variable capturing risk preferences is not sufficient within this research. When considering the results, I believe that part of the risk preferences are captured by the gender dummy and the self-employment dummy used in the analysis. To illustrate this, I will elaborate on the parameter estimates of the gender control variable. On average, being male increases one's chance of taking on a positive voluntary deductible with 2.69 percentage points, being a heavy smoker with 4.48 percentage points and being a heavy drinker with 8.14 percentage points compared to when being female. Besides this, it is associated with an increased chance of being inactive of 5.37 percentage points compared to the reference group of females. I believe that these parameter estimates can be (partly) explained by an – on average – lower level of risk aversion amongst males. This effect of risk aversion on individual lifestyle through gender can both be theoretically underbuilt (Barsky et al., 1997) and is empirically identified (Spigner et al., 1993). Similarly, I expect the group of self-employed individuals on average to be less risk averse than the group of individuals with a regular employment contract. Furthermore Kapteyn and Teppa (2011) find that risk aversion is positively related with age. This is a possible explanation for the small but negative marginal effect of age on unhealthy lifestyle uptake that I identify, contradictory to results of Dave and Kaestner (2009).

VII. Discussion

Before closing off with the final conclusions that can be derived from this study, I want to reflect on the weaker points of the research conducted. Both by mentioning limitations and by using these limitations to make suggestions for future research concerning the areas of lifestyle behaviour and *ex ante* moral hazard in health insurance. The critical points I pay attention to, revolve around limitations of data that I use and consequently on some of the assumptions that I make during the analysis. From my point of view, the methodology of the research I conduct is strong and the dataset I construct is rich. Therefore I hope that this study can function as a foundation to build future improving and extending research on.

Limitations and points of improvement

Starting my discussion with a consideration of the dataset used, first I shortly want to mention the general shortcoming of using survey data. The reliability of the results I produce depends not only on the accurateness of the analysis conducted, but also on the willingness and ability of survey respondents to answer accurately. Besides this, a survey panel can suffer from selection bias. As the LISS panel is a carefully assembled panel consisting of a supposedly representative sample of Dutch society over the years that have past, this is less of an immediate concern.

Having mentioned this, to assure the data I use is representative for Dutch society, I verified the summarizing statistics of my dataset with general Dutch averages. This leads me to several points of weakness in the data used. The main point of attention is the extremely high proportion of individuals that are physically inactive. The percentage of around 35% inactive people that I find in every year investigated, is almost three times as high as the Dutch average confirmed by the CBS. This is a strong indication that there exists a bias in my dataset on this point of consideration. A first possible explanation for this bias is the inaccurate answering of respondents on this particular matter. This can be caused either by unclear formulation of the questions concerning physical activity, or a lack of ability or motivation of respondents to answer accurately to questions on this topic. A second possible explanation for the unrealistic results on inactivity is the existence of a selection bias. Implying that the LISS panel at this point does not consist of a sample representative of Dutch society. For example there might exist a survey bias, in this case meaning that the type of individual that is more likely to take the time to answer a survey might be less likely to be

physically active. Inconsequently of the cause, existence of this type of bias leads to unrepresentative results in my analysis of the relation between the voluntary deductible choice and the propensity to be inactive. As discussed in the introductory section, inactivity is in fact one of the main risk factors when it comes to the development of NCD's. For this reason I emphasize that the investigation of impact from financial incentives on this lifestyle factor is very relevant, making this an interesting topic for future research covering a different dataset.

A second shortcoming of the dataset I use concerns the construction of control variable *risk aversion* and instrumental variable *financial literacy*. Both variables are based on survey data from a single year, respectively 2009 and 2011. Consequently, I have included the assumption of constant risk aversion and constant financial literacy in my analysis. The strong assumption of constant risk aversion has been adopted in previous research as well (Alessie et al., 2018). The fact that risk aversion does not relate significantly with any of the lifestyle factors investigated, raises questions. Since I do control for a rich range of socio-demographic variables, any time varying effects concerning risk aversion preferences should be captured by for example the age control variables. The number of respondents on the risk preferences survey is not nearly as large as the total number of respondents covered by my research. This makes it impossible to control for risk aversion across the complete panel investigated. A second disadvantage of the risk aversion measure I use is that it is constructed based on lottery questions. Kapteyn and Teppa (2011) emphasize that the use of a lottery based risk aversion control variable is not ideal to cover risk aversion. Factor-analysis of intuitive statements concerning risk taking (that can be answered with either agree or disagree) usually provide measures for risk aversion with the most explanatory power (Kapteyn and Teppa, 2011). Due to data limitations I was forced to make use of a constant, lottery based measure. Concluding, the measure of risk aversion used in this study is not optimal. An improvement would be to develop a non-constant measure that does not build on lottery questions. Similarly, I do not posses the necessary information on financial literacy for all respondents covered by my analysis. I suspect that this is the reason that financial literacy does not relate significantly to voluntary deductible choice in my analysis. Besides this, I doubt the consistency of financial literacy over time. On the one hand I do find it unlikely that financial literacy level decreases over time, on the other hand it might be somewhat likely that financial literacy increases over time. To conclude, due to data limitations I have chosen to exclude the constant measure of financial literacy from the instrumental variables I use. This measure lacks relevance to suffice as instrumental variable for voluntary deductible choice. However, I expect that a more sophisticated measure of financial literacy is appropriate to instrument for

voluntary deductible choice. With an improved measure of financial literacy that covers a sufficiently large proportion of the research sample, this can suffice as an appropriate instrumental variable to research *ex ante* lifestyle moral hazard.

Extending the research in this area

After having discussed the limitations of my research and related points of improvement for future research, I will now discuss possible extensions of research in this area. The relationship between lifestyle and financial incentives remains extremely interesting and a lot still is to find out. There are two important extensions to this research that I want to discuss.

In this study, only the option to choose a zero or non-zero voluntary deductible is taken into consideration. Though in reality there are multiple levels of voluntary deductible available. Only since recently an increasingly substantial amount of individuals opts for a maximum voluntary deductible. Therefore, I have not extensively researched the choice for a maximum voluntary deductible. Being afraid of finding biased results from researching a too small sub-sample over too little time. If the observed trend of an increasing proportion of Dutch citizens opting for a maximum voluntary deductible perseveres in the years to come, it is interesting to research the relation between opting for different levels of voluntary deductible and unhealthy lifestyle variables. This way, insights in the effect of financial incentives on lifestyle behaviour can be deepened. This would open up the possibility to formulate more detailed policy recommendations.

Besides this, a valuable extension would be to include the lifestyle variable *unhealthy food consumption.* Due to data availability limitations I have no insight in the intake of saturated fats, salts or sugar on the individual level. As discussed in the introductory section, these food consumption variables are highly related to the development of NCD's. Making this an interesting extension area of research. Especially now that topics such as introduction of a sugar tax – as introduced on soda in Belgium, Hungary, the UK and France (FD, 2016) – are emerging in political circles. More insight on the relation between unhealthy food consumption and financial incentives is more than welcome in the current debate.

VIII. Conclusion

Policy recommendations

Starting of by briefly sketching the current Dutch situation evolving around the health sector and its finances, I want to mention the following. The total budget necessary to finance the health costs increase with 7% in 2019 (FD, 2018). Main reasons are the extension of the basic health insurance package, coping with employment shortages in the sector and the obligatory deductible that stays frozen at 385 euros. The pressure on the system further increases, and so do the premiums Dutch individuals pay in 2019. The research I conduct, results in three main findings that policy makers and health insurers ought to take into consideration when developing policy and strategy. The voluntary deductible indeed works as an incentive to decrease the individual likeliness of adopting unhealthy lifestyle habits in terms of being obese, being a heavy smoker and being a heavy drinker. These findings provide policy makers and health insurers with a huge knowledge gain. Providing new possibilities of working towards a tailor-made plan for a future with a minimum amount of new NCD cases and simultaneously a minimum amount of healthcare costs growth.

Taking the perspective of health insurers, it is rational to further increase advertisement for a positive level of voluntary deductible. Extra focus should be on making the voluntary deductible a more understandable concept. This way, a larger proportion of Dutch individuals can incorporate the decision on whether a voluntary deductible would suit them rather than just going with the default option of a zero voluntary deductible. Besides these advertising efforts, health insurers might even consider to offer a larger discount on contracts for individuals opting for a voluntary deductible than currently is happening. A combination of these measures makes opting for a voluntary deductible relatively and observably more beneficial to the Dutch citizen. This way, health insurers can stimulate perseverance of the current trend involving the increasing popularity of the voluntary deductible. Even the consideration of opting for a voluntary deductible might increase consumer awareness of their individual health situation and lifestyle. Actually opting for such a deductible increases consumer financial responsibility, thus working as an incentive to decrease unhealthy lifestyle behaviour.

When considering the current situation from a political perspective, the results from my research lead me to seriously urge for the consideration of an income bound further increase in the obligatory deductible. This will result in an immediate decrease in pressure on

the health budget, both for insurers and the government. Besides this short-term advantage, this measure further strengthens the financial incentive presented to Dutch citizens in order to encourage a healthy lifestyle. Individual propensity to being obese will decrease significantly and so will the individual propensity to heavy smoking behaviour and heavy alcohol consumption. The past years there has been debate on the large differences in minimum and maximum deductible that people are allowed to choose, and on the increasing premiums that are paid on the obligatory deductible (FD, 2016). I argue that if the obligatory deductible increases with a certain amount for the people that can bear the additional financial risk, whilst decreasing the options for voluntary deductible with the exact same amount for this group, the premium asked for all individuals should not have to increase by the amount that is established for 2019. Thus I propose to increase the obligatory deductible for higher income households without offering additional premium discounts to these individuals in specific, whilst simultaneously freezing the maximum *total* level of deductible for all individuals at 885 euros. Under this policy, all individuals are allowed to pick a level of voluntary deductible, but the total deductible may not exceed 885 euros. This way, increasing the Dutch obligatory deductible with an income bound amount will both work as an additional incentive to encourage healthy lifestyle adoption and as a measure to decrease inequality, facilitating not only a transfer of money from the healthy to the sick (general principle of insurance) but also from the wealthy to the poor.

Concluding remarks

This research focuses on lifestyle variables that are highly correlated with the development of NCD's. It concerns being obese; being inactive; being a heavy smoker and being a heavy drinker. Using different models to estimate the relation of these lifestyle variables with the choice for either a zero or non-zero voluntary deductible leads me to the following conclusions. Even after controlling for the selection sources brought forward by previous literature and empirical evidence, there exists endogeneity when estimating the relation between three out of four researched lifestyle variables and voluntary deductible choice. It therefore is sensible to make use of an instrumental bivariate probit model for the lifestyle variables concerning obesity, heavy smoking behaviour and heavy drinking behaviour. To capture the causal relation between the voluntary deductible choice and each lifestyle variable, I apply the bivariate probit model with instrumental variables. The instrumental variables *complementary insurance* and *extreme pessimism* are significantly negatively correlated with the choice for voluntary deductible. The third instrumental variable

financial literacy lacks relevance; therefore I do not use this variable in the analysis. Applying the bivariate model results in a causal negative relationship between the voluntary deductible choice and lifestyle variables being obese, being a heavy smoker and being a heavy drinker. This leads me to accept sub hypotheses 1, 3 and 4. The likeliness of adopting an unhealthy lifestyle on these variables negatively depends on the uptake of a voluntary deductible. Making a deductible exceeding the established obligatory amount an appropriate incentive to encourage a healthier lifestyle. I reject sub-hypothesis 2. No relationship between the choice for a voluntary deductible and the propensity to be inactive can be identified. I expect this to be a consequence of bias in my dataset, believing that the proportion of inactive individuals in my sample is unrepresentatively large for Dutch society.

Finally, I accept my main hypothesis. There exists lifestyle *ex ante* moral hazard on the Dutch health insurance market. The voluntary deductible works as an incentive to decrease unhealthy lifestyle behaviour. This means that a higher level of deductible can suffice as an incentive for Dutch citizens to live healthier and decrease at least three out of four of the main researched risk factors of the development of NCD's. The identification of a lifestyle *ex ante* moral hazard effect is in agreement with findings of Stanciole (2008) and Dave and Kaestner (2009), and in contradiction to findings of Courbage and de Coulon (2004). In line with Stanciole (2008) I confirm that indeed it is necessary to allow for endogeneity from multiple unobservable sources of selection when identifying the *ex ante* moral hazard effect. Making the research method applied within my research appropriate to identify *ex ante* lifestyle moral hazard. Similarly to Stanciole (2008) I identify *ex ante* moral hazard on the variables concerning obesity and heavy smoking, contradictory I also confirm existence of *ex ante* moral hazard on lifestyle variable heavy drinking whilst being unable to confirm existence of *ex ante* moral hazard when considering physical inactivity. This research extends and confirms findings of Stanciole (2008) on the American health insurance market by performing similar analysis on the Dutch market for health insurance, including more reliable instruments and covering a larger research period.

IX. References

Alessie, R., Angelini, V., Mierau, J. O., & Viluma, L. (2018). Moral hazard and selection for voluntary deductibles. *Working paper, University of Groningen.*

Arrow, K. J. (1959). Toward a theory of price adjustment. *The allocation of economic resources*, 41-51.

Barsky, R. B., Juster, F. T., Kimball, M. S., & Shapiro, M. D. (1997). Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study. *The Quarterly Journal of Economics, 112*(2), 537-579.

Beaglehole, R., Bonita, R., Horton, R., Adams, C., Alleyne, G., Asaria P., & Cecchini, M., (2011). Priority actions for the non-communicable disease crisis. *The Lancet, 377*(9775), 1438-1447.

Bloom, D. E., Cafiero, E., Jané-Llopis, E., Abrahams-Gessel, S., Bloom, L. R., Fathima, S., ... & O'Farrell, D. (2012). *The global economic burden of noncommunicable diseases* (No. 8712). Program on the Global Demography of Aging.

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

Centraal Bureau voor de Statistiek (CBS). (2017). *Gezondheid, leefstijl, zorggebruik en –aanbod, doodsoorzaken: kerncijfers.* Den Haag: author unknown.

Centraal Bureau voor de Statistiek (CBS). (2017). *Ruim 700 euro eigen betalingen aan zorg in 2015.* Den Haag: author unkown.

Cherkin, D. C., Grothaus, L., & Wagner, E. H. (1990). The effect of office visit copayments on preventive care services in an HMO. *Inquiry*, 24-38.

Courbage, C., & De Coulon, A. (2004). Prevention and private health insurance in the UK. *The Geneva Papers on Risk and Insurance-Issues and Practice*, *29*(4), 719-727.

Dave, D., & Kaestner, R. (2009). Health insurance and ex ante moral hazard: evidence from Medicare. *International journal of health care finance and economics*, *9*(4), 367.

Ehrlich, I., & Becker, G. S. (1972). Market insurance, self-insurance, and self-protection. *Journal of political Economy*, *80*(4), 623-648.

Finkelstein, A., & McGarry, K. (2006). Multiple dimensions of private information: evidence from the long-term care insurance market. *American Economic Review*, *96*(4), 938-958.

Geneau R., Stuckler, D., Stachenko, S., McKee, M., Ebrahim, S., Basu, S., & Beaglehole, R. (2010). Raising the priority of preventing chronic diseases: a political process. *The Lancet*, *376*(9753), 1689-1698.

Gorter, J., & Schilp, P. (2012). Risk preferences over small stakes: Evidence from deductible choice.

Hemenway, D. (1990). Propitious selection. *The Quarterly Journal of Economics*, *105*(4), 1063-1069.

Hoeks, L. (Red.), (2016). Britse suikertaks dankzij tv-kok Jamie Oliver. Amsterdam: Financieel Dagblad (FD).

Kapteyn, A., & Teppa, F. (2011). Subjective measures of risk aversion, fixed costs, and portfolio choice. *Journal of Economic Psychology*, *32*(4), 564-580.

Kenkel, D. S. (1994). The demand for preventive medical care. *Applied Economics*, *26*(4), 313-325.

Kenkel, D. S. (2000). Prevention. *In Culyer, A. J., & Newhouse, J. P. (Eds.), Handbook of health economics. Elsevier*. *1B*(31). Amsterdam: North Holland.

Leu, R. E., Rutten, F. F., Brouwer, W., Matter, P., & Rütschi, C. (2009). The Swiss and Dutch health insurance systems: universal coverage and regulated competitive insurance markets. *The Commonwealth Fund*, *104*, 1-40.

Light, D. W. (1992). The practice and ethics of risk-rated health insurance. *Jama*, *267*(18), 2503-2508.

Loewenstein, G., Brennan, T., & Volpp, K. G. (2007). Asymmetric paternalism to improve health behaviours. *Jama*, *298*(20), 2415-2417.

Manning, W. G., & Marquis, M. S. (1996). Health insurance: the tradeoff between risk pooling and moral hazard. *Journal of health economics, 15*(5), 609-639.

Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. *Commons, ML.; Mazur, JE.; Nevin, JA*, 55-73.

Morse, R. M., & Flavin, D. K. (1992). The definition of alcoholism. *Jama*, *268*(8), 1012-1014.

Nederlandse Zorgautoriteit (NZa). (2017). *Marktscan Zorgverzekeringsmarkt 2017*, Utrecht: Nederlandse Zorgautoriteit.

Newhouse, J. P. (1993). Free for all? Lessons from the RAND health insurance experiment. In *A RAND study.* Cambridge, London: Harvard University Press.

Pauly, M. V. (1968). The economics of moral hazard: comment. *The American Economic Review,* 531-537.

Pronk, N. P., Anderson, L. H., Crain, A. L., Martinson, B. C., O'Connor, P. J., Sherwood, N. E., & Whitebird, R. R. (2004). Meeting recommendations for multiple healthy lifestyle factors: prevalence, clustering, and predictors among adolescent, adult, and senior health plan members. *American journal of preventive medicine*, *27*(2), 25-33.

Rothschild, M., & Stiglitz, J. (1976). Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information. *The Quarterly Journal of Economics, 90*(4), 629-649.

Shavell, S., (1979). On moral hazard and insurance. *The Quarterly journal of economics, 93*(4), 541-562.

Spigner, C., Hawkins, W. E., & Loren, W. (1993). Gender differences in perception of risk associated with alcohol and drug use among college students. *Women & health*, *20*(1), 87-97.

Stanciole, A. E. (2008). Health insurance and lifestyle choices: Identifying ex ante moral hazard in the US market. *The Geneva Papers on Risk and Insurance-Issues and Practice*, *33*(4), 627-644.

Stock, J. H., & Watson, M. W. (2012). *Introduction to econometrics: Global edition*. Boston, MA: Pearson Education.

Ten Kate, M. (Red.), (2016). Hoe solidair is het Nederlandse zorgstelsel nog? Amsterdam: Financieel Dagblad (FD).

Van Dale middelgroot woordenboek Engels Nederlands. (2009). Utrecht: Van Dale.

Van der Leij, L. (Red.), (2018). Zorgkosten stijgen volgend jaar met €5 mrd. Amsterdam: Financieel Dagblad (FD).

Van Rooij, M., Lusardi, A., & Alessie, R. (2011). Financial literacy and stock market participation. *Journal of Financial Economics*, *101*(2), 449-472.

Visser, M. (Red.). (2017). Wie gezond is neemt vaak het hoogst mogelijke eigen risico. Amsterdam: Trouw, de Persgroep Nederland.

Walsh, R. (2011). Lifestyle and mental health. *American Psychologist*, *66*(7), 579.

World Health Organization (WHO). (2003). *Diet, nutrition, and the prevention of chronic diseases: report of a joint WHO/FAO expert consultation* (Vol. 916). World Health Organization.

World Health Organization (WHO), (2005). Preventing chronic diseases: a vital investment: WHO global report. In *Preventing chronic diseases: a vital investment: WHO global report.*

X. Appendices

Appendix A. Descriptive and summarizing statistics: CBS

Table 5 Development obligatory Dutch deductible

Year	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
Deductible	155	165	170	220	350	360	376	385	385	385	385

Note: The levels of deductible are displayed in euros. The deductible under year 2019 is scheduled for the coming year.

Table 6 Proportion of unhealthy lifestyle occurrence per year, in % of total population (CBS)

Year	*Category*			
	Obese	*Inactive*	*Heavy smoker*	*Heavy drinker*
2009	12.1			10.4
2010	11.4			9.4
2011	11.4			9.4
2012	12.0			12.9
2013	11.8			12.4
2014	13.6	12.6	4.1	8.7
2015	13.7	12.4	3.8	9.5
2016	14.5	13.1	3.7	8.1
2017	14.2	12.1	3.3	8.6
Population	>12 years	>12 years	>12 years	>20 years

This table consists of data from CBS (*Centraal Bureau Statistiek)* reported lastly in 2018, Statline. In this table, the proportion of the Dutch population that adopts unhealthy lifestyle variable(s) can be overviewed per year.

Appendix B. Analysis

Table 7 Univariate probit regressions

Explanatory variables	*Dependent variables*			
	Obese	*Inactive*	*Heavy smoker*	*Heavy drinker*
Has vol. ded.	-0.068	0.031	-0.013	0.064
	(0.062)	(0.040)	(0.099)	(0.085)
	[-0.0145]	[0.0109]	[-0.0012]	[0.0048]
Age	0.025*	-0.002	0.107***	0.095***
	(0.014)	(0.008)	(0.026)	(0.018)
	[-0.0019]	[-0.0016]	[-0.0012]	[-0.0003]
Age^2	-0.000**	-0.000	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Male	-0.123*	0.151***	0.232**	0.871***
	(0.067)	(0.035)	(0.099)	(0.082)
	[-0.0263]	[0.0537]	[0.0216]	[0.0559]
Kids	0.019	0.034	-0.268**	-0.025
	(0.075)	(0.042)	(0.119)	(0.083)
	[0.0040]	[0.0120]	[-0.0234]	[-0.0018]
Married	-0.007	0.017	-0.425***	-0.303***
	(0.073)	(0.041)	(0.102)	(0.084)
	[-0.0014]	[0.0060]	[-0.0435]	[-0.0246]
Education low	0.143*	0.181***	0.176	0.010
	(0.078)	(0.043)	(0.118)	(0.097)
	[0.0315]	[0.0658]	[0.0172]	[-0.0007]
Education high	-0.218***	-0.214***	-0.181	0.015
	(0.081)	(0.043)	(0.121)	(0.085)
	[-0.0453]	[-0.0755]	[-0.0163]	[0.0012]
Student	-0.247	-0.164	-0.145	
	(0.350)	(0.175)	(0.474)	
	[-0.0466]	[-0.0560]	[-0.0122]	
Self employment	0.146	0.134*	0.289*	0.316**
	(0.140)	(0.070)	(0.169)	(0.136)
	[0.0333]	[0.0489]	[0.0325]	[0.0293]
Unemployed	0.114	0.013	-0.054	-0.066
	(0.096)	(0.056)	(0.125)	(0.125)
	[0.0256]	[0.0047]	[-0.0049]	[-0.0048]
Retired	0.092	0.077	0.013	0.180
	(0.095)	(0.054)	(0.129)	(0.133)
	[0.0200]	[0.0274]	[-0.0012]	[0.0145]
Log income	-0.119*	-0.049	-0.317***	0.176**
	(0.067)	(0.039)	(0.089)	(0.085)

	[-0.0254]	[-0.0174]	[-0.0297]	[0.0132]
Countryside	-0.187**	0.147***	0.242**	0.118
	(0.087)	(0.047)	(0.122)	(0.098)
	[-0.0375]	[0.0533]	[0.0255]	[0.0094]
Good self assessed	-0.185***	-0.110**	-0.393***	-0.250***
health	(0.070)	(0.044)	(0.097)	(0.088)
	[-0.0421]	[-0.0397]	[-0.0443]	[-0.0214]
Emotional problems	-0.108	0.094*	0.088	0.194**
	(0.075)	(0.048)	(0.093)	(0.089)
	[-0.0221]	[0.0341]	[0.0086]	[0.0164]
Chronic condition	0.090	-0.072*	-0.258**	-0.125
	(0.067)	(0.039)	(0.104)	(0.080)
	[0.0196]	[-0.0256]	[-0.0228]	[-0.0091]
Diabetes	0.558***	0.008	-0.127	-0.261*
	(0.109)	(0.068)	(0.197)	(0.145)
	[0.1495]	[0.0030]	[-0.0109]	[-0.0163]
Cancer	0.200	0.012	0.187	0.081
	(0.155)	(0.105)	(0.247)	(0.223)
	[0.0468]	[0.0043]	[0.0200]	[0.0065]
Lung disease	0.094	0.089	-0.132	0.134
	(0.139)	(0.093)	(0.214)	(0.169)
	[0.0211]	[0.0323]	[-0.0113]	[0.0110]
Heart problems	0.438***	-0.044	-0.021	0.054
	(0.073)	(0.045)	(0.116)	(0.084)
	[0.1067]	[-0.0154]	[-0.0019]	[0.0042]
Risk averse	-0.010	-0.031	-0.119	0.070
	(0.073)	(0.039)	(0.104)	(0.081)
	[-0.0022]	[-0.0109]	[-0.0116]	[0.0052]
Constant	-0.600	0.298	-0.627	-4.924***
	(0.632)	(0.371)	(0.938)	(0.806)
Pseudo R^2	0.075	0.019	0.121	0.119
N	11,497	11,497	10,030	12,431
N-clusters	2,298	2,298	2,115	2,416

*N*otes: * $p<0.1$; ** $p<0.05$; *** $p<0.01$. Cluster-robust standard errors in parentheses. Marginal effect in percentage point in square parentheses. The relation between age and lifestyle is modeled with help of a quadratic function. I performed the same analysis with inclusion of year dummy variables, this did not affect the results. For this reason I report the results of this model without inclusion of year dummy variables.

Table 8 Bivariate probit regressions with instruments

Explanatory variables	*Dependent variables*				
	Has vol. ded.	*Obese*	*Inactive*	*Heavy smoker*	*Heavy drinker*
Has vol. deductible		-1.215***	-0.230	-1.041***	-1.141***
		(0.240)	(0.366)	(0.307)	(0.333)
		[-0.2970]	[-0.0816]	[-0.1216]	[-0.1325]
Complementary insur.	-0.415***				
	(0.070)				
	[-0.0950]				
Extreme pessimism	-0.647*				
	(0.343)				
	[-0.1482]				
Age	0.017	0.029*	-0.001	0.083***	0.128***
	(0.013)	(0.016)	(0.010)	(0.028)	(0.036)
	[-0.0029]	[-0.0034]	[-0.0023]	[-0.0019]	[-0.0015]
Age^2	-0.000**	-0.000**	-0.000	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Male	0.117*	0.005	0.126***	0.226**	0.750***
	(0.060)	(0.072)	(0.048)	(0.105)	(0.128)
	[0.0269]	[0.0011]	[0.0448]	[0.0262]	[0.0814]
Kids	0.051	0.054	0.032	-0.221*	-0.035
	(0.070)	(0.079)	(0.055)	(0.130)	(0.098)
	[0.0119]	[0.0134]	[0.0114]	[-0.0244]	[-0.0040]
Married	-0.161**	-0.087	-0.006	-0.457***	-0.396***
	(0.066)	(0.078)	(0.054)	(0.109)	(0.092)
	[-0.0379]	[-0.0215]	[-0.0022]	[-0.0587]	[-0.0505]
Education low	0.017	0.141*	0.166***	0.166	0.076
	(0.079)	(0.083)	(0.056)	(0.127)	(0.105)
	[0.0038]	[0.0352]	[0.0597]	[0.0202]	[0.0091]
Education high	0.219***	-0.142	-0.147**	-0.009	0.062
	(0.070)	(0.091)	(0.060)	(0.135)	(0.105)
	[0.0516]	[-0.0341]	[-0.0521]	[-0.0010]	[0.0072]
Student	-0.122	0.238	-0.308	-4.052***	-3.748***
	(0.380)	(0.480)	(0.402)	(0.375)	(0.351)
	[-0.0264]	[0.0633]	[-0.1009]	[-0.0684]	[-0.0691]
Self employment	0.119	0.171	0.198**	0.291	0.328**
	(0.128)	(0.159)	(0.094)	(0.199)	(0.159)
	[0.0285]	[0.0442]	[0.0726]	[0.0401]	[0.0456]
Unemployed	0.045	0.122	0.025	0.050	-0.203
	(0.098)	(0.109)	(0.073)	(0.143)	(0.146)
	[0.0104]	[0.0309]	[0.0088]	[0.0060]	[-0.0212]
Retired	-0.006	0.126	0.062	0.005	0.204
	(0.094)	(0.102)	(0.071)	(0.148)	(0.134)
	[-0.0013]	[0.0314]	[0.0222]	[0.0006]	[0.0250]

Log income	0.125*	-0.036	-0.098*	-0.245**	0.257**
	(0.064)	(0.073)	(0.054)	(0.111)	(0.101)
	[0.0286]	[-0.0088]	[-0.0349]	[-0.0286]	[0.0298]
Countryside	-0.127	-0.287***	0.137**	0.174	0.088
	(0.085)	(0.095)	(0.062)	(0.138)	(0.119)
	[-0.0279]	[-0.0648]	[0.0495]	[0.0220]	[0.0106]
Good self assessed	0.161**	-0.121	-0.100	-0.250*	-0.085
health	(0.080)	(0.085)	(0.062)	(0.132)	(0.109)
	[0.0351]	[-0.0305]	[-0.0359]	[-0.0324]	[-0.0102]
Emotional problems	-0.037	-0.069	0.153**	0.034	0.016
	(0.096)	(0.095)	(0.071)	(0.108)	(0.125)
	[-0.0084]	[-0.0164]	[0.0556]	[0.0041]	[0.0019]
Chronic condition	-0.182***	0.008	-0.047	-0.306**	-0.170*
	(0.066)	(0.075)	(0.054)	(0.122)	(0.096)
	[-0.0407]	[0.0020]	[-0.0166]	[-0.0339]	[-0.0192]
Diabetes	-0.021	0.518***	0.037	0.017	-0.100
	(0.118)	(0.118)	(0.083)	(0.211)	(0.147)
	[-0.0048]	[0.1482]	[0.0131]	[0.0021]	[-0.0110]
Cancer	-0.283	0.087	-0.071	0.114	0.144
	(0.192)	(0.174)	(0.128)	(0.276)	(0.216)
	[-0.0570]	[0.0220]	[-0.0249]	[0.0143]	[0.0182]
Lung disease	0.090	0.071	0.111	-0.187	0.193
	(0.150)	(0.165)	(0.116)	(0.275)	(0.168)
	[0.0214]	[0.0177]	[0.0403]	[-0.0194]	[0.0251]
Heart problems	-0.053	0.368***	-0.003	-0.001	-0.021
	(0.074)	(0.083)	(0.056)	(0.128)	(0.092)
	[-0.0120]	[0.0974]	[-0.0010]	[-0.0001]	[-0.0025]
Risk averse	-0.087	-0.055	-0.002	-0.079	0.079
	(0.067)	(0.077)	(0.052)	(0.110)	(0.094)
	[-0.0203]	[-0.0136]	[-0.0008]	[-0.0094]	[0.0090]
Physiotherapy	-0.024	0.109*	-0.027	-0.096	0.044
	(0.058)	(0.062)	(0.046)	(0.089)	(0.081)
	[-0.0055]	[0.0271]	[-0.0094]	[-0.0109]	[0.0052]
Constant	-1.631***	-0.835	0.693	-0.408	-5.985***
	(0.603)	(0.681)	(0.479)	(0.978)	(1.409)
ρ		0.661***	0.178	0.635**	0.727**
		(0.157)	(0.206)	(0.185)	(0.217)
N	5,205	5,205	5,205	4,611	5,205
N-clusters	1,713	1,713	1,713	1,575	1,713

*N*otes: * $p<0.1$; ** $p<0.05$; *** $p<0.01$. Cluster-robust standard errors of the coefficients in parentheses. Marginal effect in percentage point in square parentheses. The relation between age and lifestyle is modelled with help of a quadratic function. I performed the same analysis with inclusion of year dummy variables, this did not affect the results. For this reason I report the results of this model without inclusion of year dummy variables. [a] These are estimates from the instrumental bivariate probit regression for the uptake of a voluntary deductible and being obese.