Sandra Margarida Bernardes de Oliveira

The Demand for Preventive Health Care and Health Insurance: evidence for Portugal.

Tese de Doutoramento na área científica de Economia, orientada pelos Professores Doutores Pedro Pita Barros e Óscar Lourenço, e apresentada à Faculdade de Economia da Universidade de Coimbra

Março de 2014

· U C ·

Universidade de Coimbra

FEUC FACULDADE DE ECONOMIA
UNIVERSIDADE DE COIMBRA

Sandra Margarida Bernardes de Oliveira

The Demand for Preventive Health Care and Health Insurance: evidence for Portugal

Thesis presented to the Faculty of Economics to the University of Coimbra in fulfilment of the requirement for the degree of Doctor in Economics.

Supervisors: Professor Pedro Pita Barros and Professor Óscar Lourenço

Coimbra, 2014

To Maria, with love

Financial Support

This work was carried out with the support from Fundação para a Ciência e Tecnologia - POPH-QREN (reference number SFRH/PROTEC/50189/2009) and Centro de Estudos e Investigação em Saúde of University of Coimbra (CEISUC).

ACNOWLEDGMENTS

The elaboration of this thesis was only possible due to numerous contributions, both personal and institutional, to which I express my gratitude.

Firstly, I want to thank my parents, for encouraging me to proceed in my academic life and improve my skills, despite the implicit sacrifices for them. To my sister and godmother for being always present in my life and for all the care and love that they devoted to Maria. This work would not be possible without your support. It is impossible not mention my grandfather António a reference to me and the best of all of us. My nephew Santiago: thanks for all the joy and happiness that you brought to my life.

To my PhD supervisors, Professor Pedro Pita Barros and Professor Óscar Lourenço, who had always been available to hear my doubts, to help me in moments of indecision and especially for believing in me in difficult moments. I am deeply grateful for all the pertinent remarks, the rigorous criticisms and valuable suggestions that enriched this research. Without their availability, scientific accuracy and professionalism, I would have certainly not been able to complete this thesis.

To my friends – the ones of a life and those I was fortunate to know at the University and at work – words of thank are not enough. I will be always grateful to Ana Maria, Nélia, Rita, Simone for being with me all the time. To Carla Vivas, Susana Leal, Goreti and Cláudio for sharing with me this process of doubts, questions and moments of nervousness. Susana, thanks for all and invaluable help, constant support and for never giving up on me. Thanks for your friendship. To Micaela Antunes and Ana Poças my dearest colleagues of PHD that always care and hold up for me. Thanks for your rare generosity and friendship.

A special reference to Maria Manuel Vairinho, my friend, colleague and also family, with whom I learned and shared so much about almost everything. I am a lucky person for having, found such an extraordinary person, which makes me different and better.

To all the departments and offices of the Management and Technology School of Santarém, to all who work in these services and the Professors I worked with, for providing me the necessary resources and support to elaborate my thesis.

To Renato for all the time you put me the screws on, which makes me crazy, but results. Thanks for believing and making me believe.

And last, but not the least, to my sun António, MY GIFT.

RESUMO

A procura deliberada por cuidados preventivos de saúde deverá ser uma mudança expectável no comportamento dos indivíduos. Decorrente de uma envolvente contextual desafiante e que obriga a uma eficiente alocação de recursos, conduzindo os indivíduos a fazerem melhor uso da informação, rendimento e tempo.

Ao longo desta tese apresentamos três ensaios distintos com vários elementos em comum: o estudo e análise de comportamentos individuais no domínio dos cuidados preventivos de saúde e uma dimensão de estilos de vida, utilização de dados individuais e recurso a modelos de regressão não lineares.

Os ensaios, mesmo com objetivos diferentes, visam todos contribuir para algo fundamental: a compreensão dos fatores que ajudam a explicar as atitudes individuais relacionadas com comportamentos pró-ativos no domínio da saúde (prevenir), e, por sua vez, como estes esforços de prevenção influenciam algumas decisões individuais.

Iniciamos a investigação (primeiro ensaio) com a análise do papel da informação no comportamento do indivíduo face ao processo de decisão associado à disponibilidade para pagar por cuidados preventivos de saúde. Avaliamos se a quantidade e tipo de informação detida pelo indivíduo influência a disponibilidade para pagar pela vacina da gripe A. Recorremos à especificação de um modelo de duas partes para analisar a disponibilidade para pagar, condicional num conjunto de variáveis de controlo, entre elas, as variáveis de informação. Os resultados mostram que o tipo de informação e o veículo privilegiado de informação utilizado pelos alunos, bem como o *background* dos mesmos, em termos de curso frequentado, influencia a disponibilidade pagar por cuidados de saúde preventivos.

O segundo estudo muda o enfoque e explora a prevenção como elemento determinante de um comportamento, absentismo laboral. A nossa intenção foi a de analisar se o consumo de cuidados preventivos ajuda a explicar na taxa de absentismo. A evidência veiculada por esta análise sugere que um comportamento preventivo não influencia a decisão de faltar ao trabalho. Contudo, esta análise contém algumas limitações, designadamente na qualidade dos dados para medir comportamento preventivo.

Finalmente, regressamos à prevenção como comportamento individual a explicar e analisamos as determinantes da decisão simultânea dos indivíduos de adotar comportamentos preventivos primários e secundários. Daremos especial enfoque ao papel do desemprego neste consumo. O interesse desta investigação reside no facto: 1) o desemprego pode conduzir o individuo a um pior estado de saúde geral; 2) por outro lado, numa situação de desemprego os indivíduos ficam com mais tempo livre para atividades promotoras de saúde, quer prevenção primária quer secundária; 3) se os desempregados usarem o seu tempo nessas atividades promotoras de saúde, poderá ter um efeito positivo no estado de saúde dos indivíduos, que de algum modo pode amenizar o efeito negativo do desemprego nesse mesmo estado de saúde. A principal mensagem a extrair deste estudo é que o desemprego influencia a procura de cuidados de saúde, embora o mesmo não se possa afirmar para a procura simultânea de cuidados preventivos.

Apesar de algumas limitações, principalmente relativas à qualidade dos dados disponíveis, o desenvolvimento desta tese permite concluir que os comportamentos preventivos dos indivíduos são explicados multifactorialmente e que, pelo menos do ponto de vista teórico, também influenciam outras realidades da vida dos indivíduos.

Palavras-chave: Prevenção relacionada com a saúde, disponibilidade para pagar, estilos de vida, absentismo, método de avaliação contingente, modelos de contagem.

ABSTRACT

The deliberate search for preventive health care should be an expected change in the behaviour of individuals. Due to a challenging and engaging context, an efficient allocation of resources is expected, leading individuals to make better use of information, income and time.

Along this thesis we will present three distinct essays, each of them following a different methodological approach. However, they all have a common element: the analysis of the consumption of preventive health care and its relationship with individuals' health behaviour, the use of individual data and nonlinear models.

The essays, even with different goals, were developed to contribute for something essential: an understanding of the factors that help explain individual attitudes toward proactive health behaviour (prevention), and in turn, how these preventing efforts could influence individual decisions.

We started the research (first essay) exploring the impact of information possessed by the individual on his behaviour during a pandemic of H1N1 virus. We assess whether the information and the vehicle of information influence the willingness to pay for H1N1 flu vaccine.

We developed a questionnaire, which we gave to a sample of 4193 university students (from Portugal, Greece and Spain). Its objective was to collect data on the WTP for an influenza shot during a pandemic period, along with other information about the individual. We adopted a Two-part model to analyse the research question. The results show that students tend to react differently depending on the source of information and whether they have a pre-disposition for buying the vaccine. The main medium of information appears to reinforce this decision.

The second essay changes the focus and explores prevention as a crucial element in absenteeism behaviour. Our intention was to examine whether the use of preventive care explain the absenteeism rate. The evidence confined on essay analysis suggests that a preventive behaviour does not influence the decision to miss work. However, this analysis has some limitations, in particular, the quality of data to measure preventive behaviour

Finally, we return to prevention and individual behaviour to explain and analyze the determinants of simultaneous decision of individuals to adopt primary and secondary preventive behaviours. We will give special focus to the role of unemployment in consumption.The interest of this research was: 1) unemployment may lead the individual to a poorer general health status, 2) on the other hand, unemployed individuals have more free time to use preventive health care, either primary or secondary prevention, 3) if the unemployed use their free time in these health-promoting activities, could have a positive effect on the health status of individuals who somehow can mitigate the effect of unemployment. The main result from this study is that unemployment influences the demand for preventive health care, although the same is not true for the simultaneous demand for preventive care.

Despite some limitations, especially regarding the quality of available data, the development of this thesis support the conclusion that preventive behaviours of individuals are explained by multifactors and that, at least from a theoretical point of view, also influences other dimensions of lifestyle.

Keywords: Health related prevention, willingness to pay, lifestyles, absenteeism, contingent valuation method, count data models.

CONTENTS

ACNOWLEDGMENTS ..... **VII**

RESUMO ..... **IX**

ABSTRACT ..... **XI**

CONTENTS ..... **XIII**

LIST OF TABLES ..... **XV**

LIST OF FIGURES ..... **XVII**

CHAPTER 1 – INTRODUCTION **1**

CHAPTER 2 – THE INFLUENCE OF INFORMATION IN THE WILLINGNESS TO PAY FOR INFLUENZA PREVENTION: EVIDENCE FOR PORTUGAL, SPAIN AND GREECE **7**

2.1 INTRODUCTION 7

2.2 LITERATURE REVIEW 13

2.3 EMPIRICAL STRATEGY 16

2.4 DATA AND VARIABLES 21

2.4.1 Measurement Instrument *21*

2.4.2 Sample and Procedures *22*

2.5 RESULTS 32

2.6 MAIN CONCLUSIONS AND DISCUSSION 39

2.6.1 Final Note *41*

CHAPTER 3 – DOES PREVENTIVE CARE EXPLAIN ABSENTEEISM? EVIDENCE FOR PORTUGAL **43**

3.1 INTRODUCTION 43

3.2 LITERATURE REVIEW 46

3.3 THEORETICAL FRAMEWORK 49

3.4 PORTUGUESE SICKNESS BENEFIT SYSTEM 52

3.5 DATA VARIABLES AND METHODOLOGY 53

3.5.1 Dataset *53*

3.5.2 Empirical model *55*

3.6 RESULTS 67

3.7 CONCLUSION, DISCUSSION AND LIMITATIONS 73

CHAPTER 4 -SIMULTANIOUS PREVENTIVE HEALTH BEHAVIOURS AND THE STATE OF EMPL **79**

4.1 INTRODUCTION 79
4.2 CONCEPTUAL FRAMEWORK 82
4.3 EMPIRICAL STRATEGY 84
4.4 DATA AND VARIABLES 86
4.4.1 The dependent variables - Physical Activity and Consumption of preventive health care. 87
4.4.2 Control variables 90
4.5 RESULTS 94
4.6 CONCLUSION AND DISCUSSION 103

CHAPTER 5 – CONCLUSION 107

5.1 GENERAL CONCLUSION 107
5.2 LIMITATIONS AND FUTURE RESEARCH 109
5.3 FINAL NOTE 110

REFERENCES 111

APPENDIXES 127

LIST OF TABLES

Table 2.1 - Summary statistics of WTP by country (Dependent Variable) 25
Table 2.2 - Definition of the control variables 26
Table 2.3 - Summary statistics of the covariates (mean and standard deviation). 31
Table 2.4 – Results of the first part of the TPM 34
Table 2.5 - A two-part model - Second part an Ordered probit and a GlM with a link function and a Gamma distribution for the positives 35
Table 3.1 - Definition of the control variables 59
Table 3.2 - Mean and standard deviation of control variables 65
Table 3.3 - Goodness-of-fit criteria 68
Table 3.4 - Estimation results for ZINB 69
Table 4.1 - Number of individuals that that take part (or not) in physical activity 89
Table 4.2- Number of individuals that take (or not) secondary prevention 90
Table 4.3- Definition of control variables 92
Table 4.6 Bivariate probit parameters for the all sample 95
Table 4.7 - Bivariate probit for Secondary prevention for all sample 97
Table 4.8 Bivariate probit for Physical activity age 18-39 99
Table 4.9 Bivariate probit for Secondary Prevention age 18-39 100
Table 4.10 - Bivariate probit for Physical activity age >40 101
Table 4.11 - Bivariate probit for Secondary prevention age >40 102

LIST OF FIGURES

Figure 1.1 - Outline of the thesis 2
Figure 2.1 - Empirical distribution of WTP 19
Figure 4.1 - Percentage of individuals who engages in physical activity 89
Figure 4.2 - Percentage of individuals that take secondary prevention 90

CHAPTER 1

INTRODUCTION

Over the last century, the type of health problems that developed countries had to face has changed. There are substantially less cases of serious infectious diseases, and many of those that do occur are usually curable. Chronic diseases, diseases related to old age and cancers have become more common and are increasingly the major causes of morbidity and mortality in developed countries. Modern medical technology can contribute to the cure or to the progression of these diseases, but are also much expensive, and is not a solution.

Facing a high demand for health care and a limited budget, most countries need to prioritize health care efforts among their population. Recent research has underlined and put the focus on prevention, mainly because some of these diseases are caused by behavioural or environmental factors (Frank, 2004). In fact, it is well known that some of these factors are preventable. Preventative care is being considered an increasingly important part of the strategy to improve our society's overall health. The factors that encourage its use deserve to be scrutinized more closely. Equally important is to evaluate whether preventive behaviour influence other individuals activities, and if so, the channels through which such influence occurs. In this thesis, we see health prevention behaviour as being influenced by a number of factors, but also influencing individual's decision.

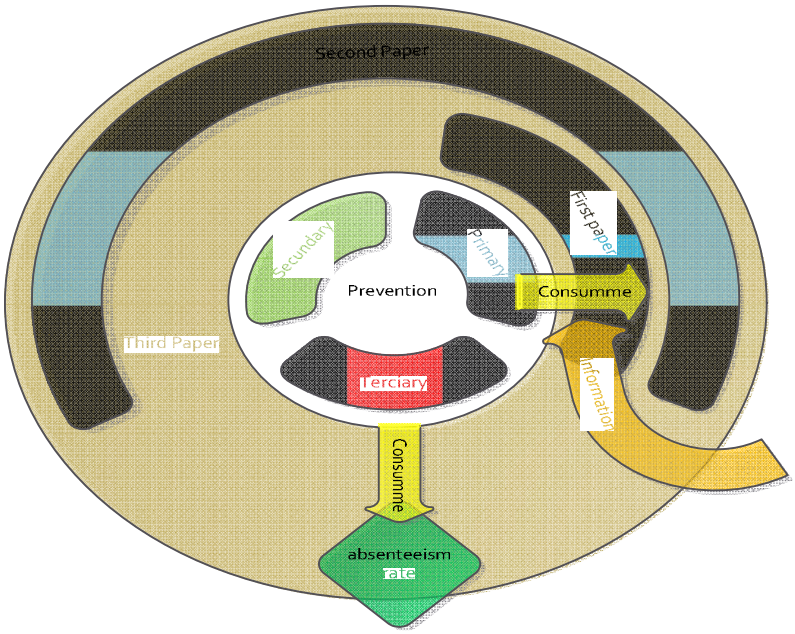
Kenkel (2000) suggest three different categories of health prevention: primary, secondary and tertiary. The author defines primary prevention as actions that reduce the occurrence or incidence of disease; secondary prevention as actions that reduce or eliminate the health consequences of a disease once it has occurred and tertiary prevention as actions that reduce disability associated with a chronic illness.

The thesis is organized into three major chapters autonomous but interrelated – the common strand is the individual's health related preventive behaviour. The schematic shown in figure1 outlines the interrelations among the different essays developed.

We explore the role played by socioeconomic factors, health behaviours and others factors could have in the comprehension of individual attitudes through the intentional demand for preventive health care. For that we based our analysis on well-known theoretical models like those developed by Grosman and Allen. The intentional demand for preventive care could be in our opinion linked to the increasing empowerment of the individual in the final decision of use health care, and particularly preventive health care (this apply to the first and third essay). So, it is important to assure that not only information is given, but how is that done and finally the quality of that information. Again, the intentional demand for preventive care could also be a way of improving overall health status and with that decreasing the missing days due to illness and improve the productivity of individuals.

In the first essay we test the impact that information has on the individuals' decision regarding willingness to pay for a vaccine (primary prevention – university students). The aim of the second essay understands the link, if there is one, between the consumption of preventive (secondary prevention) health care and the illness absenteeism rate of the Portuguese workforce. In the third essay, we search for the determinants of the consumption of primary and secondary prevention.

Figure 1.1 - Outline of the thesis



The first part of our research is devoted to understanding whether or not individuals are willing to pay for preventive care and we focus most of our attention on the H1N1 vaccine. In most cases the influenza-flu shot is free of charge for citizens eligible to be included in the risk group.

At first, the rules for belonging to the H1N1-influenza vaccine risk group were quite similar. However, public alarm and the decision to declare a pandemic made quite a difference. National governments began enlarging the groups of eligibility, which were different from country to country. Its sale was controlled by the Government and it was not possible for the common citizen to buy it.

It was important to find a comprehensive framework capable of explaining the utilization of preventive health care. This makes even more sense considering the eligible groups were different from country to country. This research make possible to compare the behaviour of individuals belonging to different countries. Based on Arrow (1963) and Hibbard and Peters (2003) we find that information is a key issue in the utilization of medical care. It is important not only the provision of information but the process of using data to inform, that is actually quite complex. Therefore, it is clear that information might play a crucial role in the decision making process during a public health crisis. In this specific case, we were facing a global threat which was producing a never before seen stream of information. It was reported with various degrees of accuracy and frequently associated to dramatic scenarios.

Based on various health studies, several researchers claim that the willingness to pay (WTP) is greatly influenced by the degree of information provided by the survey and also by the manner in which the information is presented (Smith, 2008; Pedersen et al., 2011; Neumann et al., 2012; Olsen et al., 2004). Thus, these two points are crucial for our research.

The main aim of this study is the search for potential interdependencies between information held by the individual (searched by him and to which he is exposed to) and his decision process regarding the WTP for preventive health care (PHC).

We analyse data consisting of university students in Portugal, Spain and Greece from different faculties. The universities are potential outbreak centres due to their population, high levels of social contact and, contrary to what happens with seasonal influenza, students represent a potential H1N1 risk group (high levels of morbidity and mortality). Students could

be vehicles of transmission, given that their behaviour could have an impact on the larger communities of which they are a part (family, friends and colleagues).

Since higher education institutions have the potential of becoming serious outbreak centres during a pandemic, the knowledge and attitudes of university students on the subject are important factors for the health authorities to take into account when designing directives. We believe that identifying the main vehicle of information is essential because students can influence those surrounding them. This also allows public authorities to respond to the different groups accordingly. If we discover the main information vehicle, it will be easier to determine the allocation of resources and how to communicate practices and policies more efficiently to the different groups. As a result, educational interventions could lead to increased consumer awareness regarding the cost and/or the probable willingness to pay for health care.

Our second main point of thesis is the study of the relationship between the absenteeism rate due to illness and the consumption of preventive health care. The literature has already identified some of the determinants of illness absenteeism.

Currie and Madrian (1999) suggest that health can affect the ability and desire to work – "the work capacity". The consumption of preventive care may be a way of improving work capacity and does not present a conflict of interests for the two parties involved (employer and employee). Additionally, the individual's welfare is understood as a function of two components: consumption (which requires work) and leisure. Maximization means finding the quantity of work that offers the best possible combination of these components (Allebeck and Mastekaasa, 2004).

Thus, we conclude that health affects work participation directly through productivity, and indirectly by changing the trade-offs between income and leisure (raises health stock, decreases wasted days, increases time for health, home and income production). We can infer that individuals who take preventive actions (like regular check-ups and screening cancer programs) have a lower chance of not attending work because this preventive behaviour will reduce the probability of illness and the corresponding health loss.

Our main results show that preventive behaviour can contribute to a decrease in the illness absenteeism rate of the Portuguese workforce. However, secondary prevention revealed no influence.

Finally in the last essay, we found there was a lack of consensus in the literature regarding the effects of individual's economic activity on health behaviour (especially preventive behaviour). This requires further analysis and, in our opinion, warrants an additional study. In our case, we attempt to model simultaneous decisions regarding primary and secondary prevention, while taking into account socio-economic, health variables and health behaviour, focusing especially on the effect of unemployment.

The interest of this research was: 1) unemployment may lead the individual to a poorer general health status, 2) on the other hand, unemployed individuals have more free time to use preventive health care, either primary or secondary prevention, 3) if the unemployed use their free time in these health-promoting activities, could have a positive effect on the health status of individuals who somehow can mitigate the effect of unemployment. We look at how this is related to employment status using a national representative survey.

The contribution of our study is to model the decision to consume primary or secondary prevention while taking into account the unobservable factors that can simultaneously influence both decisions. We use a system of probits, where decisions to take preventive actions, both primary and secondary, can be considered simultaneous.

Our results suggest that socioeconomic factors and health behaviour have a role to play in the possible design of a profile for simultaneous demand of preventive care.

Regarding practical and theoretical contributions extracted from this long work, our time spent researching was made richer by the choice to use different types of methodologies in the search for answers to fundamental and interesting research questions. Each chapter presents a different econometric model, from a logit to a Two-Part Model (including a two part model with a GLM in the second part) and a system of probits. All the chapters have a theoretical framework which goes beyond Cost valuation, Allen's model and the Grossman model.

Summing-up the outline of thesis: 1) three main chapters follow a paper like structure. This includes an introduction, review of the literature, empirical strategy, method (measurement instrument, sample and procedures), results, conclusion, limitations and a final note; 2) we present a global analysis, discuss the results and present a final conclusion. We start by analysing the global results. Then we present the general conclusions along with some policy implications, while also outlining the study's limitations and future lines of research. Finally, we finish with a short comment.

In what follow, chapter 2 presents the chapter regarding the influence of information on the willingness to pay for influenza prevention: evidence for Portugal, Spain and Greece; chapter 3 presents the essay “Does preventive care explain absenteeism? Evidence for Portugal”; finally chapter 4 offers the last essay “Simultaneous preventive health behaviour and the state of employment”.

CHAPTER 2

THE INFLUENCE OF INFORMATION IN THE WILLINGNESS TO PAY FOR INFLUENZA PREVENTION: EVIDENCE FOR PORTUGAL, SPAIN AND GREECE[1]

2.1 Introduction

The announcement of a new influenza virus was made in April 2009 by the Centres for Disease Control and Prevention. In June 2009, the World Health Organization (WHO) declared it as a Phase-six pandemic.[2]

During the autumn of 2009, most European Union (EU) member states included the 2009–2010 pandemic H1N1-influenza vaccine in their influenza vaccination programs. However, the groups targeted for pandemic vaccination programs differed among member states. Moreover, the different countries did not include the vaccine in their general vaccination programs - the vaccine´s initial distribution was to a set of high risk groups: firstly, because the pharmaceutical industry was not capable of supplying the quantity of vaccines needed; secondly, because of the high cost of the vaccine for a period of economic recession and finally, because there were some doubts and concerns regarding the effectiveness of the

[1] An initial version of this Chapter was presented at the European Health Economics Conference (EHEC) Biannual Conference, in Zurique, in July 2012. We are grateful to the participants for their helpful comments and suggestions.

[2] A pandemic will be said to have begun when a new influenza virus subtype is declared to have reached Phase 6. Phase 6 is defined as "Increased and sustained transmission in the general population" WHO pandemic influenza guidelines, 1999–2009 (p. 9).

vaccine (Chor et al., 2009; Lau et al., 2009; Vírseda et al., 2010; Poland, 2010). Furthermore believes about the efficacy of the vaccine influence the decision for accepted or reject it (Redellings et al., 2012).

A Public Health crisis like the one which occurred in 2009 needed a global, national and local fast response capable of reducing the health risks for the population, as well as its financial burden on the economy (avoiding the spread of disease, its consequences in terms of morbidity and mortality and consequent losses in productivity). The relevance of this response is even more important because: 1) the main concerns of citizens were related to the effectiveness of the vaccine, the fact that an important share of the population was not included in the H1N1 vaccination national programs, and the high level of geographical spread; and 2) this pandemic crisis generated an unprecedented stream of information, reported with different degrees of accuracy by the television, radio, web, official sources of information and frequently associated to dramatic scenarios. Both facts contributed to public panic and general alarm, generating a need to identify and better understand the role of the Medias in the decision of demand or not preventive health care.

Arrow (1963) argued that consumers in the health care market often demand medical care despite having limited information on product characteristics and prices, and frequently rely on the providers to act as their agents. The provision of information is arguably sufficient to let consumers make a rational decision. However, according to Hibbard and Peters (2003), the most important factor is the process of using data to inform, which is actually quite complex. The same authors claim that "to use comparative information, health care consumers must be able to take in and process the information, correctly interpret it, identify the important factors to integrate into a decision, weight those factors in ways that match the individual's needs and values, make trade-offs, and bring all the factors together into a choice" (Hibbard and Peters, 2003, p.315). The literature also points out that as the number of options and the volume of information increase, the ability to use all of it in the choice decision declines and individuals tend to let a single factor dominate their decision, while leaving out other important elements (Montgomery and Svenson, 1989). The pre-pandemic and pandemic period is complex and unfamiliar for individuals. It is characterized by the preferences of health care consumers unstable and by the empowerment of individuals who become more involved in their care. Despite the empowerment of the patient and the individual search for information recent research notes that the Internet has not completely replaced information-seeking from health professionals (Hardyman et al., 2005; Basch et. al., 2004). But again this period is atypical,

and the closest provider the General Practice (GP), also had different perspectives about the effectiveness of the vaccine. These particulars increase the need of knowing how the information is presented and used especially because the final decision relies on individuals. We could infer that pandemic situations are in fact unpredictable and call for rapid responses, for that and when possible it is important to identify misconceptions, the acceptance or not regarding the vaccine and the role of Media in this process.

Having in mind the need for a proactive public policy the main aim of this chapter is to study and analyse potential interdependencies between information held by the individual and his decision process regarding willingness to pay (WTP) for preventive health care (PHC).

A survey on WTP for a H1N1 influenza vaccine was conducted so as to measure the good's potential value to the individual if the hypothetical were to occur in the future. It is worth mentioning that the distribution, as well as the control and eligible groups were under the supervision of the Health Ministry. Furthermore, the vaccine was only available on the public market. For measuring WTP we used a contingent valuation method (CV), in which health care economics and economists in other areas have shown increasing interest. CV asks people to give their maximum WTP for a service or a good of a hypothetical market.[3] The literature on this subject suggests that the WTP is greatly influenced by the degree of information provided by the survey design and also by how that information is presented (Olsen et al., 2004; Smith, 2008; Pedersen et al., 2011; Neumann et al., 2012).

In this study, the importance of information in the decision process is analysed in a different perspective relative to other studies. As has been stated, the CV method considers a hypothetical market in which individuals, for the most part, do not have complete knowledge about the good. This could challenge a key axiom of economic theory – as the characteristics of a good change so does its value, *Ceteris paribus* (Chestnut et al., 1996 and Chiu et al., 1999). In the case of the 2009-2010 H1N1 pandemic, a global alert was issued and the general population were constantly being updated on the number of infected people, deaths and

3 For a review of the literature see Diener et al. (1998) and Smith (2008).

characteristics of the vaccine. This guarantees a clearly defined scenario and insures the results are economically valid. In short, our explanatory essay considers a market which is still hypothetical for private consumers but in which the good is already known, which is fundamental for the validity of the WTP results.

It is therefore vital for the CV survey to be behavioural in design, seeking to measure the intention to pay for a specific programme, with specific attributes; in a clearly defined and realistic manner (Smith, 2003). One of the key factors is the type of questions which make up the survey. For instance, Johannesson and Jonsson (1991) concluded that discrete valuation questions work better than open-ended CV questions, especially because they give some information about the price to the individuals. We used the second method and feel comfortable doing so for the following reasons: 1) the good are already provided in the public market; 2) individuals already have information on the price (public) of the good given that the value that national governments paid for the vaccines was made public. The latter could also be a way of eliminating the usual bias associated to a hypothetical market (Smith, 2003).

The metric to assess information and knowledge held by the individuals is a crucial issue for the design of public intervention policies. By definition,[4] knowledge is understood as the expertise and skills acquired by a person through experience or education. Knowledge acquisition involves complex cognitive processes:[5] such as perception, learning, communication, association and reasoning.

Information is tangible and appears in the form of items or objects outside the human mind, and can be defined also as knowledge that can be transmitted without loss of integrity (Kogut, and Zander, 1992). Information *per se* contains no knowledge, and both data and information require knowledge to be interpretable, in order to create new knowledge (Nonaka and Takeuchi, 1995). Knowledge and Information are not understood by all equally. Their absorption depends on cultural, vocabulary or own assumptions (tacit) and new understanding

4 Compact Oxford Dictionary, http://www.askoxford.com.
5 In organizations, knowledge can be divided into tacit knowledge, which involves senses, skills and intuition; and into explicit knowledge, which is formulated and captured, thus existing in the form of books and manuals (Von Krogh et al., 2000).

coming from the act of reflection (Schön, 1983). Reflection, in all its modalities (action, dialog, etc.), can transform tacit into explicit understanding. Therefore, it enables us to recognize how to learn. In this context, information plays an important role as a catalyst for reflection that may, by reaching its consumer, expand or relocate his knowledge state.

To say that there is an ongoing, worldwide revolution in health information is a vast understatement. Today, it is not uncommon for health information consumers to hear of a breaking story in the news, research about it on the Internet, and then contact their GP with several questions (Cline and Haynes, 2001; Fox and Rainie, 2002; Murray et al., 2003, Nelson et al, 2004). At all stages of a disease, from prevention to diagnosis, treatment and end of life, effective health communication can empower people to make informed health-related decisions and to engage in behaviour that can improve their health. Indeed, the results from 25 years of health communication research began with a noticeable impact on mortality rates for diseases such as cancer, as recipients of health communication messages. People have begun to adopt healthier lifestyles and engage in routine preventative screenings (Hiatt et al., 2001). In fact, the impact of health information could be even greater:[6] current trends in these studies and practice of medical communication indicate that practitioners have moved away from a paternalistic model of patient-provider interaction. Instead, contemporary models emphasize the incorporation of values into decision-making processes and the participation of informed patients in medical interactions (Ballard-Reisch, 1990; Emanuel and Emanuel, 1991; Epstein et al., 2004).

The main contribution of our work is analysing the role that information could play in the willingness to pay for preventive health care, not only because this field requires further research, but also because we believe it can contribute to more effective public policy. Moreover, public authorities should be aware of how health care consumers obtain the information they need in order to achieve their goals.

6 For a review of the literature on measures and methods see: Anker et al. (2011).

We analysed data taken from a sample consisting of university students from various faculties in Portugal, Spain and Greece. Universities are potential outbreak centres due to their population, high levels of social contact and, unlike what happens with the seasonal flu shot, they are a risk group for H1N1 (high levels of morbidity and mortality). As a vehicle of transmission, students' behaviour could have an impact on the larger communities of which they are a part (Van et al., 2010; Park et al., 2010).

As means of obtaining information[7] our survey includes: the media (TV, radio, internet and newspapers), time spent searching for information, the course attended by individuals and, as an a priori measure, whether or not the individual took the influenza shot (as a proxy of previous experience).

For all of the sources mentioned above, it became essential to understand the knowledge and attitudes of the university students towards a pandemic so that the health authorities could design directives in the case of a public health crisis. We believe that the identification of the main vehicle of information is central because students are capable of influencing those around them. Authorities will therefore respond to different groups accordingly. Finding the main information vehicle makes it easier to determine where resources must be allocated, and it also improves the efficiency with which practices and policies are transmitted to groups. This creates a more efficient market, since the information is well targeted and consumers have more information at their disposal.

The survey was carried out in the already mentioned countries in 2009. In Portugal data was collected by the author of this thesis, while, in Greece and Spain it was collected by Professors Nikolaos Georgantzis and Antonios Proestakis, and Professor Aurora Garcia Gallego, respectively. The universities enrolled in the survey were the New University of Lisbon, The University of Coimbra and The University of Granada. For each country and

7 We define the search of information as the deliberate behaviour of searching for information. This is in line with the work of Niederdeppe et al. (2007).

University we collected data from the faculties of medicine, economics, law, dentistry and pharmacy.

Our dependent variable (WTP), which measures de willingness to pay for a H1N1 influenza vaccine, has some particular characteristics: it is non-negative, displays many zeros, is skewed in the non-zero range and is intrinsically heteroscedastic (variance increases with the mean). Therefore, as a methodological strategy to analyse the data we initially consider two alternative econometric specifications: a Tobit model and a Two-Part model that allow us to identify the covariates that influence the individual's WTP.

For regressors, the model includes the measures of information mentioned above, along with an array of variables capturing individual's health state socioeconomic position.

This study has the potential of contributing to the design of a public answer to a public health crisis, given that it investigates the potential change in the behaviour of individuals, which are understood as an accumulation of experiences, education, attitudes, and opinions resulting from their life experiences (Poland, 2010). As a result, educational interventions could lead to increased consumer awareness regarding cost, and help find the probable willingness to pay for health care. This is especially true for a period where efficiency should be a goal for the Health Care System. It also allows the study of individuals' health conditions by checking if they belong to a risk group, if that group is eligible for seasonal influenza vaccination and knowing how they searched for information on H1N1.

We structured the chapter as follows. First, we briefly revisit the literature. Next, we discuss the conceptual relationship between WTP and information. Then we explain the method used, present and discuss the main outcomes and make some inferences on the main limitations of the study. Finally, we present some possible pathways for future research.

2.2 Literature Review

Researchers' interest in the search for health information and its impact on the demand for health care is growing. This is a result of an increase in the number of individuals who seek information on health care issues and who now have access to multiple information sources, including online resources (Neuhauser and Kreps, 2003).The main reason for this increase

may be the newfound role of communication platforms, with special focus on the internet and also the growing interest individuals have in being able to influence a medical decision.

The seminal work of Kenkel (1990) explored the role health information has on the demand for medical care. The main purpose of this paper is empirical, look to information and visits to doctor as endogenous variables. In that study, to measure information it was used a proxy measured by individual's responses to nine questions. The questions try to capture the information and knowledge individuals have on cervical cancer and Pap-smear testing services and were collected in the HPKAP survey. The author used data taken from a nationally representative household survey, conducted in the US. Kenkel concluded that patients who are better informed and are more knowledgeable tend to seek preventive care and favour health behaviours that improve their health. Another study, developed by Hsieh and Lin (1997) also produced similar results to those found by Kenkel (1990, 1991).

Parente et al.'s (2003) study is in line with those mentioned before. However, a direct measure for knowledge was used, which was a novelty. They explicitly examined the impact of elderly peoples' knowledge of Medicare benefits on the demand for preventive health care (influenza vaccination and mammography). The results suggest that (even controlling for prior use) knowledge of the insurance benefit is one of the strongest factors affecting the use of preventive care regarding the elderly. This is different to Kenkel's (1990) study that measured consumers' health knowledge through the symptoms associated with diabetes, heart disease, cancer and tuberculosis and the Hsieh and Lin (1997) study that measured the effect of information on demand for preventive care (tests for blood pressure and blood sugar levels, and urinalysis) among elderly people in Taiwan. From these studies, we can infer that information, or knowledge, has a role to play in the demand for preventive care.

Despite the apparent consensus on the effect of information on the demand for medical care, if we search for this effect using different study designs (e.g. clinical trials), this consensus

vanishes. Carney et al. (2000) reviewed these studies and concluded that self-care information decreases the demand for medical care.[8]

There is a possible explanation for these different results. It might be relevant to look at the different measures used to capture health care seekers' information and knowledge, as well as the measures used in the different studies.

Considering the goal of our study, it is important to know what consumers use as their main channel for self-care information. For instance, Wagner et al. (2001) modelled the demand for reference books, on-call nurses and computers as means of obtaining health information. The results of this experiment suggest an increase in the use of reference books (14.9%), telephone advice nurses (6.3%) and computers for health information (4.7%). They also found advertising to be a contributing factor for the use of self-care books. However, it seems to have little effect when it comes to the use of on-call nurses and computers. Additionally, Damman et al. (2009) found complementary results for the analysis of this phenomenon by looking at the seekers of self-care information in a different perspective. Using semi-structured cognitive interviews, they asked interviewees to think out loud about a series of questions, while prompting them with three web pages containing comparative health care information. The results showed several barriers in the decision making process, particularly regarding the amount of information and the interpretation of detailed information.

As a global conclusion, we can state that information and knowledge may have an impact on the demand for preventive care; that the demand for self-care information has increased in the last few years and that the way in which information is presented to people is as important as its dissemination.

8 Self-care information is the information found by consumers using, for example, medical reference books, telephone advice, nurses, media and so on.

2.3 Empirical Strategy

As mentioned in the previous section, our aim is to analyse the impact that health information has on consumers' willingness to pay for an influenza vaccine. The outcome variable is non-negative, presenting a significant fraction of zeros and its distribution has positive skewness.[9] We estimated the skewness and kurtosis of the willingness to pay and the results confirm a positive skewness and a considerable non-normal kurtosis (skewness=4.432 and kurtosis=29.074). These particular characteristics rule out the possibility of using simple linear regression models.[10]

The standard linear model is not an answer, especially when dealing with the skewness of the data, because it may predict negative or nonsensical values and the zero mass may respond differently to covariates. Extensive literature, both methodological and applied, on how to model these type of variables has been developed over the last few decades (Duan, Manning et al., 1983; Manning, Newhouse et al., 1987; Keeler, Manning et al., 1988; Manning and Marquis, 1996; Mullahy, 1998; Ruiz et al., 2007).

Given the characteristics of the dependent variable, we have two main avenues for modelling our data: a Tobit Model or a Two-Part Model.

The Tobit model works well when the variable is censored, but it is hypersensitive to minor departures from the normality assumption.[11] In addition, the specification is underpinned on the assumption that a common data generating mechanism generates the zeros and the

---

[9] Health care data has been extensively studied by other authors. See, for example: Duan (1983), Manning (1998), Mullahy (1998), Blough et al. (1999) and Jones (2000) among many others possible references.

[10] For details, consult selected parts of Jones (2000) and Jones and O'Donnell (2002).

[11] As we possessed knowledge on the skewness, we try to use one of the alternatives suggested by the literature: the log model (this type of data is often modelled better as a lognormal). However, it did introduce two complications: a nonzero threshold and a lognormal WTP (Cameron and Triverdi, 2009). After the log transformation we still had negative skewness (lnWTP: skewness=-1.18; kurtosis=7.75).

positives observations (Manning, 2010). For the presented reasons we exclude the use of the Tobit Model as the preferred model.

Nevertheless, in the context of individual health care choices, the decision process may involve two stages and different factors can influence each one differently (Zweifel, 1981). The specification of the empirical model which explains the willingness to pay for the vaccine should recognize the possible two-stage decisional structure. Thus, we adopt a two-part model. Cragg (1971) argued that individuals, when making decisions about the purchase of durable goods, have to make two sequential decisions. In the first stage of the decision process the individual decides whether or not he should buy the durable item, while in the second stage his decision concerns the quantity he should buy. Therefore, two different stochastic models are in play when explaining consumer behaviour. We believe that the same idea can be applied successfully to our data.

The two-part model specification is also interesting from a purely statistical point of view. The first part consists of a binary model that identifies the factors that distinguish between users' (positive observations) and non-users' (zeros) WTP. The second part determines the factors that influence the WTP of those who have a positive WTP (Manning, Newhouse et al., 1987; Deb and Trivedi, 1997; Deb and Trivedi, 2002).

These models are tailored to accommodate the proportion of zeros present in the data.

The general formulation of the two-part model we present follows Cameron and Triverdi's (2005, p.545) process.

Let an individual with fully observed outcome be called a participant in the activity being studied. Define a binary indicator variable d = 1 for participants and d = 0 for non-participants. Suppose that WTP > 0 is observed for participants and WTP = 0 is observed for non-participants. For non-participants we observe only Pr [|d = 0|X]. For participants the conditional density of WTP given WTP > 0 is specified to be f (WTP|d = 1, X), for some choice of density f (·).

In the two-part model framework the density function for *WTP* is then given by

$$f(WTP|x) = \begin{cases} Pr[d = 0|x] & if\ WTP = 0 \\ Pr[d = 1|x] f(WTP|d = 1, x) & if\ WTP > 0 \end{cases} \qquad (1)$$

Usually the participation decision *d* is modelled using either a probit or logit model. A latent variable formulation considers d = 1 if $I = x' \beta + \varepsilon$ exceeds zero, which ensures participation. To ensure positive values for the participants, the density f (WTP|d = 1, x) should be for a positive-valued random variable, such as the log-normal, or an appropriate density for f (.). This could lead to different model specifications for positive values.

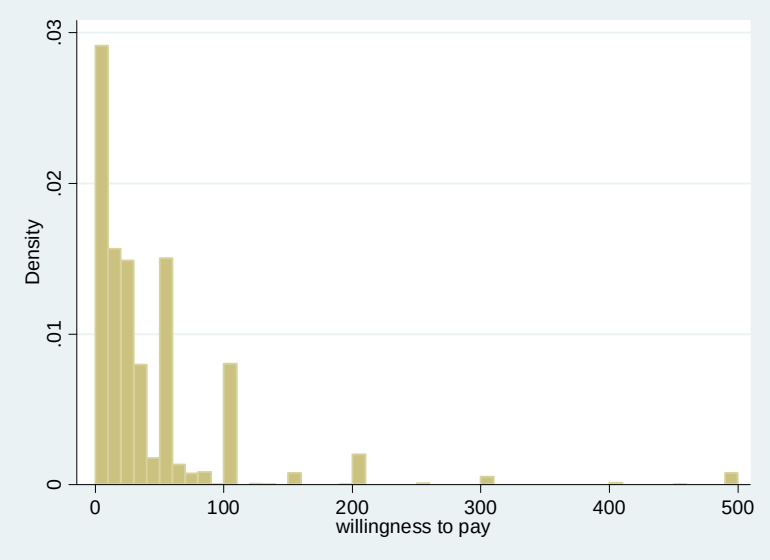
In our specification, we constrained the covariates to be the same in both parts of the model. The estimation can be performed in two steps because the maximum likelihood could be separated. Is straightforward as it separates into a estimation of a discrete choice model using all observations and as a estimation of the parameters of the density f (WTP|d = 1, x) using only observations with WTP > 0. This means the estimation can be performed in two steps. A first step uses the full sample and estimates a binary choice model that analyses, from a regression point of view, the probability of observing a zero or non-zero WTP. The second stage estimates a truncated-at-zero model using only the individuals with positive values for WTP.

For everything explained above, we implemented a two-part model to estimate the impact of information on the WTP, as done by Neuman et al. (2012) and Hammitt and Zhou (2006).

Regarding the second part of the model, and given the characteristics of the dependent variable, we follow two distinct modelling approaches: first, we transform the initial dependent variable (continuous) in an ordinal variable, estimate an ordered probit and, following this, the second approach is estimate a GLM for the continuous variable. The following section explains why we chose these models while illustrating the formulation.

Focusing on the first approach, an exploratory analysis of the dependent variable makes us realize that the variable presented local points with a mass of answers.

Figure 2.1 - Empirical distribution of WTP



We take into account five local points (easily observed on graph 2.1): 30€, 100€, 200€, 300€ and 500€ (the WTP is measured in Euros).

We estimate the level of consumption using an Ordered Probit, given that WTP>0.

Let WTP_i be a categorical ordered random variable representing the willingness to pay of individual i, from 1 to 5. Let WTP_i^* be an unobserved latent variable, representing the true willingness to pay, and X_i is a column vector containing the set of covariates which explain WTP_i^*. We assume that the latent variable WTP_i^* is generated by a linear regression structure such as

$$WTP_i^* = X_i'\beta + \varepsilon_i \qquad (2)$$

where $\varepsilon_i \sim N(0,1)$ is a random error and β is a column vector of the model's coefficients.

The regression model described by the structural model in (2) cannot be estimated because the dependent variable WTP_i^* is latent and therefore unobserved. In order to estimate the parameters β in equation (2), we have to define a rule that relates the two variables (the observed and the latent variable) and can assess the impact each regressor has on the latent variable WTP_i^*. So we conceptualize that the observed responses are the result of a mapping between WTP_i^* and WTP_i, as shown in (3).

$$WTP_i = \begin{cases} 1 \text{ if } 0 < WTP_i^* \leq 30 \\ 2 \text{ if } 30 < WTP_i^* \leq 100 \\ 3 \text{ if } 100 < WTP_i^* \leq 200 \\ 4 \text{ if } 200 < WTP_i^* \leq 300 \\ 5 \text{if } 300 < WTP_i^* \leq 500 \end{cases} \quad (3)$$

Where WTP_i (i=1,...5) are the latent variable's cut-off points that characterize the transition from an observed categorical score to the next.

The other approach, as has been stated, consisted of using GLM to model the second part of the TPM (i.e. positive WTP values). This approach has some interesting advantages: firstly, the interpretation of the results (the scale we look for when modelling WTP is Euros while the scale of estimation in the Tobit would be log-Euros). Secondly, the quality of the retransformed parameter: we have focused on GLMs so as to avoid the difficulty of implementing the Duan (1983) transformation (as was done by Basu and Rathouz (2005) and suggested by Manning (1998) and Mullahy (1998).

$$E(WTP|X) = \Phi(x\hat{\alpha}) \times \exp(x\hat{\beta}_{GLM}) \quad (4)$$

Finally, the transformed data will only have an approximately normal distribution. In a GLM the use of a link function relates μ(x) to a linear specification xT β of covariates.[12] As stated by Basu and Rathouz (2005), the retransformation problem is eliminated by transforming μ(x) instead of dependent variable, and allowing for heteroscedasticity. Log link models with a gamma error distribution are the most common GLM applications in health economics. Like in Blough et al. (1999), Manning and Mullahy (2001), Basu et al. (2004, 2005) and Manning et al. (2005), we also use a GLM application with a log link and a gamma error distribution.

The interpretation of non-linear regression models is not a simple task. This is essentially for two reasons. The first is that in exponential regression models, the conditional mean function, which is usually the descriptor of the distribution that the investigator is interested in, is

12 Therefore directly making straightforward inference and avoiding the need of (re-)transformation as claimed by Cantoni and Ronchetti (2006).

nonlinear. As is readily acknowledged, the β_k parameter no longer measures the impact of a one-unit change of the regressor x_k on the mean function. Therefore, a more elaborated interpretation is required. The second difficulty of interpreting results in a nonlinear model is the fact that the researcher is usually interested in analysing some other descriptors of the conditional distribution of the count variable. For this reason, it is of interest to calculate the marginal effects.

2.4 Data and variables

The data were collected making use of a questionnaire developed by the researchers and applied in Portugal, Spain and Greece.

In what follows we present the measurement instrument, sample and procedures.

2.4.1 Measurement Instrument

The questionnaire was divided in three parts. The first part collected health information, the second part information on health habits and the willingness to pay for the vaccine, and the third part focused on socioeconomic factors. A questionnaire consisting of 17 questions was designed to represent the different dimensions of preventive care, and the willingness to pay for it (see Appendix I). It collected individual level data on demographics, workplace, seasonal influenza vaccination history, health behaviours, perceived threat and information on H1N1, as well as the willingness to pay for the pandemic vaccine. The questionnaire included open-ended, multiple choice and Likert Scale type questions.

Most of the questions were adapted from the literature (Doebbeling et al., 1997; Pareek et al., 2008), while others were developed by the authors for this particular research. The main reason for including new questions was that the other measurement instruments did not

include a sufficient number of items and were not specific enough:[13] it was necessary to measure the perception of individuals in terms of information acquired as well as their willingness to pay. Secondly, the nature of our sample (university students) implied some questions had to be changed.

The design of the questionnaire was piloted on representative members of the study's population. The results of the pilot study indicated the existence of some ambiguity regarding the interpretation of certain questions. Therefore, some of the questions were reformulated. In a second stage, the new version (which included the changes) was again tested on a sample of students.

The data assortment occurred during a period where individuals faced the perceived risk of getting the H1N1 virus. This may have caused the results to have an upwards bias. The individuals were asked the maximum value, in Euros, that they were willing to pay for the influenza vaccine against H1N1. This question came in the middle of the questionnaire, after a series of questions designed to obtain the students' knowledge on the H1N1 virus, as well as questions on their favourite means of seeking information on the subject.

The final part of the questionnaire contained questions regarding the socioeconomic background of the students and their parents.

2.4.2 Sample and Procedures

A convenience sample consisting of 4,193 individuals (66% female) was collected by the authors and by professors Nikolaos Georgantzis (Universidade of Granada), Antonios Proestakis (Universidade of Granada) and Aurora Garcia Gallego (UJI and Universidade de Granada), who kindly consented the use of the data.

The sampled individuals were undergraduate students from Portuguese, Spanish and Greek universities (1,153 from Portugal, 1,408 from Greece and 1,274 from Spain). The students

[13] A common procedure is to use a questionnaire already validated but given the specificity of our research we have to use different resources.

included in the sample came from different cities, universities and faculties. The faculties included were economics, law, medicine, pharmacy and dentistry. The classes were chosen randomly from each faculty, irrespective of the curriculum year and the kind of class. All the students were invited to participate in the study while attending classes in the various faculties. Despite the study being totally voluntary, it is worth mentioning that no student refused to participate.

All answers were kept anonymous and participants were informed there were no right or wrong answers and that they should answer as frankly as possible. When completed, the questionnaire was returned to the researcher who was present throughout the entire process.

As for the characterization of individuals included in the sample, the mean age was 20.9 years (SD: 4.6) and 66% were female. In addition, 94% of the respondents were full time students.

The individuals' self-risk perception was very low: 79% stated that they were probably, or almost certainly, not included in the national vaccination program; 13% reported belonging to a risk group and only 19% took yearly or sporadic seasonal flu shots. We also found some differences in attitudes and behaviour: 53% of students stated they wash their hands more frequently and 23% said they increase their search for information after gaining knowledge on the H1N1 virus.

The results of the socio-demographic variables in our sample give us confidence on the representativeness of our sample (university students). In fact, the majority of the university population is female, and students in southern European countries do not traditionally work while attending university. Despite the low perception of self-risk and the lack of prevention regarding seasonal flu, there was a change in attitudes and behaviour worth noting (especially concerning the search for information and new hygiene habits).

2.4.2.1 Dependent Variable

This study did not investigate the factors which determine the WTP for the treatment of a specific disease. Instead, it assessed the benefits of preventive care actions and practices.[14] The dependent variable adopted in this essay is the willingness to pay for the H1N1 vaccine.

As mentioned before, we used the CV method for measuring WTP. We chose the CV method because it is the most widely used method for evaluating non-measurable economic benefits or costs, as explained in section 2.1. Mitchell and Carson (1989) claim that CV has greater sensitivity to individuals' conditions and preferences relative to the opportunity cost method and the replacement cost method.

The formulation of the question which is used is fundamental to measure accurately the WTP. Recent literature has suggested the use of the double-bounded dichotomy method, pointing out that it is more efficient. However, it can also introduce a level of bias in the final value of the WTP. A starting bid can influence an individual's judgement once he is facing a hypothetical market. Therefore, we opted for open-ended questions, since there was already some information on the market regarding prices charged (Smith, 2003).

The formulation of the open question in the questionnaire was as follows:

"If you're not eligible for the State's free vaccination program, what would be the maximum price you would be willing to pay for this vaccine? Please complete the sentence:

I would pay up to _____€ for the vaccine."

After analysing the collected data we concluded that some values should be considered absurd (if we consider the rational axioms of economics).[15] For instance: €3,000,000 or €10,000,000. Therefore, we adopted a procedure for identifying potential outliers in our data.

[14] See Zethraeus (1998), Donaldson et al. (1997), Johannesson et al. (1993) for examples of studies on WTP for preventive actions or goods.

[15] Some of the individuals reported a WTP of € 3.000.000 and € 10.000.000. This could be seen as absurd; or there may have been a non-detected problem of mystification; the individuals could have thought that the price being asked was for the total number of vaccines.

For this we used Hadi's method (1992, 1994).[16] The procedure identifies outliers as all observations (n=156) equal or greater than 190 euros. This being said, we did not consider it necessary to exclude from our dataset as outliers observations with a value up to 500 euros (n=138). In fact, when including these observations in the sample, no statistically significant differences were observed. Therefore, we decided to truncate our sample on the right side to values equal or less than 500 euros. After dropping the identified outliers, the sample was reduced to 3,817 valid observations.

The WTP for H1N1 vaccine is clearly higher in Greece than in the Iberian countries; amongst the Iberian countries, Portuguese individuals are more willing to pay for the influenza vaccine.[17] The results for Greece are somewhat surprising, as they were already facing an economic depression cycle during this period. The results are reported in the table 2.1 below.

Table 2.1 - Summary statistics of WTP by country (Dependent Variable)

Dependent Variable WTP	Mean	SD	N	P50	P25
Portugal	€37.22	58.99	1150	€20	€8
Spain	€23.17	42.79	1268	€10	€0
Greece	€51.50	75.06	1399	€30	€10

2.4.2.2 Control variables

Table 2.2 describes the control variables included in the model. Whenever necessary, we give a more detailed explanation of some of the variables.

[16] We performed the *hadimvo* Stata command which includes Hadi's method (1992, 1994).The *hadimvo* comand identifies multiple outliers in multivariate data using the method of Hadi (1992, 1994), creating newvar1 equal to 1 if an observation is an "outlier" and 0 otherwise. Optionally, newvar2 can also be created containing the distances from the basic subset.

[17] The authors have also performed a Non-parametric test on willingness to pay by country and found the existence of a country effect.

The regressors used to explain WTP are grouped into three groups. One first group encompasses the variables meant to capture the individual's health status and potential behaviours, the second group include the information variables, and finally, the last group is dedicated to socio economic and demographics regressors.

In what follows we present some further details about each variable. We begin with the variables adopted to reflect the health status and potential health behaviours. With this first group of variables we try to measure the perceived susceptibility to infection, asking students if they were willing or not to take the H1N1 vaccine. We create a new variable *vacyes* using only those individuals who reported to be vaccinated for sure or very likely.

We capture the knowledge of students about if they below to the risk groups and the motive for that. In addition we ask if the individual take the seasonal vaccine. For that we use an ordinal variable and later we create a new one (Seayes) that capture the information reported by the individuals that always or sometimes take the seasonal flu shot.

Table 2.2 - Definition of the control variables

Variables	Description
Health Variables	
Vac	ordinal variable - Willingness for having the H1N1 vaccine[18]
Vacyes	=1 if the individual said he/she is going to be vaccinated for sure or very likely
Risk_group	=1 if the individual reported belongs to a risk group[19]
Season	Ordinal variable - with three categories
Seasyes	=1 if the individual reports taking sometimes or always the seasonal flu shot
Information Variables	
econ	=1 if the individual studies economics

[18] Ordinal variable with five categories: 1/Vaccinate: 1 (yes, for sure) ;2 (Very likely) ; 3 (Probably not) ; 4 (Sure not); 5 (not).
[19] If the individual reports belonging to the risk group due to occupation, age, or for being sick.

Variables	Description
Med	=1 if the individual studies medicine (the omitted category)
Law	=1 if the individual studies law
Phar	=1 if the individual studies pharmacy
law2	= 1 if the individual studies are correlated with law
Dent	= 1 if the individual studies dentistry
Minfo	minutes spending searching for information per day
Grinfo	ordinal variable that evaluates the additional minutes consumed in the searching for specific H1N1 information per day[20]
grin1	=1 if people have increased info (a lot or a bit) due to virus (vs decrease or remain the same)
med1tv	=1 if the primary mean of information is tv (the omitted category)
med1rad	=1 if the primary mean of information is radio
med1internet	=1 if the primary mean of information is internet
medpap	=1 if the primary mean of information is newspaper
e_dead	ordinal variable that measures the perception of the individual about the future deaths due to H1N1 virus
e_dead1	=1 if the expected deaths increases a lot
e_dead2	=1 if the expected deaths increases a little
e_dead3	=1 if the expected deaths constant
e_dead4	=1 if the expected deaths decreases a little
e_dead5	=1 if the expected deaths decreases a lot
Socio economic and demographic variables	
age	age in years
male	=1 if the individual is a male
occ	Occupation - binary variable that evaluates if the student works
wage	Wage of a worker student
size	Number of people that belong to the family aggregate
f_sal1	=1 if the monthly income is less than 350 Euros

[20] Categories: 1 (increased a lot) ; 2 (increased a little) ; 3 (kept the same) ; 4 (decreased a little) ; 5 (decreased a lot).

Variables	Description
f_sal2	=1 if the monthly income is more than 350 and less than 900 Euros
f_sal3	=1 if the monthly income is more than 901 and less than 2001 Euros
f_sal4	=1 if the monthly income is more than 2001 and less than 3500 Euros
f_sal5	=1 if the monthly income is more than 3500 Euros
Portugal	=1 if the individual lives in Portugal
Spain	=1 if the individual lives in Spain
Greece	=1 if the individual lives in Greece (the omitted category)

Information Variables

The amount and the quality of information provided could change individual's WTP for the H1N1 vaccine. Moreover, the provision of information plays an important role in protecting the public from being exposed to dubious procedures. Consequently, the use of these information variables is of critical importance and fundamental for our essay.

The set of information variables used in the model intend to capture individuals' main channel of information and identify the existence or not of a new behaviour. The channel used may have an influence on WTP. Brodie et al. (1998) argue that the media coverage may be one of the factors influencing peoples' anxiety. However, at the time in question we could not be certain of the extension of this influence on public attention and attitudes. The authors argue that the available information also differs according to the main media source people use. For instance, Jadad and Gagliardi (1998) add that the use of the internet enables consumers to gain free access to an expanding volume of information that can improve their decisions, but can also generate problems and difficulties. It is of great importance that individuals are able to perceive what valuable information is, and what is not.

The original variable in the questionnaire is ordinal ask for the main media used as source of information, has four categories (newspaper, radio, tv and internet) and was transformed into sixteen dummies. We used four dummies, since we only consider the primary channel of information. More details can be seen in table 2.2.

Students included in the research had an academic background linked to economics, law and the medical sciences. The decision to include different backgrounds was based on a statement by Fast et al. (1989, p.70): "some forms of consumer education attempt not only to acquaint

consumers with the availability of lesser-known forms of information, but also to help consumers to evaluate the relative quality of various information sources". We have different students' profiles and the possibility of diagnosis different behaviours in the intention of paying or not for the H1N1 vaccine.

We used a variable to capture the amount of minutes per day that individuals devote to the search of information. So as to reinforce the role of these variables, we considered two different ones: the acquisition of general information and the deliberate acquisition of information on the H1N1 virus. These behaviours could specify a perceived risk of get infection or a preventive action. Unfortunately we are not capable of separate these effects. The variable that measure the search for specific information about H1N1 is ordinal, the authors of the study had constructed a new one taking in account only two categories (see table .2.2).

The questionnaire leads the students to point the potential variability they expect in the number of deaths next year. Again we were using an ordinal variable with five categories that we believe gives information about the individual perceived severity regarding the H1N1 virus. We had five dummies but again we only used one, e_deaths1.

The last set of variables is about socio economic and demographics variables we control for the individual age and gender, covariates that the literature proved to influence the demand for preventive care. The student occupation and the size of the students household.

As income variable we used father's income. Household income is an indicator of students' purchasing power, which can influence the WTP for the vaccine. Therefore, it is an important covariate to be included in the model. However, as a way of not to jeopardize the rate of response to this variable, the questionnaire captured income through a categorical variable. Income is evaluated by a categorical ordinal variable (containing five categories). The father's monthly income (dummy variable) leads to the loss of 1000 observations. We do not include the mother's income because this could lead to the loss of even more observations;

furthermore, it is not statistical significant.[21] In our sample, only 6% of students work and have a low income. As a result, we do not use this information. We use the father's income because traditionally he is the main provider of the financial support for the family (especially in southern countries). It was not possible to calculate the equivalent aggregate income because, despite having information on the size of the family, we did not have information on the number of children living in the household.

Before going any further in our analysis, we present a thorough characterization of the sample. Table 2.3 shows the most common summary statistics, (e.g. average, standard deviation) for all explanatory variables involved in the analysis.

[21] This conclusion was withdrawal after several estimations of the model.

Table 2.3 - Summary statistics of the covariates (mean and standard deviation).

Dependent Variable WTP	Mean	Std. Dev.
Health Variables		
Vacyes	.20	.40
Risk	.119	.324
Seasyes	.188	.39
Information Variables		
Eco	.25	.432
Med	.228	.42
Law	.247	.431
Pharm	.25	.432
law2	.007	.083
Dent	.016	.128
Hourinfo	.956	1.05
grin1	.021	.145
med1tv	.642	.479
med1rad	.035	.184
med1net	.24	.426
med1paper	.064	.245
e_dead1	.076	.266
Socio economic and demographic variables		
Age	20.9	4.66
Male	.338	.473
Size	4.11	1.02
f_sal1	.012	.112
f_sal2	.095	.294
f_sal3	.349	.476
f_sal4	.190	.392
f_sal5	.106	.308
Portugal	.301	.458
Spain	. 332	.471
Greece	.366	.481

As expected, differences were observed depending on the course attended, belonging to a risk group, the primary channel for information used, father's income and country of origin.[22] About 20% of the inquired students expressed their willingness to take the H1N1 vaccine, 19% said they took yearly or sporadic flu shots, and only 12% reported belonging to a risk group. In regard to academic background, we have four main sources (economy, medicine, law and pharmacy) which represent around 98% of the sample. If we look separately at each area, their representativeness is quite similar (small differences were found).

As for the search for information, 96% of individuals said they would do so more. Only 2% of students searched for specific information on the H1N1 virus. The main channel for information used by students was television (64%), followed by the internet (24%), whereas newspapers and radio had only residual values. If we wish to sum up the socioeconomic and demographic characteristics, the average student was 21 years old, 66% were female and the majority belonged to a household aggregate made up of four people. In regard to the father's income, 35% reported earning a monthly salary between 901 and 2001 Euros, whereas 20% belonged to the group earning between 2001 and 3,500 euros. Only 10% belonged to the group earning more than 3,500 euros.

Finally, 36% of the students were Greek, 33% were Spanish and 30% were Portuguese. However, if we take into account the residence criteria then the proportion of students if very similar.

2.5 Results

The model is built around the framework of the Hurdle model. In an attempt to unveil the profile of individuals and the factors that explain their decision process, we begin this section

[22] The mean by country if we use the logarithm of WTP in the sample was: Portugal (3.113215), Spain (2.951316) and Greece (3.52984). In fact these means are quite similar to those reported in table 2.3 even though they have more zeros due to the variables construction.

by studying students' WTP implied by these specifications. We focus our analyses on the effect of different information channels and academic backgrounds have on students' WTP.

As mentioned in the empirical strategy section, we estimated a TPM, and in the second part of the model we follow two different approaches; an Ordered Probit and then a GLM model.

Table 2.4 outlines the output of the estimation using a Two-part model, for the whole sample (first-part of the model). The main results are the following:

1) The estimated coefficient is positive and statistically different from zero for the group of individuals who state they are very likely, or will definitely, take the seasonal vaccine. There is a higher probability that these individuals will pay for the H1N1 vaccine, which is understandable when we take into account the age of the individuals, given that they may belong to the risk group. As for the health variables, both the variable that measures the willingness to take the new H1N1 vaccine and the variable that defines whether or not the individual belongs to the risk group are not statistically significant.

2) Regarding the information variables, student background is a key covariate which assumes an important role in the decision to pay for the vaccine. Economics, law, pharmacy and dentistry students have a higher probability of being willing to pay for the vaccine than medical students. It is possible that the medical students, given the formal education they received, were able to process more effectively all the information produced during this period. Individuals who increased their search for general information (measured in minutes) have a greater willing to pay for the new vaccine than those who did not. The way in which the media presents the news, especially in terms of accuracy, can influence individuals' perception. We found the same pattern in regard to the search of specific information on the H1N1 virus (positive sign and statistical significance). The above explanation, when combined with the capacity to process technical information, could explain this behaviour.

In terms of the media, we only considered the primary channel of information. The results show that those who obtain their information from the internet and newspapers have a lower willing to pay for the vaccine relative to those who get their information from television. The variable representing the use of radio was not statistically significant.

The variable representing the number of expected deaths was also not statistically significant.

3) In relation to the socioeconomic and demographic variables, age and income can predict the probability of a positive WTP. The younger the individual, the greater his probability to have a positive WTP. This could have a rational explanation: the virus showed greater incidence in young people contrary to what is expected with the influenza virus. Income is a categorical variable and the results show that individuals whose fathers' income is in the lower two brackets have a higher probability of a positive WTP than those with more income. The literature on the subject demonstrates the existence of a strong link between formal education, income and health care consumption. Based on this, we assumed that family background could also influence decision on having a positive WTP for the vaccine. However, gender and family size are not relevant factors for the decision as neither variable present statistical significance.

As for country effects, Portuguese students are not willing to pay for the vaccine and, do not present statistical significance; on the other hand Spanish students are willing to pay for it, but less than the Greek students.

Table 2.4 – Results of the first part of the TPM

	Logit model		
Dependent Variable WTP	**Coef.**	**Std. Err**	**z**
Health Variables			
Vacyes	-0.101	0.126	-0.8
Risk	-0.001	0.156	-0.01
seasyes	0.483*	0.136	3.55
Information Variables			
Eco	0.916*	0.148	6.2
Law	0.707*	0.141	5
Pharm	0.396*	0.133	2.98
law2	0.387	0.582	0.66
Dent	1.203**	0.609	1.98
Hourinfo	0.180*	0.058	3.08
grin1	0.830***	0.482	1.72
med1rad	-0.112	0.276	-0.41

med1net	-0.481*	0.113	-4.27
Medpap	-0.706*	0.186	-3.79
e_dead1	0.098	0.184	0.54
Socio economic and demographic variables			
age1	-0.569**	0.223	-2.55
Male	-0.150	0.104	-1.44
Size	0.065	0.048	1.35
f_sal1	0.837***	0.471	1.78
f_sal2	0.450**	0.206	2.19
f_sal3	0.107	0.115	0.92
f_sal4	0.210	0.135	1.56
Portugal	0.133	0.145	0.92
Spain	-1.168*	0.118	-9.9
Const	2.260	0.509	4.44

Notes: Coefficients market with *, ** and *** are significant at 1, 5and 10 per cent level, respectively.

In what concerns the results for the second part of the TPM model, which provide estimates on the impact of different covariates on the WTP for those who are willing to pay for the H1N1vaccine, we present estimates for the two alternative models specified. The left hand panel shows the ordered probit while the right hand panel reports de GLM estimates.

Table 2.5 displays the results for the coefficients resulting from the estimation of the Ordered Probit and the GLM model.

Table 2.5 - A two-part model - Second part an Ordered probit and a GlM with a link function and a Gamma distribution for the positives

	Ordered Probit			**GLM**		
Dependent Variable WTP	**Coef.**	**Std. Err**	**z**	**Coef.**	**Std. Err**	**z**
Health Variables						
Vacyes	0.041	0.063	0.65	0.071	0.081	0.87
Risk	0.094	0.074	1.28	-0.029	0.071	-0.41
Seasyes	0.072	0.054	1.33	0.110***	0.064	1.73
Information Variables						
Eco	0.222*	0.067	3.32	0.310*	0.078	3.96

Law	0.227*	0.067	3.41	0.215*	0.074	2.9
Pharm	0.054	0.070	0.78	0.138***	0.079	1.74
law2	-0.602	0.513	-1.17	0.017	0.476	0.04
Dent	0.150	0.162	0.93	0.320***	0.176	1.82
Hourinfo	0.052**	0.021	2.48	0.067**	0.034	2
grin1	-0.128	0.145	-0.84	-0.242**	0.107	-2.26
med1rad	0.110	0.114	0.96	0.066	0.149	0.44
med1net	-0.041	0.054	-0.76	-0.106***	0.058	-1.82
med1pap	-0.125	0.096	-1.29	-0.214**	0.109	-1.96
e_dead1	0.205**	0.082	2.5	0.144***	0.082	1.76
Socio economic and demographic variables						
age1	-0.473*	0.115	-4.1	-0.344*	0.117	-2.95
Male	-0.28494*	0.049	-5.78	-0.284*	0.061	-4.66
Size	-0.0308	0.023	-1.36	-0.041***	0.024	-1.7
f_sal1	-0.443*	0.199	-2.23	-0.446**	0.196	-2.27
f_sal2	-0.243*	0.081	-3.01	-0.285*	0.103	-2.77
f_sal3	-0.057	0.055	-1.03	-0.139**	0.068	-2.06
f_sal4	0.028	0.063	0.45	-0.071	0.070	-1.02
Portugal	-0.381*	0.055	-6.89	-0.362*	0.065	-5.56
Spain	-0.552*	0.058	-9.5	-0.555*	0.065	-8.52
Const				4.878	0.264	18.51
cut1	-1.286	0.260				
cut2	-0.235	0.260				
cut3	0.456	0.260				
cut4	0.803	0.262				

Notes: Coefficients market with *, ** and *** are significant at 1, 5 and 10 per cent level, respectively.

The results in a global perspective between the two models present a major difference in what concern to statistical significance of the covariates favouring the GLM model. In this field we find the main differences in the set of information variables, in the opposite are the socio-economic variables. One possible explanation for the differences observed in the statistical significance is that the aggregation process that the variable WTP was subjected to estimate the order probit specification may lead to loss of information and to less variability. However,

in terms of signs of the coefficients they are similar across the two models, given some consistency of the results obtained using both specifications.

From the set of health variables none of them are statistically different from zero. Thus the perceived risk seems do not have importance in explaining the decision of paying for the vaccine against H1N1.

Looking to the information variables: 1) the course attended by students help in the attempting to find the possible determinants of the willing to pay by students. In this case students of law and economics who are willing to pay and more willing when compared to medicine students. In fact the coefficient estimated for the WTP of economics and law students are quite similar. 2) In regard to the primary channel for information used by the students, the variables are not statistically significant. 3) Those who consume more general information have a greater WTP but the coefficient for H1N1 information research has not statistical significance. The variable "expected deaths" has a positive sign and is statistically significant. The positive sign means that students expecting an increase in deaths are less willing to pay for the vaccine. [23] The expected sign was the contrary meaning that students will pay more if there is an increase in the number of expected deaths. The perceived risk will be higher and students may invest more in prevention. The other variables of this set do not present statistical significance.

For this sub-population, gender and age explain the WTP for the H1N1 vaccine: younger female participants are more willing to pay than older male participants. In terms of income students whose fathers have lower income are less willing to pay than those whose fathers are at the higher categories of income.

Being a Portuguese or Spanish student means be less will to pay for the vaccine than the Greek students. Attending the magnitude of the covariate we could add that the Portuguese students are more willing to pay than the Spanish ones.

[23] It is important to note that this variable is ordinal: if the variable is 1 a high increase in expected deaths is expected. The variable ranges from 1 to 5 (an expected big increase to a big decrease).

In the GLM (table 2.5) for the positives we find some differences: students who had already taken the seasonal vaccine have a greater WTP. The rationale behind this is easily understood: if these individuals are already certain they will take the seasonal vaccine then their perceived risk is higher. Also could be seen has a previous experience that influence the actual student behaviour. Law, Economics, Pharmacy and Dentistry students are more willing to pay for the H1N1 vaccine than medical students. This is somewhat expected if we take into account their academic background. We might expect medical students' trainee work in hospitals to increase their perceived risk regarding contagious diseases. But immunisation (even for the seasonal vaccine) was never consensual amongst health professionals. The magnitude of the dentistry covariate is somehow surprising, once is one of the highest and there is great similarity with the medicine course in terms of contents. In the opposite side are the pharmaceuticals students which are comprehensive having in mind their background and the product of this hypothetical market.

The more minutes people spend looking for general information, the higher their WTP for the vaccine. However, the search for specific information on the H1N1 virus produces the opposite result. We could accept that the combined effect of this search increase individuals' knowledge and enable them to make the right choices. Students using TV as their primary source of information are more willing to pay than those using the internet and newspapers. This could be related to credibility and whether or not the information is reliable. It is also possible that student's value more treated information and transformed in images.

The variable "expected deaths" present positive sign and statistically significant as in the ordered probit alternative model.

Women and younger people have a greater WTP. Women typically consume more health care and preventive care. Students from large families and students whose fathers have low incomes are less willing to pay. If we look at the fathers' salary as a proxy for their level of education (in this case the level of education would not be higher), the results show an income effect, but also that those with lowers levels of education usually consume less health care (a fact supported by the literature).

In terms of country effects, Greek students have a greater WTP than Portuguese and Spanish students.

2.6 Main conclusions and discussion

This study analyses whether information vehicles and/or students' background suitably explains their WTP for a vaccine against the H1N1 virus. In order to collect primary data and set up a model, an appropriate set of questions was developed using the CV method.

We estimated a TPM using two approaches in what regards the estimation of the second part: in the second part of the model we used an Ordered Probit and a GLM.

In its first stage, the use of the two-part model allowed us to establish the factors which determine whether or not the individual is willing to pay for the H1N1 vaccine. The second part (for the positives) enabled us to outline the main determinants of WTP against the H1N1 virus.

Regarding the first objective, we found that individuals who took the seasonal vaccine were more willing to pay for this new vaccine. Academic background is also an important determinant in the decision; all students, other than medical students, are more willing to pay for the vaccine. The search for more information (general and specific) is also an important factor and has a positive impact on the decision. The publics' exposure to messages, topics, and information on health care differs depending on which media source the individual turns to most often and has an influence on his WTP for preventive health care. We also found an age, income and country effect.

The results from the model using only positives gave us more valuable information. The results between the Ordered and the GLM are neither very different in terms of statistical significance nor in expected signs. We consider the GLM approach to be more feasible, since the estimated coefficients are robust. This, allied with the results of the Ordered Probit in section 2.5, convinced us to illustrate the individuals' profile based on the GLM model.

Those who take the seasonal flu shot have greater WTP. If we accept this as a prior experience (researchers defend that the H1N1 vaccine is similar to the seasonal flu shot), it is easier to understand consumer behaviour, which is understandable given this group may be a part of the risk group.

The information/knowledge received by the formal education sector (course in which the students are enrolled) improves consumers' decision-making ability and contributes to market efficiency. In regard to the search for informal information, we found a curious result:

positive sign for the search of general information and a negative sign for the search of specific information. This presents a picture of a coherent strategy for the search and processing of information by the individuals. We may have found an awareness and preference for certain kinds of information, which may improve the efficiency with which consumers handle information.

The impact the media has on the willingness to pay is clear: with the GLM, students that use TV as their primary channel of information have a greater WTP than those who use the internet or read newspapers. This may be related to individuals' subjective memory.

The age and gender effects are similar in both models. Women and younger people tend to purchase more preventive health care. In this case, young people were part of the risk population.

In both models we found an income effect. Students whose fathers have a medium-to-low salary are less willing to pay for the vaccine. Students with more people living at home are less willing to pay; this may be explained by the income effect.

Greek students have a greater WTP than Portuguese students. This is surprising given that the Greek economy was already in recession during this period. We think that there are cultural reasons for this. For example: Portuguese people expect health care goods to be free or have a low price.

Finally, focusing on our main goal, we can conclude that the WTP for preventive health care is influenced by information in different ways: the most frequent channel used to obtain information, academic background and the deliberate search for information.

Our study has some limitations, and future research may lead to further testing of the results here obtained. First, although our measurement instrument is shown to be reasonable, it needs further improvement in order to better understand health care consumers' decision-making process. Some research had already been done after the pandemic, with different kinds of samples (national representative, focus on students) and mostly cross-section (Redelings et al., 2012; Park et al., 2010; Galarce et al., 2011; Naing and Tan, 2011; Maurer et al., 2010; Akan et al., 2010; Van et al., 2010). We conduct a follow-up study only for the Portuguese students, with a measurement instrument with more questions, as well as different questions, so as to reach a model with better fit. The preliminary results of the follow-up study which

was conducted in Portugal are not surprising: there was some negative attitude and resistance towards vaccination. These were based on worries about the safety and efficacy of the vaccine, which undoubtedly affected the WTP for the vaccine.

However, we feel it is necessary to have at least one more wave to make a panel survey possible. Secondly, the sample may produce biased results given that the questionnaires were given to a restricted set of students. It is a convenience sample with the limitations that all of us know, but is also the most commonly used in studies like ours - an exploratory study. Future work may collect a more representative sample. Lastly, in terms of empirical strategy we believe we may have two different populations and, this being so, a Finite Mixture model would be appropriate and could be a way surpassing the weaknesses of our instruments. Future studies should be carried out using this method so as to improve the quality of the results (even though we had already used a robust estimator like GLM). It would also be interesting to analyse in further detail one channel of communication and its impact on the willingness to pay for preventive health care.

2.6.1 Final Note

The economic value of health care services is now being carefully scrutinized for curative and preventive health care in order make possible the survival of the NHS. This means a demonstration of individuals' WTP for health care services will be required. Despite the limitations mentioned above, the study provides a valuable contribution: if a crisis similar to the influenza pandemic of 2009-2010 occurs again, a preventive public intervention, which takes into account the information phenomenon and its impact on individuals' WTP for health care, could be set into action.

Public authorities could have new strategies at their disposition by: 1) giving clear, reliable and timely information to the primary sources of information used by individuals; 2) calling the attention of individuals to the use of more scientifically based information which provides the real risks of the pandemic.

CHAPTER 3

DOES PREVENTIVE CARE EXPLAIN ABSENTEEISM?

EVIDENCE FOR PORTUGAL[24]

3.1 Introduction

The two thousand largest companies in Portugal lost 7,731 million days of work as a result of illness and 1,665 million days of work as a result of accidents, in 1993. This represents 5.5% of all working days to these companies (Preventing absenteeism at the workplace, (1997)).[25] What is more, according to the Portuguese Social Report (2007) the number of non-worked hours is roughly 102,424 thousands and almost 50% of all non-worked hours are due to illness. These figures illustrate that illness is one leading cause of absenteeism in Portugal.[26]

This reality, illustrated with hard data, justifies the growing interest of researchers on the relationship between absenteeism and its determinants, focusing particularly on the effect of health status. Generally, studies in this field find a statistical association between health and absenteeism rates from workplace.[27] Several possible explanations have been proposed to explain such a relationship. These include individual's characteristics, such as socioeconomics, and health related behaviour. Recently, an increasing interest in well-being and health at work has emerged.

---

[24] An initial version of this Chapter was presented at the HESG Summer conference 2010 (Health Economics Study Group) - 23 to 25 July, Cork University Ireland and at XXX Jornadas de Economia de la Salud - 22 to 25 July, Valência - Espanha. We are grateful to the participants for their helpful comments and suggestions.
[25] Source: European Foundation for the Improvement of Living and Working Conditions.
[26] Absence - non-attendance at work when attendance was scheduled or clearly expected (Social Report, 2007).
[27] See for a review of literature for instance (Allebeck and Mastekaasa, 2004).

Empirical literature linking health to labour market activity has grown over the last 25 years and there is large evidence showing that health affects wages, labour force participation, the number of hours worked, retirement, job turnover, and benefit packages. Currie and Madrian (1999) provide a literature review on these issues. The same authors suggest that health can affect the ability and desire to work – the so called "the work capacity". Thus, both employers and employees have an interest in improving work capacity, given that poor health reduces productivity and results in lower wages due to changes in time allocation. Additionally, because workers with poor health are more costly to employ, wages may be lowered. In the labour economics literature this is known as "reservation wages". Consumption of preventive care may be a way of improving work capacity. Additionally, the individuals´ welfare is understood as a function of two components: consumption (which requires work) and leisure. Therefore, maximization means finding the volume of work that offers the best possible combination of these components (Allebeck and Mastekaasa, 2004). Thus, the mentioned authors conclude that health affects work participation directly through productivity, and indirectly by changing the trade-offs between income and leisure (an increase in the health stock means a decrease in lost days, leading to an increase in time for health, home and income production).

The purpose of this study is to explore the possible association of consumption of preventive health care (e.g. clinical exams – secondary prevention) and absenteeism[28]. As we see it, the channel through which preventive behaviour can be associated with absenteeism rates is through the individual´s behaviour or components of individual´s health status not controlled for in the regression model. As is well known, an individual's health stock is systematically influenced by random events, e.g., in the working context, accidents. It may be conceivable that individuals who consume systematically preventive health care may have a more cautious behaviour, making him less prone to accidents, being in the workplace or anywhere else, thus influencing absenteeism rates. On the hand, regression models never capture all individual's

[28] Secondary prevention: its aim is reducing the consequences of illness, without actually affecting the probability of its occurrence. It promotes the use of screenings set out to detect diseases in an earlier and asymptomatic stage (Kenkel, 2000; Downie, 1999).

heterogeneity. Health is a multidimensional concept difficult to define, and even more difficult to measure. Therefore, all regression models control incompletely for the individual's health status. We hypothesize that individuals who consume preventive services may present some particular characteristics, (personal choices, like medical care consumption, education, lifestyles) which make him better off in some health status dimensions not fully accounted in the regressors. In conclusion, we are using the individuals health related preventive behaviour as a means to include in the covariates some dimensions of individuals behaviour/health status that might influence absenteeism.

Conceptually, the consumption of preventive health care (PHC) may have a three-fold effect: 1) Using the intuition of the Grossman model, PHC is an input into the individual's health production function and its marginal productivity is positive. The earlier detection of the disease could lead to a reduction in the number of medical acts. Thus, PHC influences the individual's health stock, which in turn affects absenteeism rates (fewer lost days is reflected in increased income); 2) if PHC consumption affects the individual's health stock it may also increase the effective time endowments and affect the marginal rate of substitution between goods and leisure. The induced health gains may influence the current behaviour of those who think both in the present and in the future. This decision will depend on how much value people put on their future quality of life; 3) however, the consumption of PHC could, in the short-term, induce the increase of absenteeism rate given that the individual possibly might have miss work to consume it.

Summing up, absenteeism is an important issue for employers, governments and society, motivating the shedding of light on the social and health factors which influence it.[29]

This chapter presents the results on an empirical application aiming at unveiling the conceivable association between absenteeism and the consumption of preventive health care.

[29] Portuguese absenteeism rate was 6,7% in 2007, (data from the "Social Report 2007", Ministry of Solidarity, Employment and Social Security – Portugal, 2007).

The application developed in this chapter will also analyse, and estimate, the impact of other factors that may influence absenteeism rates, other than PHC, namely, socio-economic factors, observed health status dimensions, gender and area of residence, among other factors.

In order to accomplish our goals, we use a nationally representative sample analysing the data using alternative econometric models of the count data family, namely, hurdle and zero inflated models (ZI).

The present study is important for several reasons. By providing empirical evidence on this subject, we are contributing to the literature on the effect of PHC on the illness absenteeism rate. Secondly, it will shed light on a subject for which there is little empirical evidence in Portugal. Highlighting the causes of Portuguese illness absenteeism is relevant because it allows us to better understand the preventive behaviour of Portuguese workers.

The present chapter proceeds as follows. In Section 3.2, we present a synthetic literature review. Section 3.3 briefly discusses the theoretical framework that supports the empirical application. In section 3.4 we provide some information about the Portuguese sickness benefit system. Section 3.5 describes the empirical model and the dataset we used. Section 3.6 presents the results and, finally at section 3.7, closing the chapter, we offer a conclusion and discussion.

3.2 Literature Review

As presented in the section 3.1 absenteeism in economic terms is the result of a discrepancy, regarding contractual hours, between the individuals' marginal rate of substitution and the economic rate of substitution[30]. More, an individual marginal rate of substitution will be particularly high when the individual is sick. The literature review that follows is based on

[30] See for instance Brown and Sessions (1996).

this assumption. We start by exploring the relationship between absenteeism and worker health status and later we explore the impact of other determinants of the phenomenon.

Despite the literature search done, we did not find studies that devoted their attention to the effect of consume of preventive care (clinical exams – secondary prevention) and absenteeism rate. The focus of the studies found was mainly on the individual health status and the impact on the absenteeism rate.

One of the de most important work in this research field was developed by Vistnes (1997) who found evidence suggesting that health status can explain absenteeism rates more consistently than economic factors. Workers who prefer not to be absent from work, or who have good health, tend to be more productive and earn higher wages (Johansson and Palme, 2005; Johansson and Palme, 1996). Barmby and Larguem (2009) extended the previous work of Vistnes(1997) and developed a model of absence from work, incorporating an epidemiological structure. The author found that their measure of sickness had a significant effect on the workforce's probability of being absent. Gilleski (2010) went further and argued that individuals who have serious episodes of illness probably do not react to changes in economic incentives, as supported by Brown and Sessions (1996). The latter found evidence that as sickness increases, leisure time may become more valuable. This is because of a need to recover and due to work becoming increasingly onerous (Brown and Sessions, 1996). Moreover, there is evidence that smoother episodes of illness (flu and colds) have considerable influence on short-term illness absenteeism (Leigh, 1989 and 1991).

The work of Chatterji and Tilley (2002) pioneered a stream of absenteeism literature called presenteeism, it suggests worker attendance, as well as their absence behaviour, influences contract design. If contractual requirements are stringent enough to make workers attend work when they are ill, it could lead to illness being more readily communicated to other workers and have effects on productivity.

There is evidence suggesting that health status can explain illness absenteeism better than economic incentives. Moreover, presenteeism may also generate an externality capable of affecting the entire workforce. Workers and employers face considerable costs associated to absenteeism. As sickness absenteeism increases, so does the value of leisure time. Therefore, it is best to reduce and prevent illness absenteeism. The consumption of preventive health care can be seen as an input for the production of good health. Furthermore, to a certain extent

the individual also determines his or her health, the probability of being ill and his impact on absenteeism.

Other determinants of absenteeism

As has been shown, workers maximize their expected utility function which contains consumption and total leisure time as arguments. Health, whose role we have already explored, is also a function of both personal and of the firm's characteristics. In the following section we explore the role of some of these variables.

Regarding gender, there is evidence that women's sickness absence rate is consistently higher than the corresponding rate for men (Barham and Begum, 2005; Vistnes, 1997; Barmby et al., 1991). Gilleskie (2010) claims that women are about 1.5 times more likely to be absent from work and that they experience a 50 per cent higher percentage of lost scheduled work time than men. Age could have an ambiguous effect: if used as a proxy for experience, it is associated to lower rates of absenteeism. This is because older workers expect longer unemployment spells if fired and, for this reason, are more careful with their work behaviour (Dionne and Dostie, 2007; Barham and Begum, 2005; Vistnes, 1997). However, Ercolani (2006) observes a higher pattern of lost man hours with age.

Variables that measure family responsibilities such as the number of children and marital status can also have an impact on absenteeism. Increased family responsibilities may discourage absenteeism, especially if the employee does not receive paid sick leave or if there are penalties associated to missing work (Vistnes, 1997; Dionne and Dostie, 2007). Vistnes (1997) found that having children increased the probability of women being absent, but not mens. Region of residence could also be connected to the absence rate. Distance to the work place may be a determinant of absenteeism (Barham and Begum, 2005).

Occupational and industry variables are thought to play a role. The industries with the highest absence rates are typically found in the public sector (Scoppa, 2008; Ercolani, 2006). Part time jobs tend to be unstable and are negatively related to the absence rate (Bradley et al., 2007; Hernanz and Toharia, 2006). Belonging to a union has a positive impact on the absenteeism rate (Leigh, 1981; Allen, 1984). Moreover, the industry in which a worker operates and the occupation he has may be correlated with the firm's cost of absenteeism.

Therefore, these factors may reflect the penalties associated to absenteeism, an employee's ability to work despite being ill and occupational injury rates (Vistnes, 1997).

Previous research on firm size has suggested that workers of large firms feel a sense of alienation and are more likely to be absent from work (Fenn and Ashby, 2004; Barmby and Stephan, 2000; Winkelmann, 1999; Vistnes, 1997). Dionne and Dostie (2007) also found lower illness absenteeism rates in small firms, even after controlling for flexible work arrangements.

Gilleskie (2010) argues that the limited literature on absenteeism has used hazard function models to depict transitions between health states, which may be the reason for absenteeism. The author also states that absenteeism could be partly anticipated, which would allow employers to make arrangements and slight changes to production plans. The higher costs of absenteeism are associated to involuntary and unexpected absenteeism (especially when the employees are in a good health state).

3.3 Theoretical Framework

Our primary research question is: does the preventive health care consume among Portuguese workers affect their absenteeism rate due to illness? A priori, the effects that the consumption of preventive care has on employee absenteeism are theoretically unclear. This is because both negative and positive effects may be occurring if we take into account the length of time. If the consumption of preventive health care improves health, also the stock of health while focusing on the long term, then better health leads to fewer illnesses, which in turn translates into fewer missed work days. On the other hand, workers that consume preventive medical care (e.g. secondary prevention) must miss work to do so, especially in the short run, thereby increasing their absenteeism.

Thus, the net effects of the consumption of preventive medical care on absenteeism are unclear.

Most of the existing literature on illness absenteeism is focused on the classic labour-leisure choice model developed by Allen (1981). Barmby et al. (1991), Johansson and Palme (1996), Dunn and Youngblood (1986), Delgado and Kniesner (1997) and Vistnes (1997) use,

implicitly or explicitly, this model in a labour supply perspective on observed worker absence. In the following section we explain the intuition behind the Allen Model using the work of the above researchers as well as some insights into the Grosman model.

The Allen model assumes perfectly competitive, profit maximizing firms. Firms will hire up to the point where the output price times the incremental value of another unit of labour equals the wage.

According to this labour supply framework, absenteeism is the workers response to the firm's set remuneration contract (contracted hours and hourly wage rate). Workers decide whether to accept the contract or not, and even if they accept it, they can deviate from the contracted hours by supplying fewer of them.

The intuition behind Allen's model is that if the contracted working time is higher than the desired working hours, employees have an incentive to miss work. Thus, there is a potential utility gain resulting from absenteeism.

If absence occurs, the firm must be compensated so as to guarantee the continuity of the relationship. In addition to lost earnings, the worker pays a lump sum penalty (D) for each scheduled work period missed. This penalty is observed in the form of a decreased probability of receiving a merit promotion or merit wage increase, and an increased likelihood of being dismissed. In our study, an increase in the consumption of PHC can be seen as an input for the decrease of this penalty, at least in the long run. Preventive health behaviour could be a vehicle for the better health of workers, leading to more healthy days. On the other hand, in the short run, individuals with healthy behaviour may have to be absent from work so as to take part in secondary prevention. Time absent from work is denoted as t^A. If the individual has poor health, then the amount of working days is less than if the individual is healthy.

$$D = D(t^A)\ D' \geq 0,\ \ D'' \geq 0,\ \ D(0) = 0$$

Workers who miss more days pay larger penalties.

The formal analysis of the model is developed below and follows the work of Allen (1981) and Brown and Sessions (1996):

$$\mu=\mu(x\ ,\ l\ ). \qquad (1)$$

Where x is a vector of consumption of intermediate goods and l is the leisure good.

Individuals face a budget constraint represented by:

$$R + w(t^c - t^A) - D(t^A) - x = 0\,. \qquad (2)$$

R is income from sources other than work and w is the exogenous real wage (the price is normalized to one). Therefore, income from other sources (R) plus receipts from work hours minus the loss of income due to absenteeism is totally spent on the consumption of goods and intermediate goods (for the production of health).

Workers face a time constraint, since:

$$t - t^c - t^L = 0\,, \qquad (3)$$

t represents the total amount of time in the period under consideration and t^L is the number of leisure hours when $t^A = 0$ $(t^A + t^L = L)$. This time constraint could possibly be greater since health produces time and healthy time. By substituting (2) and (3) into (1) and differentiating (1) in respect to t^A , we arrive at the first-order equilibrium condition:

$$U_L - (w + D')U_x = 0, \qquad (4)$$

U_K represents the partial derivative of U with respect to $k = L, x$. A worker will be absent on any given day as long as the extra leisure is more valuable than the sum of the wages he would have earned that day minus the loss in future earnings due to absenteeism. However, we must also remember that as the individual gets healthier, the utility of leisure decreases.

By differentiating the first-order conditions (2), (3) and (4) and applying Cramer's Rule, we can show that

$$\frac{\delta t^A}{\delta w} \gtrless 0, \qquad \frac{\delta t^A}{\delta R} > 0, \qquad \frac{\delta t^A}{\delta t^c} > 0, \qquad \frac{\delta t^A}{\delta D} < 0.$$

An increase (decrease) in the wage rate ($\frac{\delta t^A}{\delta w} \gtrless 0$) produces an income effect which increases (decreases) the tendency to be absent if we consider leisure to be a normal good. It also creates a substitution effect that decreases (increase) the tendency unequivocally. An increase in non-labour income ($\frac{\delta t^A}{\delta R} > 0$) leads to more demand for all normal goods and

services, including time absent from work. If the number of contracted work hours changes ($\frac{\delta t^A}{\delta t^c} > 0$), the number of absent hours moves in the same direction. Increased penalties for absenteeism ($\frac{\delta t^A}{\delta D} < 0$) reduce the number of missed days.

3.4 Portuguese sickness benefit system

Some information on the organizational and institutional arrangements of the Portuguese sickness benefit system will help the discussion which will follow the presentation of the empirical results presented in this chapter. This section was written based on the statutory order 28/2004 of February 4, 2004 of the Portuguese Government.

The Portuguese sickness benefit system is a public and compulsory scheme for employees and the self-employed, but it is not universal. It also includes a voluntary scheme for qualified individuals who are not covered by the compulsory social protection system (e.g.: people who work at home). Civil servants and lawyers are covered by special schemes and have their own specific schemes.

In Portugal, the majority of workers have a sickness benefit scheme attached to employment, which probably increases labour participation. In order to qualify (for the benefit scheme) there is a 6-month affiliation period, with registered remuneration having occurred on at least 12 days for the 4 months preceding the incapacity to work. There is no waiting period for cases of hospitalization during the period of maternity allowance or in cases of tuberculosis (otherwise, the waiting period is 3 days).

The incapacity to work must be certified by a doctor, who, issues a certificate that has to be sent to the social security within five days following its occurrence.

The benefits depend on the previously registered earnings and the duration of the incapacity. The daily benefit paid by social protection is set up by applying a percentage. This percentage varies according to the incapacity's duration and the average daily wage for the 6 months preceding the 2 months in which the illness began. It start at 65% for incapacity periods lower than (or equal to) 90 days, 70% for periods between 91 and 365 days, and 75% when the incapacity period is longer than 365 days. The maximum duration of these benefits is 1095 days, after which a disability pension applies.

The sickness benefit system for civil servants is the largest subsystem. The civil servants represent around 14% of the working population.[31] The main differences of this subsystem compared to the general sickness benefit system are: 1) there is no waiting period; 2) there may be a 1/6 cut in income, but if the worker proves his incapacity to work daily payment does not change.

An important question that must be understood is whether the Portuguese sickness benefit system possesses any characteristics that contribute to the absenteeism rate. Frick and Malo (2008) tried to identify the determinants of individual absenteeism for EU14 countries based on two indicators: the generosity of sickness benefits and the strictness of employment protection. Regarding the first indicator, Portugal is in an intermediate position. However, for the second indicator Portugal has one of the highest positions. The first result suggests Portugal does not have a more generous benefit system, which justifies the Portuguese absenteeism rate being near the EU average. On the other hand, Portuguese workers take on average more than eight days of health-related leave, which could be associated to the strictness of the employment protection.[32] In fact, the strictness of the Portuguese labour market is constantly referred to as one of weaknesses of the economy by employers and foreign investors.

3.5 Data, Variables and Methodology

3.5.1 Dataset and dependent variable

The data used in this chapter was taken from the last edition of the National Health Survey (4th NHS), conducted between February 2005 and January 2006 and designed to be representative of the Portuguese population. The survey reflects the geographical structure of

[31] Information for 2005 using data from the census of the Portuguese Ministry of Finances.
[32] See for example The Fourth European Survey on Working Conditions launched in 2005 by The European Foundation for the Improvement of Living and Working Conditions.

the population according to the 2001 Census and collects information from householders residing in various Portuguese regions. It gathered information of 41,300 individuals.

The dataset includes information on individuals' socio-demographic and economic characteristics (age, sex, marital status, education, income, region of residence, health insurance status), health status (self-assessed health, chronic conditions, functional status, stress, body mass index, etc.), medical care utilization (number of doctor visits in a three-month period), variables reflecting the individual's lifestyle (tobacco consumption habits, physical activity, etc.) as well as some characteristics from the labour market (main occupation, condition regarding work, number of hours worked, activity sector).

In terms of preventive medical care, the dataset contains information on primary and secondary prevention. We explored the possibility of using some indicators of secondary prevention, namely: use of the flu shot, cholesterol screening tests, blood-pressure tests, mammography and pap smears.

The final sample used in this empirical application was obtained after defining the population of interest and making some data cleaning procedures. Firstly, we restricted our population to active individuals in the labour market (individuals of working-age who have an occupation in the labour market). Bearing in mind the legal working age and the retirement age in Portugal, we included in our study individuals aged between16 and 65. Secondly, we dropped from or analysis all individuals who were: 1) employees of the armed forces, 2) unemployed, 3) those, disabled in the long term and 4) those searching for their first job.

The generation of the secondary prevention indicators (whether the individual made blood pressure tests or a cholesterol screening test) severely reduced the sample size. Because of sampling design issues, the prevention related questions were not responded by all respondents. The survey design only asked these questions to those who were interviewed between the 27th and the 39th week. Moreover, we also had to deal with missing observations for some relevant questions. Thus the final working sample consists of 5,090 individuals aged 16 to 65.

In what concerns the dependent variable adopted in our model it should reflect absence from work due to the occurrence of an illness. One of the questions in the national Health Survey is: "In the last two weeks, how many days left to do some things that usually does, whether at home, at work or in free time [in day-to-day] due to illness, injury, violence or time to use

health care (such as GPvisits, analysis, radiographs, treatments, admissions, etc.)?" For those who answered affirmatively, the following question was asked: "How many days of school or work have you missed?" Therefore, the dependent variable (absence) in this study is a non-negative integer that measures the number of missed work/school days, for health reasons, in the two weeks preceding the questionnaire.

Absence (the dependent variable) has a mean of .185 days, a minimum of 0 and a maximum of 14, presenting a large percentage of zeros.

From the data analysis we concluded that 889 (17.4%) individuals reported missing work at least once in the two weeks before the questionnaire. If we observe the distribution of the dependent variable we can see that 494 (roughly 10%) of the individuals missed one, two or three days of work, and 185 (3.6%) claimed to have missed work for 14 days. In short, the variable's statistical distribution shows a high number of zeros, a considerable mass of individuals that missed at least three days of work, and a mass of individuals that reported missing all the days in the two weeks of reference. These summary results are in line with data for Portugal in the Fourth European Survey on Working Conditions, launched in 2005 by The European Foundation for the Improvement of Living and Working Conditions.

3.5.2 Empirical model

As is widely recognized in the econometric literature dealing with count data regression models, the natural starting model to analyse data with this characteristics is the Poisson regression specification. However, as is widely acknowledge, the Poisson regression model (PRM) is not the best first choice when the dependent variable presents the property of 'overdispersion' (occurs when the conditional mean exceeds the conditional variance), and that of 'excess zeros' (occurs when there are a large mass concentrated at the zero) as it appears to the case with our data after a brief look at the unconditional summary statistics. The issue is that the PRM does not incorporate individual's unobserved heterogeneity, presented in the data whenever the covariates do not represent the full variability of individual behaviour (Gurmu and Trivedi, 1992, 1996; Gourieroux and Visser, 1997; Mullahy, 1997). Failure to include unmeasured factors in the specification of the model leads to loss of efficiency, induces biases in the variances and, consequently, on testing procedures (Cameron and Trivedi, 1996; Gourieroux and Visser, 1997).

It is worth to mention, again, that our dependent variable includes only non-negative and integer outcomes, a significant fraction of zeros and a distribution characterized by positive skewness, being this in line with other studies developed to analyse the same type of data (Delgado and Kniesner, 1997; Fric and Malo, 2008).

The stringent conditions imposed by the Poisson regression model and the characteristics of our dependent variable motivate us to look at more general count data models that can be better specifications to analyse the data. The Negative Binomial regression model (NB) is generally the specification that the empirical economists consider as a suitable alternative to the PRM. This being said, empirical evidence shows that the NB model, again, might not be the most adequate modelling strategy when the dependent variable is characterized by a large proportion of zeros. Several authors, among them, Gurmu and Trivedi (1996) and Bago d'Uva (2005) show that making the unobserved heterogeneity flexible without accounting for the excess zeros, results in models that do not have a good fit to the data.

Therefore, empirical health economists have pursued the formulation of more general count data models, which in some way may overcome the issues raised in practical applications of the NB regression model. Models of the Zero-Inflated family are one of such general formulations popular to deal with count data characterized by a large mass at zero.

In our case, ZI models can have a double justification, economic and a statistical. Zero Inflated (ZI) models incorporate the idea that there are two sources of zeros, that is, there are two different data generating process to explain the occurrence of a zero (Winkelmann, 2003). In our context, the underlying hypothesis of the model is that there are two alternative processes generating zero absent days. On the one hand, a zero could result from a usual distribution (count data process) that governs the generation of both zeroes and positives ("imperfect state"). This statistical process governs the absenteeism of those individuals who may decide whether to be absent or not. The alternative class of zeros, also called by Lambert (1992) as "perfect state" comes from a process that produce only zeros. In this type of zero, Frick and Malo (2008) defend that individuals will have zero absent days because they follow an absolute rule of no voluntary absenteeism (even if they are ill, they choose not to seek care so as to not miss work). As is easy to conclude, Frick and Malo (2008) also adopt, and defended, this model in the context of analyzing absenteeism data.

Therefore, a ZI model should be considered in our modelling alternatives. Vistnes (1997), Delgado and Kniesner, (1997), Winkelmann (1999), Frick and Malo (2008), Huver et all. (2012) and Stauband and Winkelman (2011) used a Zero Inflated Negative Binomial (ZINB).

Besides the behavioural justification for the utilization of use of ZI models to analyse our research question, we note that the use of a ZI model may also found its justification as a statistical artefact to inflate the observed zeros. It is known that one statistical effect of the ZI model is that increases the proportion of the zeros that the model predicts.

The probability function of ZINB is a finite mixture of a general count data probability density function (pdf) and a degenerate distribution concentrated at zero[33]. Below we show the pdf of a general ZI model,

$$Prob[Y_i = 0] = q_i + (1 - q_i)R_i(0) \quad (1)$$

$$Prob[Y_i = j > 0] = (1 - q_i)R_i(j) \quad (2)$$

$R_i(.)$ is a count data pdf, which in our case is a negative binomial distribution. If $Y_i = 0$ then two possibilities arise to generate this observation: q_i the probability of being at the "perfect state" only zeros are generated, and $(1 - q_i)$ stand for the situations where Y_i could be a positive observation. The analytical pdf of the negative binomial model can be seen at Cameron and Trivedi (2005).

In the ZI specification one individual may belong to one of two alternative populations. Either the individual is a "perfect state", in which case he has zero days of absenteeism, with probability one, or he is in an "imperfect state", and can have any number of days off work. In our ZI specification we let the probability of belonging to one of the two alternative populations to depend on a set of covariates. We adopted a simple logit model to model this feature of the ZI model.

[33] See Winkelmann, (2003) and Cameron and Triverdi (1998) for a deep discussion about zero-modified count data models.

Y_i stands for our dependent variable (number of days absent from work) and q_i depends on a specific vector of independent variables, modelled as a logit. We use the same vector of independent variables in the inflated model and in the count model. Namely: preventive, socioeconomic and demographic, occupational variables; variables that assesses health conditions and the National Social Protection system. Further details will be given in section 3.5.3.

An alternative specification that can also be suitable to explain our dependent variable is the hurdle model. In this formulation, besides the statistical justification to its use, there is also an economic reasoning supporting the model. In the first stage the individual decides whether or not to miss work, with one set of factors explaining such decision, while in a second stage, the worker decides the number of days to be absent. A second set of covariates might explain this second decision. However, we adopted the same set of covariates in both parts of the model.

The specification of the hurdle model is as follows,

$$\Pr(Y = 0|X) = f_1(0|\theta_1) \qquad (3)$$

$$\Pr(Y = y|X) = \left(1 - f_1(0|\theta_1)\right) f_2(y|\theta_2, y > 0) \quad y>0 \qquad (4)$$

The decision to miss from work is modelled using a logit model, while the second stage, decision on how many days to miss, given she/he decided to be absent, is modelled using a truncated at zero negative binomial model.

The estimation of both models (ZI and Hurdle) is performed using stata. In the case of the hurdle model, its estimation was done in two separate steps. In the first step we used a logit model to explain the (binary) decision to be absent from work, while in the second stage, and using only those individuals with positive absenteeism, we used a truncated binomial model. Cameron and Trivedi (2005) shows that when the unobservables of both parts of the model are statistically independent the estimation of the model is separable.

In the results section we will use statistical criteria to assess the model that better explains the data, being that the model whose parameters are analysed.

3.5.3 Control variables

Table 3.1 presents the covariates used in our analysis seen as determinants of absenteeism. These were grouped in several clusters: National Social Protection system, socioeconomic characteristics (marital status, age, number of children, gender, educational attainment, income), health status (several indicators of the presence of chronic condition), health related behaviour (smoker, drugs prescription, meals), occupational status (stress, working hours in a week, agriculture industry, construction industry, manufacturing industry, other industry), geographic variables (Norte, Centro, Lisboa, Alentejo, Algarve, Açores, Madeira). Finally we included a covariate to capture the individuals' behaviour regarding the utilization of preventive medical care. The full list of control variables are presented and discussed below.

When selecting the control variables, we took into account Allen's absenteeism model, as well as the main factors that influence absenteeism, which we identified in section 3.2. Table 3.1 shows the complete set of covariates included.

Table 3.1 - Definition of the control variables

Variable	Variable definition
Rheumatism	=1 if the individual has rheumatism
Depressive disorder	=1 if the individual has a depressive disorder
High blood pressure	=1 if the individual suffer from hypertension
Chronic pain	=1 if the individual has a chronic pain
Diabetes	=1 if the individual has diabetes
Obesity	=1 if the individual has obesity
Asthma	=1 if the individual has asthma
kidney stones	=1 if the individual has a kidney stones
Renal failure	=1 if the individual suffer from renal failure
Cancer	=1 if the individual suffer from cancer
Cerebral haemorrhage	=1 if the individual had a cerebral haemorrhage
Emphysema	=1 if the individual suffer from emphysema
Stroke	=1 if the individual suffer from stroke
Smoker	=1 if the individual smokes
Drugs prescription	=1 if the individual takes prescription drugs
Meals	=1 if the individual makes at least three meals
Stress	=1 if the individual has been taken sleeping pills or anxiety pills in the last two weeks

Working hours in a week	number of hours usually worked in a week
Agriculture industry	=1 if work at agriculture (omitted category)
Construction industry	= 1 if work at the construction
Manufacturing industry	= 1 if work at the manufacturing
Other industry	= 1 if work at other industries than the referred above
General Social Security	=1 if is beneficiary of the national health service
Public servant social security	=1 if is beneficiary of the civilian servants health insurance schemes
Married	=1 if the individual is married
Divorce	=1 if the individual is divorced
Single	=1 if the individual is single
Widow	=1 if the individual is widow/widower (omitted category)
Age1	If the individual has between 16 and 29 years(omitted category)
Age 2	If the individual has between 30 and 49 years
Age 3	If the individual has between 49 and 65 years
Children	=1 if there is children in the household
Male	=1 if the individual is male
Educmax	Number of years of schooling completed with success
Income	equivalent monthly income in euros
Norte	=1 if the individual lives in the region “Norte”
Centro	=1 if the individual lives in the region “Centro”
Lisboa	=1 if the individual lives in the region “Lisboa”
Alentejo	=1 if the individual lives in the region “Alentejo”
Algarve	=1 if the individual lives in the region “Algarve”
Açores	=1 if the individual lives in the region “Açores”
Madeira	=1 if the individual lives in the region “Madeira”
Blood pressure or a cholesterol screening test	=1 if the individual made blood pressure or a cholesterol screening test in the last 5 months

Preventive variables

We included a covariate to capture the individual’s behaviour regarding the use of of preventive medical care. The traditional covariates presented in the literature are flu shots, cholesterol screening, blood pressure check-up, breast self-exams, mammograms, pap smears, and prostate screening.

In this empirical study, the variables we used as a measure of preventive care were blood pressure check-ups and cholesterol screenings. The main goal of this procedure was to assure a smoother decrease in our sample and, at the same time, guarantee that preventive behaviour was captured.[34] We note that breast self-exams, mammograms and pap-smears would reduce the scope of the study to women. Although mammograms are recommended to both women and men, they are predominantly done by women. Moreover, in this regards, the national health survey only inquiries women. Questions related to prostrate screening is not included in the national health survey. Finally, flu shots in Portugal are only recommended to risk groups such as senior citizens, children from 6 to 23 months, chronic disease patients, individuals with immunization problems and health professionals that may be exposed to the virus.

In short, we created a (dummy) covariate to measure the utilization of blood-pressure check-ups or cholesterol screenings in the preceding five months as indicators of preventive behaviour of the Portuguese workforce. We consider that measuring individual's preventive behaviour using this method, and data is not the first best option, however, we were limited by data availability. Nonetheless, we believe that this variable, in fact, captures part of the individual's behaviour we are looking for, in special because high blood pressure and high cholesterol are often precursors of heart disease.

Socioeconomic and demographic

In the socioeconomic and demographic set of variables, we include age, gender, schooling, region of residence, marital status, number of children in the household and income.

The demographic and socioeconomic covariates can influence the decision to be absent. This is particularly evident when analysing the covariate age. According to Grossman (1972), age captures the depreciation of health capital, which influences health status and individual

[34] The design of the survey previews the questions about preventive health behaviour between the 27th and the 39th week.

preferences. As the rate of health depreciation increases with age, time being healthy decreases and leisure becomes more attractive, as is expected. In fact, there is evidence that lost hours increases with age. At the same time, age could be seen as a proxy for experience, which is associated to lower rates of absenteeism. This is because older workers expect longer spells of unemployment if fired and are more careful with their work behaviour. We transformed age into a categorical variable with three intervals as a way of analysing whether young people are more absent from work.

We also control for variables that measure the family's responsibilities: the number of children and marital status. Increased family responsibilities may have a direct impact on absenteeism, especially if employees lose money through the lack of paid sick leave or through penalties associated with absenteeism.

The region of residence could also affect absenteeism. Distance to work place could be an important factor when comparing the different regions.

Finally, we used income as a control variable. The labour-leisure model distinguishes between labour income and non-labour income, but we cannot adopt this procedure for our data set. It was not possible to isolate the components of non-labour income. However, we do not believe this creates a large bias in the results, given that the majority of monthly net income in Portugal is made up of wages. In the Portuguese Household Survey (PHS), income is the total monthly net income at the disposal of the household in the month before the interview (including wages, pensions, rents and all the different types of social security benefits), measured by a categorical ordinal variable with ten categories. We use the monthly equivalent income which, in a first stage, allocates an income corresponding to the midpoint of the interval, and in a second stage, interpolates grouped data by taking into account differences in the composition of households - using the modified OECD scale (OECD, 2004). This alternative to the common procedure is a more flexible modelling strategy,[35] but it is not free

[35] The common procedure of including in the model dummies for the ten categories was not attractive since it did not enable taking into account the composition of the households.

of problems. Because we assume household income to be at the midpoint of its income class, for the last class we need to assume an arbitrary value.

Health conditions

We used objective measures of health conditions such as chronic diseases, which, despite being self-reported, individuals know, with some degree of accuracy, whether or not she/he has the condition. The dummy indicators used were: diabetes, asthma, cardiovascular illness, renal failure, kidney stones, cancer, rheumatism, depressive disorder, cerebral haemorrhage, chronic pain).[36]

We could have used self-assessed health indicators (SAH) as a way to include health status in our model. However, SAH is a potentially endogenous variable. We estimate an ordered probit model for SAH, regressing SAH in the set of objective measures of health[37]. The aim of this procedure was to understand if the use of more objective measures will be capable of reflecting health status. For that we analyse the statistical significance and the expected sign of the parameters of the model. The results show that all the variables used to indirectly capture health status are good predictors of SAH. These results give us facts that reinforce our choice.

It was our intention try to capture health behaviours that might have a potential impact on absenteeism.[38] This may include smoking, whether the individuals present signs of stress and number of daily meals.

[36] Chronic diseases could also be a proxy of the stock of health in the past.
[37] See appendix IV- chapter III for the results of the Ordered Probit.
[38] Engagement in sports activities could also be a good proxy for goof health but was only available for a small part of the sample, which leaves us with an even smaller sample

Occupational variables

As occupational variables we considered four dummy variables for the economic activity (EA) of the individuals. The EA variables we used were from construction, manufacturing and agriculture. We grouped all other EA variables into the *other industries* category. We did this because some of the remaining categories were residual and because we can link this aggregated category of *other industries* with the tertiary sector. Therefore, construction and manufacturing industries basically represent the primary and secondary sectors and were chosen because they account for a large percentage of our sample and are traditionally the most hostile environments.

National Social Protection System

We used a dummy to capture whether the individual benefits solely from the general social protection system, or if he also uses other social protection system.[39] As sub-systems we consider the civil servants. A detail explanation was given in section 3.4. Civil servants have a different system protection, which is in fact more generous than the general system of protection.

Until recently, being a civil servant in Portugal meant being employed for life. It is plausible to assume that this security in the public sector decreases the penalty for being absent from work. Sickness benefit for civil servants can result in a 1/6 cut of income, but if the worker proves his incapacity to work it does not suffer any change.

[39] Since 1979 legislation established that all residents have the right to health protection regardless of economic or social status.

3.5.3.1 Summary Statistics

Before going any further in our analysis, we present a characterization of the sample, N = 5.090. Table 3.2 shows the sample mean and standard deviation for all explanatory variables involved in the analysis.

Table 3.2 - Mean and standard deviation of control variables

Control Variables	**Mean**	**Std. Devition**
Health variables		
Rheumatism	.123	.329
Depressive disorder	.081	.273
High blood pressure	.171	.376
Chronic pain	.129	.335
Diabetes	.054	.225
Obesity	.049	.215
Asthma	.045	.208
kidney stones	.046	.210
Renal failure	.010	.100
Cancer	.015	.120
Cerebral haemorrhage	.01	.097
Emphysema	.024	.153
Stroke	.008	.090
Health Behaviours		
Smoker	.22	.414
Drugs prescription	.486	.5
Meals	.922	.268
Occupational variables		
Working hours in a week	40.7	10.1
Stress	.088	.283
Other industry	.484	.5
Construction industry	.090	.286
Manufacturing industry	.097	.296

Protection System		
Civil Servants	.140	.347
Socioeconomic		
Married	.571	.495
Single	.351	.477
Divorce	.038	.192
Age2	.383	.486
Age3	.278	.448
Children	.602	.874
Male	.497	.5
Edumax1	.608	.488
Edumax2	.152	.359
Income	563.8	375.7
Geographic variables		
Norte	.151	.358
Centro	.150	.357
Lisboa	.147	.354
Alentejo	.128	.334
Algarve	.147	.354
Açores	.149	.356
Preventive Variables		
Blood pressure or cholesterol screening test	.204	.403

The prevalence of chronic diseases is most notable in those with high blood pressure (17.1%), rheumatism (12.3%), chronic pain (12.9%) and depressive disorder (8%). In terms of health behaviour, 22% of the Portuguese workforce smokes, and 92% have at least 3 meals per day.

In general, Portuguese workers work forty hours a week. The percentage of people working in the manufacturing sector is close to those working in the construction sector, around 9% of the total; 48% work in other industries and the remaining work in the primary sector. Almost 14% of the workforce is beneficiaries of the public civil servants health protection.

Men make up approximately 50% of the sample, which suggests women have gained significant space in the labour market. Regarding individual's education, the data reveals 76% of individuals in the sample have under 10 years of formal education. Concerning the

individuals' marital status, around 57% are married (the most common) and 39% report being single or divorced. 60% of our study population claim to have children.

In terms of socioeconomic and demographic characteristics, we have an aging workforce (65% are over 30 years old), with low qualifications (60% have only 3 tiers of education).

In regards to the region of residence, the numbers indicate that 15% of individuals reside in the North of Portugal, about 15% in the Centre, 14.7% in the Lisbon and Tagus valley region and approximately 27 % in the south of Portugal (Alentejo and Algarve).

Only, a mere 20% reported doing cholesterol screening or blood pressure check-ups in the last five months.

3.6 Results

The current section presents and discusses the main empirical results that were the outcome of the competing models estimated in this chapter. We start by discussing the results of the model selection method, followed by the economic interpretation of the estimates found.

As described above, we fitted the data using two alternative count data models namely a ZINB and a Hurdle model. We used information criteria to select the model with better fit.

The results for AIC and BIC statistics are presented in table 3.3.

Akaike and Bayesian information criteria is a method of penalized likelihood based on the fitted log-likelihood function (Jones and O´Donnell, 2002). As referred by Sin and White (1996), this model selection technique can be soundly applied to a variety of models, including nested and non-nested, linear and non-linear, correctly specified and wrongly-specified.

Within this model selection methodology, two statistics have received special attention from researchers in count data contexts: the Akaike Information Criteria (AIC) and the Bayesian Information Criterion (BIC).

The Bayesian Information Criterion and the Consistent Akaike Information Criteria statistics are defined as:

$$BIC = -2\ln(L) + kln(n)$$

$$AIC = -2\ln(L) + k(1 + \ln(n))$$

Above, ln (L) represents the logarithm of the likelihood function of the maximum likelihood estimator; k is the number of parameters in the model and n is the sample size. Models that have lower values in both statistics are preferred (Deb and Trivedi, 2009).

Table 3.3 - Goodness-of-fit criteria

Model	**Loglikelihood**	**AIC**	**BIC**
ZINB	-968.5943	2107.189	2662.164
Hurdle NB	-994,636	2808.776	2722.776

After observing these results, we conclude that the model with better fit is the ZINB

This section follows presenting the results of the estimation of the ZINB model as well as the analysis and interpretation of the results.

Table 3.4 presents the extended version of the ZINB that allows for the zero-inflated probability to depend on the explanatory variables. The first panel (called NB) presents the relevant estimates for the population who considers being absent as a possibility. It is the result of the NB count data model. The second right-hand panel (last three columns of the table), presents the results of the inflated model which is a logit. Roughly, it presents the impact of each covariate on the probability of the individual belong to the degenerate distribution of zero, that is, those individuals who have an absolute rule of no absenteeism.

Table 3.4 - Estimation results for ZINB

	NB			INFLATED		
Dependent variable **Missing days - absence from work**	**Coefficient**	**Std. Devition**	**Z**	**Coefficient**	**Std. Devition**	**Z**
Health variables						
Rheumatism	0.089	0.449	0.2	-0.249	0.497	-0.5
Depressive disorder	0.468	0.469	1	-0.461	0.506	-0.91
High blood pressure	-0.152	0.463	-0.33	0.613	0.470	1.3
Chronic pain	-0.680	0.470	-1.45	-0.253	0.481	-0.53
Diabetes	0.826	0.659	1.25	1.286	0.811	1.58
Obesity	-2.454*	0.558	-4.39	-5.321*	2.391	-2.23
Asthma	2.706*	0.835	3.24	0.968	0.764	1.27
kidney stones	-0.265	0.559	-0.47	-0.634	0.636	-1
Renal failure	1.547	1.469	1.05	4.870**	2.340	2.08
Cancer	-0.145	0.863	-0.17	0.267	1.195	0.22
Cerebral haemorrhage	-4.667*	1.374	-3.4	-4.943	7.285	-0.68
Emphysema	-1.491	0.915	-1.63	-1.659	1.371	-1.21
Stroke	-3.368*	0.868	-3.88	-17.570	493.236	-0.04
Health Behaviours						
Smoker	0.633***	0.353	1.79	-0.038	0.384	-0.1
Drugs prescription	1.779*	0.581	3.06	-1.195**	0.533	-2.24
Meals	0.342	0.544	0.63	0.121	0.612	0.2
Occupational variables						
Stress	-0.552	0.392	-1.41	-1.153**	0.458	-2.52
Working hours in a week	0.0001	0.016	0.01	-0.006	0.016	-0.39
Other industry	-0.330	0.706	-0.47	-0.061	0.629	-0.1
Construction industry	-0.639	0.812	-0.79	-1.539**	0.761	-2.02
Manufacturing industry	0.653	0.783	0.83	0.989	0.723	1.37
Protection System						
General Protection	-1.643***	0.930	-1.77	-1.053	0.807	-1.31
Socioeconomic variables						
Married	2.537*	0.657	3.86	5.190	3.246	1.6
Single	1.753**	0.798	2.2	5.112	3.228	1.58

Divorce	3.224*	0.912	3.53	4.446	3.299	1.35
Age2	1.556*	0.490	3.18	1.695*	0.477	3.55
Age3	0.353	0.525	0.67	1.622*	0.587	2.76
Children	0.172	0.260	0.66	0.081	0.232	0.35
Male	0.366	0.443	0.83	-0.590	0.403	-1.46
Edumax1	0.294	0.450	0.65	0.507	0.477	1.06
Edumax2	0.447	0.637	0.7	0.551	0.604	0.91
Income	-0.0003	0.0006	-0.51	-1.8E-05	0.0006	-0.03
Geographic variables						
Norte	0.439	0.570	0.77	-0.417	0.580	-0.72
Centro	-0.422	0.521	-0.81	-0.811	0.541	-1.5
Madeira	-2.342**	1.060	-2.21	-1.196	1.020	-1.17
Alentejo	0.264	0.639	0.41	0.469	0.678	0.69
Algarve	-0.606	0.645	-0.94	-0.601	0.709	-0.85
Açores	-1.872***	1.093	-1.71	-1.459	1.004	-1.45
Preventive variables						
Blood pressure or cholesterol screening test	0.092	0.994	0.09	0.548	0.91	0.6
Constant	-3.236	1.993	-1.62	-1.167	3.850	-0.3

Notes: Coefficients market with *, ** and*** are significant at 1, 5 and 10 per cent level, respectively.

In general our results are in line with the findings of other researchers who investigate similar research questions.

We start our analysis looking to the count data model results, left-hand panel. Regarding the dummy variables, obesity, asthma, cerebral haemorrhage and stroke have statistical significance, but only asthma has a positive effect on absenteeism, meaning that those individuals who report asthma, they, on average, are more absent from work. The other three variables have a negative sign, which is unexpected, therefore requires for additional explanation. These results mean that the higher the coefficients score of the former variables, the fewer predicted days of absence. It could be a case of presenteeism (particularly in the case of obesity), where worker attendance (and absenteeism behaviour) is influenced by contract design. If the contractual requirements are sufficiently stringent so may induce workers to attend work even when they are ill, which in turn could lead to illness being more

readily communicated to other workers and to associated effects on productivity (Chatterji and Tilley, 2002).

The covariates that represent the individuals' health behaviour, only smokers and those who take prescription drugs have a positive sign and are statistically different from zero. This suggests these individuals have, on average, a higher number of missing working days.

The variables indicators of stress, working hours, other industry, construction industry and manufacturing industry are used as a simple representation of the labour market and none of them are statistically different from zero. We will expect that stress and the number of working hours would have a positive impact in the prediction of missing work. Constructing and manufacturing industries are commonly the ones where there is more work accidents that lead to more absenteeism. Despite our prior believes, none of these effects were found.

Individuals who are beneficiaries of the general social security system present a negative sign with statistical significance. These individuals tend to miss work less than those who benefit from other subsystems. The negative sign could be explained by the fact of this protection system is less friendly and traditionally offering less security at work when the worker gets sick.

Regarding the effect of the covariates that reflect marital status (‘married’, "single", and "divorce”), the estimates indicate that these factors have an positive influence on the decision to miss work. This means that the probability of missing work increases with these categories of marital status if compared with the widow category.

Concerning the effect ‘age’, the effects are statistically different from zero in one category, this suggests that age does, in fact, influence the decision of being absent from work. The individuals aged between 30-49 years have a higher probability of missing work when compared to those under 30 and to those above 50.

Contrary to ‘age’, the gender (the covariate ‘male’) of the individual it appears not an important explicative factor for the decision of missing from work. Women may present preventive behaviour that indirectly influences absenteeism, but the variable is not statistical different from zero.

In what regards income, this variable is not statistically different from zero in our results which reinforces the evidence that the decision of missing or not could not be related with economic factors.

The variable that measures the presence of children in the individual's household has no statistical significance, again the literature find some evidence that those who have children miss work more often, in our case does not apply.

Regarding the place of residence (the dummy variables 'North', 'Centre', "Alentejo", "Algarve", "Madeira" and "Açores"), the estimates show that only the covariates Madeira and Açores have a negative sign and statistical significance. People from these regions show a lower probability of missing work than those living in other regions of the country. The distance to the workplace could be a determinant of absenteeism, the bigger the distance the higher the probability of missing work. In this case we do not have data to make this distinguish.

In what concerns the variable used as indicator of preventive behaviour, the coefficient is positive (0.092). This means that in the population of those who consider missing work a possibility, having more preventive behaviour seems to present, on average, a higher number of days missing from work. However, the magnitude of the parameter is very low, without economic significance, and on top of that, it does not present statistical significance. Possibly, this is due to the quality of the data that does not measure adequately the preventive concept we intended to measure.

After analysing the results of the estimation for the group of "perfect state" we find some interesting results that we explore below.

In the set of health variables asthma, cerebral haemorrhage and stroke does not present statistical significance. Only obesity is statistically significance, with a negative sign. This means that overweighed individuals present a lower probability of belonging to the group who have the absolute rule of not miss from work. On the other hand, individuals that suffer from renal failure seems to be more likely to belong to the certain zero group. This is not an expected result and that is difficult to justify

The covariate drugs prescription evidence that those who commonly have more drugs prescription have a lower probability of belonging to the certain zero group.

The variables stress and construction industry and manufacturing industry are statistically different from zero and present a negative sign. The individuals that manifest signs of stress are less likely to belong to the certain group of zeros; the same apply for those who work in in construction industries. These results are in line with the literature.

We did not find statistical significance for the set of Social Protection System, geographical and marital status variables.

Concerning the variable 'age', the effects are statistically different from zero in two categories. This finding suggests that age does influence the probability of an individual belonging to the "zero group". The results show that older individuals are more likely to have behaviour of no absolute absenteeism. The reason could be the high opportunity cost if we used age has a proxy of professional experience.

In what concerns the covariate that we used to capture preventive behaviour, again, it is not statistically different from zero, meaning that it not presents any impact.

3.7 Conclusion, discussion and Limitations

The core of this work was analysing and understanding the effect that the consumption of preventive health care has on the absenteeism rate of Portuguese workers. Unveiling the impact of other determinates constituted a secondary aim of this application.

The theoretical support was based on a mixture of the Grossman Model (1972, 2000) and the Allen Model (1981). The main reason for this was that we were not able to establish a clear link between the consumption of preventive care and the decision to miss work. In order to model the decision to miss work we used Allen's Model insights and then we built a support for our hypotheses based on the Grossman model. We consider the direct consumption of preventive care (blood pressure or cholesterol screening tests) to have an ambiguous impact on the illness absenteeism rate. In the short run, it may have a positive influence on this rate, but in the long run it may improve the health stock and decrease this rate. In fact the consumption of preventive care could have a positive impact in the absenteeism rate; the consumption could imply miss work. But the consumption of preventive care could also reflects on a reduction of the number of medical acts (because of the hypothetical earlier

detection of the disease if the case) and in an increase of health stock. In short we tried to capture some dimensions of the health status that are: or unmeasured or not well represented by the control variables.

The dependent variable that we modelled was the number of days workers reported missing work for health reasons in the fifteen days prior to the survey. To estimate the effect of secondary prevention on the illness absenteeism rate, we specified and estimated various competing models. Each of them was appropriately tailored to handle count data. After having used out some statistical criteria on the competing specifications, we concluded that the best specification for our data was a zero inflated model with the negative binomial as the parent distribution. Therefore, we began by attempting to create a profile of the individuals who missed work for health reasons, focusing our analysis on the covariates that could represent direct or indirect preventive behaviour.

The models were constructed using different type of control variables: measures of health status, health behaviour, occupational variables, insurance status, socioeconomic and demographic variables and preventive health care consumption.

The major conclusions that can be drawn from the selected model are the following: 1) the set health variables chosen to represent the health status of the individuals seem to be a determinant of absenteeism. The most relevant health variables are asthma, stroke, obesity and cerebral haemorrhage. But it is important to mention that the expected sign does not always match the sign of the estimated coefficient. Individuals who claimed either being obese or having had a cerebral haemorrhage or a stroke present a lower probability of missing work. Only those that suffer from asthma present a higher probability of being absent. According to the literature we were expecting chronic diseases to have a greater impact than they actually did (Barmby and Larguem, 2009). Despite some of the most recent studies point out that illness absenteeism is more frequent in individuals who suffer from less acute diseases, especially in the short run (Gilleski, 2010). 2) The occupational variables do not contribute to the decision to be absent; 3) of the socio-economic variables, only marital status and age have a positive impact on the decision. Education, income and having children appear to be unimportant. Surprisingly, variables measuring family responsibilities, such as the number of children (which increases family responsibilities) may discourage absenteeism if the employee loses money through lack of paid sick leave or through penalties associated to absenteeism. These variables do not affect the decision to miss work by the Portuguese

workers. The same is true for income and education and the reason could be that Portuguese Social security system is relatively friendly and can reduce opportunity costs; 4) regarding health behaviour, only those that smoke and who take medication present a higher probability of missing work; 5) in terms of the impact of the Social Protection System, individuals who use the universal social protection System have, on average, a lower number of missing days, than those with the civil servant protection system; 6) of the variables used to control for the place of residence, individuals that live in Madeira and Açores miss work less; 7) secondary prevention does not have an influence on the decision to be absent for health reasons.

We have not found clear evidence showing that the direct measures of preventive care which we used had an effect on the decision to be absent for health reasons. We would expect that in the case of healthy individuals, screening tests would allow early detection of the disease, which would have an impact on the sickness absenteeism rate. However, our study is not able to capture this effect. The reason behind could be the fact of we used a cross-section survey, while these kinds of effects are more easily detected with a dynamic panel. In the case of individuals who possess a disease we can capture the effect in an indirect way - using health variables. Nevertheless, the complementary measures of preventive behaviour show us the importance of prevention when dealing with the illness absenteeism rate. Concerning the "certain" zero group health behaviours and occupational variables could be a predictor of belong to this group as well as age.

In short, sickness absenteeism in Portugal have an impact on economic activity (as is the case in other countries). For instance, even in developed countries where the mortality risk associated to chronic diseases has reduced significantly; the burden of illness is still very high in terms of pain and disability. In economic terms, we have to consider the treatment cost, the loss of productivity, early retirements and greater pressure on the Social Security System. In terms of work environment, despite the huge effort to guarantee safety in the workplace which has been made in this country (making use of legal support and severe fines for those that do not comply), we still have one of the highest rates of work accidents in the European Union. This fact also contributes to the burden of illness absenteeism. Preventive health behaviour can be very useful for the reduction of these costs, given that a large part of them can be prevented /avoided. Friendlier health social protection systems could induce more absenteeism. This is because individuals could have privileged information and can influence the GP's decision. Moreover, nowadays individuals are frequently called to participate in the

final decision. The cost of this phenomenon is very high for all the actors in the economic activity and that is the reason why some changes have been made in recent years.

The first possible drawback of this study is related to the quality of the data. A challenge that every empirical health economist faces is finding data with enough quality to estimate the models. This work stands somewhere between health and labour economics. It was impossible to find a database with information that satisfied our requests. We decided to choose a health database and tailor the data to meet our objectives. In this chapter we have worked in such a context and it was not free of problems. In fact, one of our goals was the analysis of the consumption of preventive health care, which significantly reduced our sample size and, as a result, restricted our work in many ways. The variables used also give us a parsimonious idea of the Portuguese labour market, which is far from ideal.

Secondly, we were not able to make the distinction between false negative and false positive rates in the prevention questions. The way in which the questions were formulated did not assure us that the cholesterol screening or the blood pressure tests were the result of preventive behaviour. In fact, cholesterol screening tests are frequently taken in response to an event and not as pure preventive care.

Thirdly, it was not possible to differentiate between primary and secondary prevention, nor were we able to differentiate between genders, because this would reduce the sample drastically.

Finally, it was not viable to analyse the effect of work patterns on illness absenteeism. For example, we would have liked to compare the differences between white collar and blue collar workers, part time to full time workers and labour-income to non-labour income.

Concluding, when a direct measure is used, the consumption of preventive health care appears not to have important consequences on absenteeism for health reasons. This could be a result of our study only look at a specific aspect of preventive care (cholesterol screaming, blood pressure). That is why we thought it would be relevant to go beyond these objective measures of preventive health care and also use subjective measures of preventive actions. We used some variables that gave us indications on health behaviour and consequently on lifestyle.

Policy makers should be aware of the potential gains from investing more in tertiary prevention - health promotion. Following the effort made over the last decades in primary and

secondary prevention, the Portuguese government should rethink the prevention policy and redirect its objectives to health promotion policies.

Besides the opportunities for further investigation presented by the shortcomings of this study, some further related investigation is possible. Future work is needed in this field because interesting questions remain unanswered.

- There is a small impact of non-labour income in absenteeism?
- Are there sub-populations where preventive action will have a more positive effect?
- What relationship can be drawn between presenteeism and preventive health?
- Could a relationship be drawn between absenteeism and presenteeism?

CHAPTER 4

SIMULTANIOUS PREVENTIVE HEALTH BEHAVIOUR AND THE STATE OF EMPLOYMENT

4.1 Introduction

The association between an individual's health related preventive behaviour (HRPB) and individual's employment status deserves our attention, especially in a period where the Portuguese economy is dealing with a major structural crisis which is leading to imbalances in the labour market.

In the last few decades the Portuguese economy had relatively low rates of unemployment. This trend has suffered a turnover and nowadays the country is facing an unemployment rate of 16.3%.[40]This figure is even higher if we look to at specific social groups (30% in 2011 for people under 25).[41] It is expectable that economic downturns can increase involuntary job loss that may have effects on stress-related diseases and mental health associated with economic insecurity (e.g. Brenner and Mooney, 1983; Catalano and Dooley, 1983; Fenwick and Tausig, 1994).

[40] Data for 2013; source:http://www.pordata.pt/Portugal/Taxa+de+desemprego+total+e+por+sexo+(percentagem)-550 (access 10-03-2014).
[41] Data for 2011; source: http://www.pordata.pt/Portugal/Taxa+de+desemprego+total+e+por+grupo+etario+(percentagem)-553 (access 22-12-2012).

Arinaminpathy and Dye (2010) point out that a significant reduction in household income, in the case of job loss, could make individuals avoid spending money on non-urgent health care needs and on health related preventive actions. Teffta and Kageleiry (2014) argue that the utilization of preventive care may be different during economic downturns. In his line of reasoning, individuals with fewer resources prefer to allocate them to more pressure needs. The issue is that the lag of time between the time of consumption of preventive care and its effect on health status may be very long.

Thinking under this logic, we can anticipate that those unemployed present a lower utilization of preventive medical care when compared to those employed. However, we can think in arguments that predict that those unemployed might present a more favourable HRPB (primary and secondary prevention), especially in a country with an NHS that offers free medical care at point of delivery. For instance, unemployed people have more time to allocate in health promotion activities, such as primary and secondary prevention practices.[42] Time endowments allocation could be central in the promotion of healthy activities.

Therefore, HRPB has become increasingly important in health care systems, because in difficult economic contexts like this, a healthy friendly HRPB can help offsetting the undesirable effect of unemployment in the individual health status.

The relationship between employment status and the utilization for preventive care is still inconsistent. These differences could be linked to the scale of analysis. At the individual level, the effect of unemployment on health is usually negative, while the opposite sign is expected for aggregate data.

Ruhm's (2000, 2002) findings showed that economic upturns, by decreasing the unemployment rate, have a negative effect on physical health; contrary to the notion that

[42] Primary prevention and it main purpose is reducing the probability of contracting an illness. Vaccines, physical exercise and a healthy diet are examples of primary prevention. Ehrlich and Becker (1972) call it self-protection. The second category is secondary prevention and its aim is reducing the consequences of illness, without actually affecting the probability of its occurrence (self-insurance). It promotes the use of screenings set out to detect diseases in an earlier and asymptomatic stage. Finally, tertiary prevention aims to reduce the disability caused by a disease already established (Kenkel, 2000; Downie, 1999).

unemployment worsens physical and mental health. This can be explained by two main factors: the price of leisure (as a good) and the role of income. Non-market "leisure" time increases in periods of economic expansion, making it more costly for individuals to undertake time-intensive health-producing activities such as exercise (Ruhm 2000, 2002). It is likely that income will also grow and its effect is ambiguous. It may increase risky activities such as heavy drinking, drunk driving, smoking and obesity, leading to increased deaths (due to external causes such as motor vehicle fatalities), non-fatal accidents and health problems (Evans and Graham, 1988; Ruhm, 1995, 2000, 2005; Freeman, 1999; Ruhm and Black, 2002). Other researchers (e.g. Baker, 1985; Karasek and Theorell, 1990; Sokejima and Kagamimori, 1998) focus on different variables such as work conditions and work related stress. They found the same results, especially when hazardous environments and extended work hours are involved. In the same line of study, Courtemanche (2009), Xu and Kaestner (2010) demonstrated that having shorter working hours decreases obesity. This is because it is associated to higher levels of exercise and reduced consumption.

Regarding secondary prevention, Ruhm (2000) realized that, in the United States, better health occurred during bad economic times, even though screening tests were less frequent (the same tendency was found for doctor visits and hospital stays) (Ruhm, 2003; Xu and Kaestner, 2010). Dehejia and Lleras-Muney (2004) found that pregnant women obtain earlier and more extensive prenatal care during weak economic periods. Ruhm (2007) shows a similar pattern for sophisticated heart disease treatments (e.g. coronary bypass and angioplasty) among senior citizens.

The aim of this chapter is to investigate if the individual who is unemployed have more preventive actions (HRPB) in terms of regarding primary and secondary prevention. To analyse the research question we use data taken from the Portuguese Health Survey.

Some contributions emerge from this study. First, we investigate the effects that the unemployment status has on the HRPB, here meaning use of primary and secondary prevention. Secondly, from an empirical point of view, we analyse the use of primary and secondary measures jointly, therefore, we take into account the unobservable factors that can simultaneously influence both decisions. Thirdly, it is possible that the individual's characteristics could explain why some individuals are more willing to engage in good health habits (Fukunaga et al. (1997) and Macrae et al. (1984)). Knowing the difficulty we try to identify individual characteristics that justify a more prone behaviour regarding prevention.

Finally, we verify any differences in the utilization for preventive care using sub-samples that take in account the age segment of the individuals. Ehrlich and Chuma (1990) assume that individuals value their health capital differently, depending on age.

Using Portuguese data from the last National Health Survey, the model accommodates both primary and secondary prevention as endogeneous variables. The former refers to interventions that help avoid a given disease (e.g., doing exercise), and the latter is aimed at detecting a disease in its early stage, so that early curative treatment is possible (e.g., measuring blood pressure and mammograms). Focusing on our purposes we split the sample into two sub-samples (young and old adults), so as to assess the determinants of prevention behaviour in different stages of life.

In what concerns our methodological options, we adopt regression models, and specify a system of probits for modelling the simultaneous decision of consuming primary and secondary prevention, while controlling for the characteristics of those who undertake these behaviours.

The results of this research suggest different profiles for those who consume primary and secondary prevention and that, in fact, employment status influence the use of preventive care

The chapter proceeds as follows. In Section 4.2 we briefly discuss the theoretical framework. Section 4.3 describes the empirical strategy. Section 4.4 describes the data and the dependent variables. Section 4.5 presents the results and section 4.6 offers a conclusion.

4.2 Conceptual Framework

Our main theoretical support is Grossman's (1972) seminal work on the demand for health and health care. These studies claim that individuals invest in their own human capital by improving their physical capacity or by increasing their stock of knowledge.

According to this framework, individuals possess an initial stock of health that depreciates over time and on the other hand health risks associated with preventable illnesses will increase with age. Therefore, based upon such an observation, there should be an increase use of preventive health care services with age (kenkel, 1994). Furthermore, Kenkel (1990) extended the health demand model by including health information, given that health

information may affect the perception of individuals, in a positive direction, regarding their need to utilize preventive care services. Grossman (2000) go further supporting that health investment includes all types of health promoting behaviour. For example: consumption of effective medical care, eating healthy foods, exercising, screening exams between others.[43] The model also provides a set of explanatory variables (age, income, education and marital status) that clearly contribute to the health production function. One of the most important factors for this function is the level of education.

In our study, the individual's decision rule is to maximize their expected utility in order to simultaneously determine the optimal amount of primary and secondary prevention they need. Utility theory allows individuals to choose a lifestyle which increases the stock of health capital, thereby decreasing the probability of becoming ill (knowing that the individual has a healthy diet lowers the probability of getting gastrointestinal cancer, exercise decreases the probability of getting cardiovascular disease, and screening exams makes possible early detection and increase the probability of a cure). We are conscious that slightly different reasoning exists. For instance, Luras (2001) does not reject the fact that maximization utility theory has a decision rule. However, he also looks at health norms, traditions and values which may be serving as constraints in the maximisation process. It then follows that individuals eat healthily because they like healthy food, they exercise because it is fun, and they stay away from addictive goods because they do not like drinking or smoking. Rationality may not be present and the effect on utility might follow directly from pleasure. If people take into account their norms, traditions and values they can eat fruit, vegetables and exercise to stay healthy, as well as not smoking so as to avoid lung cancer.

Health demand models are a prolific way of analysing the demand for prevention. We base our work on the fact that consumers can make two related choices: first, they decide the amount of primary prevention they are capable of obtaining by the decision to do (or not) sport; secondly, they make a binary decision about whether to engage (or not) in secondary

[43] The model does not include the difference between curative and preventive care, or between primary and secondary prevention.

prevention (e.g. the individual may decide to take risky behaviour such as not taking screening tests, even if prescribed by the doctors).

We consider it reasonable to expect there to be common elements which affect both decisions. Factors such as expectations about future health and attitudes towards risk might influence both the propensity to do screening tests and the decision to adopt a particular life style. For example, if an individual has a relative with cancer, they may be more willing to do screening tests and more conscious about the benefits of adopting a healthier life style.

We estimated a structural health model where the relation between self-protection, health, perceived risk, attitudes towards secondary prevention and socio-demographic variables were accounted for.

The main hypotheses to be tested were derived from theoretical analysis and from the review of the literature. They are presented below:

Hypothesis 1: employment status can simultaneously influence the demand for primary and secondary prevention. The expected sign is ambiguous according to the literature.

Hypothesis 2: There is an implicit relationship between chronic diseases and the decision to do primary and secondary prevention.

Hypothesis 3: The socio-economic variables play a relevant role as determinants in the two decisions considered in the bivariate probit. This hypothesis intends to list the variables that allow us to characterize the individuals who demand primary and secondary prevention simultaneously.

4.3 Empirical Strategy

In this chapter we are concentrating in understanding how individual's unemployment state affect his HRPB, here seen as both, primary and secondary prevention activities. Therefore, the econometric model has to acknowledge that the individual makes two simultaneous and interrelated decisions. Accordingly, the econometric specification adopted in the chapter consists of two interrelated equations: one first equation explains the individuals' propensity

to engage in physical exercise (primary prevention) while the second explains the propensity to use secondary prevention services.

The first equation regresses a (latent) measure of primary prevention a set of other explanatory covariates. The equation is written as follows,

$$y_1^* = \beta_1' X_1 + \mu_1 \qquad (1)$$

y_1^* is the latent predisposition of the individual to engage in primary prevention activities, X_1 is a vector containing a set of exogenous characteristics determinants of primary prevention. Moreover, β_1 are the parameters of the model and μ_1 is a random error independently and identically normal distributed across individuals.

The second equation of the model regresses the secondary prevention indicator y_2^* on a set of exogenous covariates X_2. Formally, the equation is expressed as follows,

$$y_2^* = \beta_2' X_2 + \mu_2 \qquad (2)$$

y_2^* the latent predisposition of the individual to engage in secondary prevention activities, X_2 is a vector containing a set of exogenous characteristics determinants of secondary prevention activities. Furthermore, β_2 is a vector of parameters of the model and μ_2 is a random error independently and identically normal distributed across individuals.

The former equations (**1**) and (**2**) are dependent on each other, hence constituting a simultaneous equation model. The error terms μ_1 and μ_2 are assumed to be jointly normally distributed, with null vector mean and variance-covariance matrix $\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$. The parameter ρ measures the correlation between the unobservables present in the equations.

The overall model is outlined as follows,

$$\begin{cases} y_1^* = \beta_1' X_1 + \mu_1 \\ y_2^* = \beta_2' X_2 + \mu_2 \end{cases} \qquad (3)$$

where $\lfloor \mu_1 \quad \mu_{2^*} \rfloor \sim N(\mathbf{0}, \Sigma)$.

The vectors of covariates $\mathbf{x}_1$ and $\mathbf{x}_2$ include a constant term.

The latent endogenous variables y_1^* and y_2^* are, by definition, unobserved. The corresponding observed variables denoted y_1 and y_2, are

$$y_1\left(y_1^*\right)=\begin{cases}1 & if \quad y_1^*>0\\ 0 & if \quad y_1^*\le 0\end{cases} \qquad y_2\left(y_2^*\right)=\begin{cases}1 & if \quad y_2^*>0\\ 0 & if \quad y_2^*\le 0\end{cases} \qquad (4)$$

System (3) jointly with the functional relationships (4) constitutes a bivariate probit model (Greene, 2003).

The observed variables to reflect HRPB are y_1 and y_2, are 0/1 variables, where 1 indicates the presence of the characteristics and 0 its absence. A description of these variables will be presented further.

For the estimation of this model we resorted to maximum likelihood estimation (MLE) methods. All parameters of the model are estimated simultaneously taking into account the correlation between the disturbance terms in the two structural equations of model (3). The main advantage of joint estimation through MLE is the gain in efficiency that results from the incorporation of correlation in observables across equations for a given individual (Cameron and Trivedi, 2005).

4.4 Data and variables

The data was taken from the last National Health Survey (4th NHS). This survey was conducted between February 2005 and January 2006 and was intended to be representative of the Portuguese household population. The survey reflects the geographical structure of the population according to the 2001 Census and it collects information from Portuguese households over various regions. It covers 41,300 individuals. However, for reasons to be explained below, the final sample used in this chapter comprises 7,926 observations. We restricted the population to adults over 18 years of age.

4.4.1 The dependent variables - Physical Activity and Consumption of preventive health care.

Primary prevention - Physical activity

The Centre for Disease Control and Prevention (CDC) in the United States makes the distinction between moderate and vigorous activity. Moderate physical activity means more than 30 minutes of exercise a day, more than 5 days a week. Examples of moderate physical activity includes walking briskly (3 miles per hour or faster, but not race-walking), water aerobics, cycling under 10 miles per hour, playing tennis (doubles), ballroom dancing and general gardening. Intense physical activity means exercising more than 20 minutes a day, more than 3 days a week. Examples of this are: race walking, jogging, running, swimming laps, tennis (singles), aerobic dancing, cycling at 10 miles per hour or faster, heavy gardening (continuous digging or hoeing) and hiking uphill or with a heavy backpack.

Following Tavares and Barros (2011), we constructed a binary variable that equals 1 if the individual practices moderate or intense activity and zero if not. We followed these authors to find support on the procedure for excluding the observations in which the individual reports more than 5 hours of physical activity per day. This could indicate that the individual is a professional sportsman, or that he has a strong manual job.

In summary,

$$y_1 = \begin{cases} 1 \;\; if \;\; \text{mod}\, erate\, or\, \text{int}\, ense\, activity\, Physical\;\; activity \\ 0 \;\; otherwise \end{cases}$$

Secondary prevention indicator

To measure secondary prevention we initially planned to use the information available in the sixteen section of the national health survey, entitled Preventive Care.[44] Unfortunately, the

[44] The section includes five questions related to behaviour regarding prevention: flu vaccination, cholesterol check-up, blood pressure check-up, breast cancer and cervical cancer screening.

survey's design does not allow the physical activity and preventive care sections to be simultaneously used to describe the individuals. The reason is that the data reflecting individual's physical activity was only collected between the 14th and the 26th weeks of the survey application and the data collection on preventive care (section sixteen) was only conducted between the 27th and the 39th week. Therefore, no individual answered questions simultaneously on physical activity and preventive care.

Therefore, as last resort to measure secondary prevention, and to find a proxy for this type of medical care utilization, we used the data collect on the "Health Care" section. In this section it is possible to know the reason of the last visit to the doctor. Among all the categories displayed, some of them can clearly be linked to secondary prevention, namely a routine visit, performance of complementary diagnostic exams or measuring blood pressure.[45] This strategy to measure secondary prevention presents an inconvenient that can bias our results. Our population is comprised only of those individuals who visited the doctor in the last three months.

In what follows we make a simple statistical analysis of the dependent variables by gender and age. Before that, we note that 52% of individuals declared to engage on moderate or intense physical activity but only 9% had any secondary prevention action.

By analysing physical activity by gender and age, we can observe that:

(a) the percentage of women doing physical activity decreases slightly with age;

(b) the percentage of men who do physical activity decreases with age;

(c) the behaviour of older individuals (40 years and more) is more similar between gender than of those who are younger (from 18 to 39) (Figure 4.1).

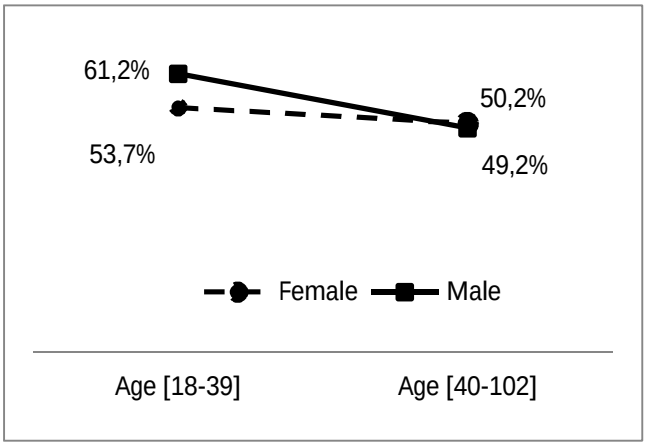
[45] The choice of answers was based on what the literature uses as variables for secondary prevention. See for instance Kenkel (1994).

Taking this information into account we will consider both age (in two classes: from 18 to 39 and over 40) and gender in explaining physical activity. Figure 4.1 also suggests there is an association between age and gender when explaining physical activity.

Table 4.1 - Number of individuals that that take part (or not) in physical activity

		Age [18-39]		**Age [40-102]**	
		Physical activity		**Physical activity**	
		0	**1**	**0**	**1**
Gender	Female	53%	46%	53%	54%
	Male	47%	54%	47%	46%

Figure 4.1 - Percentage of individuals who engages in physical activity



By analysing the variable secondary prevention (prevention 2), table 4.2, according to the gender and age of the individuals, we can observe that:

(a) the percentage of women that do prevention 2 seems to remain unchanged with age;

(b) the percentage of men that do prevention 2 increases with age;

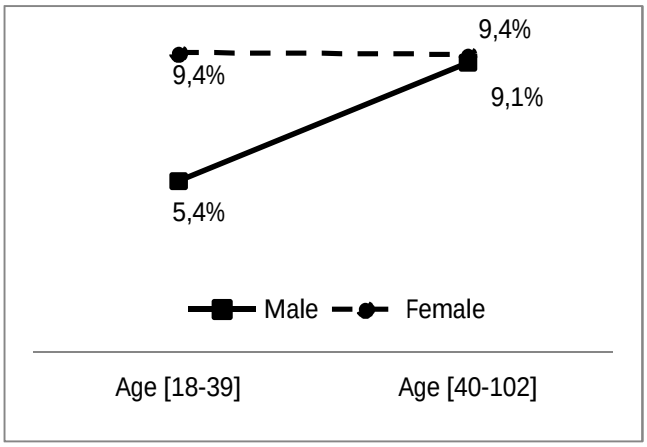
(c) the behaviour of older individuals (over 40) is more similar between gender than that of younger individuals (from 18 to 39) (Figure 4.2).

Taking this information into account we will consider both age (in two classes: from 18 to 39 years and over 40) and gender in explaining prevention 2. Figure 4.2 also suggests a link between age and gender in explaining prevention 2.

Table 4.2- Number of individuals that take (or not) secondary prevention

		Age [18-39] Secondary Prevention		Age [40-102] Secondary Prevention	
		0	**1**	**0**	**1**
Gender	Female	48%	62%	54%	55%
	Male	52%	38%	46%	45%

Figure 4.2 - Percentage of individuals that take secondary prevention



4.4.2 Control variables

In short, our first hypothesis is that decisions concerning physical activity and secondary prevention could be simultaneous. But we also examine whether chronic diseases could have an influence on these decisions. We test for health behaviour, while expecting those with healthy lifestyles do more prevention and more physical activity. We also use socio-economic variables and test for social protection system. These hypotheses intend to list the variables that allow us to characterize the individuals who do sports and secondary prevention. Therefore, we present the independent variables and the justification for their inclusion based

on the literature and on section 4.2. Table 4.3 describes the control variables used in our model.

Health variables - Chronic diseases

Chronic diseases are a major source of adult deaths (heart attacks, strokes, cancer, pulmonary disease, diabetes) and have relatively high prevalence rates (arthritis, back, mental and circulatory problems). This reflects major changes in dietary habits, physical activity levels and tobacco use as a result of industrialization, urbanization, economic development and food market globalization (WHO, 2005).

Those who are generally weaker and more debilitated have a potentially higher probability of getting other diseases. Therefore, they will probably use health care services more often. In fact, there is evidence showing that the marginal effect of a new illness, in terms of pain and severity, may be greater for those who started off with poor health (Türp et al., 2000). Furthermore, rehabilitation and treatment is harder for those who are already sick (Nordin et al., 2002). One would assume these factors would lead to more screening and preventive behaviour for individuals in poorer health (given this can be seen as a perceived risk for getting a disease). On the other hand, those who are in poorer health have less time to receive treatments and screening tests given their physical limitations. The expected results are therefore ambiguous, but since the expected benefits prevention are higher for those with pre-existent medical conditions, these individuals are more likely to engage in preventive behaviour. For instance, if an individual has a circulatory system disease he may be able to do some physical exercise and have a healthy diet; this could prevent or decrease the incidence of further diseases.

In our model we use a set of variables to control for chronic disease: high blood pressure, diabetes, asthma, cancer, cerebral haemorrhage, emphysema and stroke.

Table 4.3- Definition of control variables

Variable	Variable definition
High blood pressure	=1 if the individual suffers from hypertension
Diabetes	=1 if the individual has diabetes
Asthma	=1 if the individual has asthma
Cancer	=1 if the individual suffers from cancer
Cerebral haemorrhage	=1 if the individual had a cerebral haemorrhage
Emphysema	=1 if the individual suffers from emphysema
Stroke	=1 if the individual suffered from stroke
Smoker	=1 if the individual smokes
Obesity	=1 if the individual has obesity
Married	=1 if the individual is married
Divorce	=1 if the individual is divorced
Single	=1 if the individual is single
Age	Age of the individuals
Male	=1 if the individual is male
Educmax	Number of years of schooling completed with success of the most educated person living in the household
Income	equivalent monthly income in euros
Unemployed	=1 if the individual is unemployed

Socio-economic variables

Education

Kenkel (1994) found evidence of a positive relationship between education and the propensity to use medical care: well educated individuals use more preventive health care services because they are more aware of its benefits. Chen and Lange (2008) and Lange (2011) went further and examined how education is related to the choice of whether to undergo screening for breast, colorectal and cervical cancer. Their key results were: highly educated individuals

are more aware of the objective risks, but differences in subjective risks are more closely linked to decisions to undergo screening for them. The main conclusion is that highly educated individuals are able to process information related to medical risks more accurately, so they change behaviour to improve their performance (the authors are not able to say if schooling causes these differences). Zubarik et al. (2000) in previous research achieved similar findings: those with more education may be better informed about the potential benefits of prevention for early detection or the avoidance of disease.

We used a variable to measure the maximum number of completed school years.

Age

In the tradition of the Grossman model, Lairson et al. (2005) found that older women are subject to greater risk and demand more prevention because of its great benefit. But we could argue that older people have less incentive for doing prevention since the pay-off period is shorter and the risks of getting diseases (e.g. chronic diseases) increases with age (Shureiqi et al., 2001). Ehrlich and Chuma (1990) assume that the value that individuals attribute to health capital differs according to age and over their life cycle. So, willingness to participate in activities involving health risks can be explained as a function of this value. Greater exposure to health hazards implies a lesser tendency to invest in health, and hence a less healthy lifestyle. The expected effect of age in the demand for preventive care is ambiguous.

Income

Wu (2003) claims monetary and time costs also vary according to the individual's economic status: those who work are more likely to have higher income and can more easily afford any out-of-pocket costs (likely to increase the stock of health). On the other hand, these costs also increase the opportunity costs of leisure and the investment in health care. The overall effect is not clear, but research by Lairson et.al. (2005) suggests it is positive. We use the monthly equivalent income calculate based on the information given by the data survey.

Unemployment

In this section we did not justify the inclusion of the variables state of employment because we consider that it was already done in some depth in section 4.1.

4.5 Results

The overall purpose of this section is to present the estimation results of the bivariate probit model. We aim to understand the contribution of factors capable of explaining differences in simultaneous HRPB, namely, primary and secondary prevention. Given the aim of this chapter we will focus particularly on the impact of unemployment.

Table 4.4 Bivariate probit parameters for the all sample

Dependent variable - Physical Activity	Coef.	Std. Err	z
Health Variables			
High Blood Pressure	-.0204	.038	-0.54
Anxiety	-.101	.070	-1.45
Diabetes	-.088	.054	-1.63
Asthma	-.020	.066	-0.30
Cancer	-.216*	.105	-2.07
Cerebral Haemorrhage	-.176	.116	-1.52
Emphysema	-.040	.079	-0.51
Stroke	-.364*	.125	-2.90
Health Behaviour			
Smoker	-.031	.038	-0.81
BMI	-.003	.004	-0.75
Socioeconomic			
Single	-.124**	.059	-2.11
Married	.094**	.045	2.11
Age	-.011*	.001	-8.49
Male	.046	.031	1.47
Yearsschool	-.047*	.004	-10.69
Lnycp	-.069*	.027	-2.59
Unemployed	-.529*	.073	-7.24
Retired	-.463*	.0465	-9.96
_cons	1.612	.190	8.50

Notes: Coefficients market with *, ** and*** are significant at 1, 5 and 10 per cent level, respectively

The estimated parameters for the overall sample and for the variable physical activity are shown in table 4.4. In general we find evidence that unemployed individuals present, on average, a lower probability of doing physical exercise when compare with those employed. We also found that chronic diseases also contribute for explaining the decision of doing or not physical exercise.

Results show evidence that the socioeconomic characteristics of the individuals seem to have a determinant role in explaining the demand for preventive care.

In more detail, one of our main aims was to test the effect of the employment status on HRPB. Concerning the practice of physical exercise, our results seem to indicate that individuals who are unemployed have a lower probability of doing exercise if compared with those employed. We could advance as explanation for this result that the unemployed individual could have an increase in stress due to income reduction and reallocate his resources in what he feel has more urgent needs; or even, for free activities he did not feel prone to.

Testing the impact of chronic diseases was another goal. Unsurprisingly, we found that those who suffer from cancer or have had a stroke have a lower probability of doing exercise. The impact of having a stroke on the practice of exercise is higher than that of having cancer. Nordin et al.(2002) and Turp et al., (2000) find evidence that individuals with severe diseases will be less willing to do physical activity due to their limitations. We did not find statistical significance for the other variables used.

In the set of socioeconomic variables those who are married tend to do more physical exercise. We also observe from the results that those with a higher level of education, income, age present, on average, a lower likelihood of engaging in intense or moderate physical activities. A higher degree of education and income can be seen as a proxy for more demanding employment with further responsibilities, leaving less time for leisure activities. Concerning the covariate age, the literature suggests, as was explained before, a different pattern of preventive behaviour. These results are in line with previous studies suggesting that older people have less incentive for doing primary prevention since the pay-off period is shorter and the risks of getting diseases increase with age (Shureiqi et al., 2001). [46]

All the other variables did not have statistical significance, therefore are not an important determinant for explain the individual's decision to practice physical activity.

[46] See section 4.4.2

Table 4.5 - Bivariate probit for Secondary prevention for all sample

Dependent Variable - Secondary Prevention	Coef.	Std. Err	z
Health Variables			
High Blood Pressure	.134*	.051	2.64
Anxiety	-.130	.096	-1.35
Diabetes	-.032	.073	-0.43
Asthma	-.029	.092	-0.31
Cancer	-.017	.134	-0.13
Cerebral Haemorrhage	-.345**	.176	-1.96
Emphysema	-.317**	.126	-2.52
Stroke	.033	.158	0.21
Health Behaviours			
Smoker	-.046	.054	-0.86
BMI	.003	.005	0.57
Socioeconomic			
Single	-.122	.085	-1.43
Married	.123*	.063	1.97
Age	.002	.002	1.23
Male	-.068	.043	-1.58
Yearsschool	.033*	.006	5.56
Lnycp	.099*	.038	2.62
Unemployed	-.061	.110	-0.55
Retired	.031	.064	0.48
Cons	-2.484	.265	-9.39
rho	.017	.026	

Notes: Coefficients market with *, ** and*** are significant at 1, 5 and 10 per cent level, respectively.

The estimation results for the overall sample for the second dependent variable – preventive medical exams- show slight different results. Results are reported in table 4.5. Exploring the role of employment status in the simultaneous decision of doing prevention is one of our hypotheses. The results show that in this case employment status is not a predictor for the use of screening tests.

On the contrary, we do find evidence of the role of chronic diseases and of the socio economic factors in the utilization of secondary prevention. In the case of the last set of variables we did not find a so profound importance as for explaining the practice of exercise.

Concerning the decision to take part in secondary prevention (see table 4.5), individuals suffer from chronic diseases have different types of behaviour. People who suffer from high blood pressure have a high probability of doing prevention, but those who have had cerebral haemorrhages and emphysema have the opposite behaviour. High blood pressure is a disease but it is also a cause of other diseases. Individuals who suffer from this may be more aware of the perceived risk and be more willing to consume preventive care. In the latter cases, the illness has already occurred and the opportunity costs could increase, those with poorer health could have less time to receive treatments and screening tests given their physical limitations.

In what concern to the effect of marital status on the decision to take part in secondary prevention, we find that those who are married have a higher probability of doing secondary prevention. This could be because they have family responsibilities, which makes them more aware of the need to do prevention. The perception of risk for these individuals could be higher, or the partner could influence the decision. But again these are only hypotheses because our data do not permit a more deep analysis.

As for the level of education and income, as they increase so does the probability of doing preventive care. This is clearly stated in the literature (Kenkel, 1994; Lairson and et. al., 2005; Carrieri and Bilger (2013)): the higher the number of years in school the higher the probability of doing screening tests. In regards to the economic situation, higher income levels increase the probability of doing secondary prevention; witch is also in line with empirical results mention.

Moreover, on average, men are less likely to do prevention than women. This result is widely documented in the literature Kenkel (1994).

In the following paragraphs we make use of some of the results of our exploratory analysis presented in section 4.4.1. There is some evidence showing that age and gender influence preventive behaviour. Therefore, we split our sample into two sub-samples: the young adults, aged between 18 and 39, and the old adults, over 40. The estimation results are presented in tables 4.6 and 4.7 for the sub-sample 18-39 and in tables 4.8 and 4.9 for old adults.

Table 4.6 Bivariate probit for Physical activity age 18-39

Dependent variable - Physical Activity	Coef.	Std. Error	Z
Health Variables			
High blood pressure	.130	.115	1.13
Anxiety	-.365**	.176	-2.07
Diabetes	.0424	.198	0.21
Asthma	-.145	.114	-1.27
Cancer	.427	.363	1.18
Cerebral haemorrhage	.693	.577	1.20
Emphysema	-.192	.193	-1.00
Health Behaviours			
Smoker	.003	.055	0.05
BMI	-.002	.006	-0.33
Socioeconomic			
Single	-.200	.125	-1.59
Married	-.146	.121	-1.21
Age	.002	.005	0.50
Male	.127**	.053	2.39
Yearsschool	-.050*	.007	-6.72
Lnycp	-.133*	.047	-2.85
Unemployed	-.681*	.111	-6.12
Retired	-1.304*	.527	-2.47
_cons	1.681	.342	4.92

Notes: Coefficients market with *, ** and*** are significant at 1, 5 and 10 per cent level, respectively.

Regarding the sub-sample of young adults (table 4.6) and the decision to do exercise, we mainly found the same results. Again those who are unemployed have a lower probability of do exercise. From the set of chronic diseases only the anxiety variable presents a negative sign and statistical significance. Meaning that, these individuals have a lower probability of doing exercise. All the other variables of this set do not present statistical significance. Witch could be explained by the age of the individuals. It is uncommon for young adults to report

chronic diseases. In addition, men seem to do more exercise. The variables representing education and income present the same results regarding sign and statistical significance as for the all sample. Again these results reinforce the results for the whole sample: individuals who are unemployed have a lower probability of doing exercise than those who are employed.

Regarding the decision to take part in secondary prevention (Table 4.7), the results are mainly different for chronic diseases if we compare with the overall sample, and for the covariate gender.

Table 4.7 Bivariate probit for Secondary Prevention age 18-39

Secondary Prevention			
	Coef.	Std. Err	Z
Health Variables			
High Blood Pressure	-.065	.159	-0.41
Anxiety	-.091	.244	-0.37
Diabetes	.065	.260	0.25
Asthma	.039	.160	0.24
Cancer	.325	.390	0.83
Ch	-4.672	406.116	-0.01
Emphysema	-.406	.343	-1.18
Health Behaviours			
Smoker	-.003	.082	-0.04
BMI	-.007	.009	-0.73
Socioeconomic			
Single	-.289***	.170	-1.70
Married	-.087	.158	-0.55
Age	.008	.007	1.22
Male	-.191**	.079	-2.43
Yearsschool	.026**	.011	2.42
Lnycp	.135***	.071	1.89
Unemployed	-.163	.177	-0.92
Retired	.557	.622	0.90
Const	-2.438	.513	-4.76
rho	-.027	.045	

Notes: Coefficients market with *, ** and*** are significant at 1, 5 and 10 per cent level, respectively.

In this sub-sample the covariates that measure the impact of chronic diseases on secondary prevention does not present statistical significance. Again the reason could be the age of the individuals. In this case women tend to do more secondary prevention. This finding is in line with the literature (Kenkel, 1994). Individuals who are single have a lower probability of doing secondary prevention. The covariates education and income present the same sign than for the all sample. The remaining covariates do not present statistical significance.

The estimation results for the old adults present some differences from the global estimates regarding statistical significance of some covariates (table 4.8 and 4.9). It is our conviction that this result is linked with the characteristics of this sample.

Table 4.8 - Bivariate probit for Physical activity age >40

Dependent variable – Physical Activity			
	Coef.	**Std. Err**	**Z**
Health Variables			
High Blood Pressure	-.029	.040	-0.72
Anxiety	-.092	.076	-1.22
Diabetes	-.081	.056	-1.43
Asthma	.061	.080	0.76
Cancer	-.283*	.110	-2.57
Cerebral Haemorrhage	-.163	.119	-1.37
Emphysema	-.010	.087	-0.12
Stroke	-.333*	.125	-2.65
Health Behaviours			
Smoker	-.093***	.052	-1.79
BMI	-.008***	.004	-1.77
Socioeconomic			
Single	-.011	.080	-0.14
Married	.092***	.049	1.89
Age	-.020*	.002	-9.91
Male	-.016	.039	-0.42
Yearsschool	-.045*	.005	-8.30
Lnycp	-.061***	.032	-1.86
Unemployed	-.442*	.095	-4.65

Retired	-.348*	.050	-7.00
_cons	2.23	.256	8.68

Notes: Coefficients market with *, ** and*** are significant at 1, 5 and 10 per cent level, respectively.

We find the same results for those who are unemployed. Unemployed older adults still have a lower probability of practice exercise and the same is true for those who are retired.

The health condition having cancer or a stroke decreases the probability of doing physical exercise for older adults.

The main differences for this sub sample concern to the group of variables included to reflect health behaviours. The variable that account if the individual is a smoker also presents statistical significance and the same is true for the BMI covariate. Smokers and obese individuals have a lower probability of doing exercise. It is possible that the individual's characteristics could explain why some individuals are more willing to engage in good health habits or not. We find a similar relation that Ruhm (2000) but we do not control for exogenous shocks. Ruhm (2000) presents preliminary evidence that smoking and body weight decline while physical activity increases and diets improve during recessions.

Socioeconomic variables present the same sign and statistical significance than for the all sample.

For the old adults sample regarding the consumption of secondary prevention the results (table4.9) are similar to those obtain for the all sample.

Table 4.9 - Bivariate probit for Secondary prevention age >40

Secondary prevention			
	Coef.	**Std. Err**	**Z**
Health Variables			
High Blood Pressure	.152*	.054	2.82
Anxiety	-.081	.102	-0.80
Diabetes	-.044	.076	-0.58
Asthma	-.067	.113	-0.59
Cancer	-.014	.140	-0.10
Ch	-.335***	.178	-1.89
Emphysema	-.295**	.136	-2.18

Stroke	.028	.158	0.18
Health Behaviours			
Smoker	-.112	.072	-1.55
BMI	.007	.006	1.18
Socioeconomic			
Single	-.067	.117	-0.58
Married	.097	.068	1.42
Age	.0003	.003	0.13
Male	.012	.052	0.23
Yearsschool	.032*	.007	4.60
Lnycp	.085***	.044	1.92
Unemployed	-.047	.139	-0.34
Retired	.046	.068	0.68
Const.	-2.379	.344	-6.91
rho	.040	.031	

Notes: Coefficients market with *, ** and*** are significant at 1, 5 and 10 per cent level, respectively.

Overall the results do not differ much across the sub-samples and the main differences found are related to the subsample characteristics, namely the age of the individuals.

4.6 Conclusion and discussion

Our main focus in the current chapter was to understand if the decision to engage in HPRB depends on the unemployment status of the individual. Our main results can be summarized as follows.

First, it was clear that for all samples those who were unemployed have a lower probability of doing exercise than those who are employed, and that the state of employment is not an explanatory covariate for secondary prevention. We could argue like Brenner and Mooney (1983); Catalano and Dooley (1983); Fenwick and Tausig, (1994) stress-related diseases and mental health associated with economic insecurity could withdraw individuals from the practice of exercise. In the case of secondary prevention, since this screening tests are free of charges for the target group or have a small co-payment, the potential loss of income is not determinant for not engage in a preventive behaviour.

Secondly, we found evidence showing that chronic diseases are a factor dissuading individuals from doing more physical activity and, on the contrary, more secondary prevention. This can be explained by the fact that the marginal effect of a new illness can be greater for those who are already in poor initial health; individuals can do more screening tests but will have less propensity to do exercise if they feel ill. Our results are in line with some of the literature (Türp et al., (2000) and Nordin et al., (2002)).

Thirdly, we try to identify characteristics that permit infer why some individuals have healthy behaviour and do more prevention and others do not. For that we control for some health behaviours. In a global analysis we did not find a profile for individuals who clearly have a strategic healthy behaviour. In this field the only result that worse to mention is for smokers and obese individuals. Smokers and obese individuals have a lower probability of doing physical activity.

Fourthly, it is commonly assumed in the literature that those with a higher level of education may be better informed about the potential benefits of prevention. Thus, they do more screening tests. We found similar results regarding the literature for secondary prevention, but the opposite for physical activity, probably because these individuals recognize a higher marginal benefit from screening tests than from exercise. In addition, the risk of getting diseases such as cancer increases with age and older individuals do less exercise and invest less in their health capital. In this case age acts only as a risk factor.

We do not found a clear evidence that those who are married or single do more exercise and secondary prevention. The relation between gender and prevention was only capture in the sub-sample of young adults. We find that men do more exercise than women but on the contrary do less screening tests.

Those who have high incomes have a lower probability of doing physical activity, but have a higher probability of doing more screening exams. The opportunity cost of doing exercise could be higher for those who are more able to afford any out-of-pocket costs.

Concluding, we did not find the existence of a significant relation in the decision of undertake the two different types of prevention. We find different profiles between those who make primary and secondary prevention.

This study enclosed some limitations. In Portugal, health surveys are not very detailed. This is especially true for national representative surveys. This made it almost compulsory for us to use the fourth wave of the Portuguese health survey. Despite the richness of the data set, the way the survey is constructed prevents the use of panel data. Our study is cross-sectional and we used data from 2005, which was the most recent year we were able to work with. However, we are aware that since 2005 some behaviour may already have changed among the Portuguese population. We think that the results could be different with a more recent survey that could reflect better the actual rate of unemployment.

Concerning the decision making process of taking preventive health behaviour, the econometric approach requires a dynamic treatment, given that behaviour changes over time. Again, the way the survey is constructed does not permit the use of certain variables that in our opinion could be very useful. For instance, it would have been interesting to investigate the impact of hereditary factors such as the history of serious family illness. Individuals who have a history of serious family illness are more likely to be aware of their own health, and as a consequence may live a healthier lifestyle and use more preventive medical care.

It will be very interesting to extend this research to a dynamic panel in the future, as well as using variables usually linked to sociological models.

Despite the above mentioned limitations, we hope our study sheds light on certain aspects of the demand for preventive care and preventive health behaviour, in particular its relationship with individual's employment status.

The findings may be categorized into four areas. The first one claims that individuals who suffer from chronic diseases are less likely to do prevention. Secondly the effects of socioeconomic variables are diversed, but it is possible to identify different profiles. It is clear that different individual profiles exist regarding primary prevention and secondary prevention. Third, the participation in prevention activities varies between different procedures for the same individuals. Individuals seem to value more the practice of secondary prevention than that of primary prevention. Finally, employment status has an ambiguous role. Have a negative impact on the practice of exercise and is not an important covariate in explaining the decision for secondary prevention.

Unemployed do less exercise and could have more free time to do it. So it will be important to better understand the reasons behind this behaviour, looking closest this group of individuals.

Therefore, it will be possible to design health promoting programs for the practical of exercise focus on this group, mitigating the effect of unemployment in the individual's health. The same could be true for secondary prevention.

In fact is somehow surprising that this individuals use less secondary prevention, once they do not have to pay the co-payments. Nevertheless, could be a result of the survey used, since we used a sub-sample of only the individuals that have at least one appoint with the GP in the last three months.

The policy message, which can be drawn from this study, can be twofold. First, macroeconomic policies in Portugal have a strong role to play in improving the propensity of individuals to engage in HPRB. Policies that reduce poverty, unemployment, and improve education could be vital for enhancing the demand for preventive health care. Secondly, we were able to trace a simple profile of the individuals. Therefore, it is possible to focus health policies on particular groups, and reinforce the role of the GP as a fundamental provider of information regarding the benefits of preventive behaviour.

CHAPTER 5

CONCLUSION

5.1 General conclusion

The main purpose of this dissertation was to study individuals' prevention behaviour while looking at its three dimensions (primary, secondary and tertiary prevention) and at the different targets and vehicles that can potentiate it.

We believe this research has added further insight into focused and efficient health policy. First, it has bolstered Portuguese research in the field of prevention. Secondly, it has increased the knowledge of individuals' perceptions of prevention, attitudes and behaviour. Thirdly, only part of this research was applied to Portugal, we were also able to make an international comparison between Portugal, Spain and Greece in chapter 2. This research also developed an instrument to measure the WTP for primary prevention, and rethought the role of information in the individual's decision-making process regarding the demand for prevention. We were able to identify the primary vehicle of information used by students and we discovered that their academic background was also of importance. We compared the differences between Portuguese, Greek and Spanish students in terms of perceptions, attitudes and behaviour when confronted by a public health crisis. We looked for a change in individual behaviour in the post- pandemic period, for this a follow-up study was conducted.

We analysed the mediating role of prevention on illness absenteeism, while considering the role of primary and secondary prevention on this relation. We revealed the important role of some health behaviour variables. In addition, we studied the moderate role of organizational variables (activity sector, private or public sector) and individual variables (gender, age and schooling) on the relation between variables.

We modelled the simultaneous decision of doing primary and secondary prevention while controlling for the individual's health status. We also analysed the importance of healthy

behaviour and studied the link between prevention and socio-economic status, with special focus on employment status.

In short, our first topic of interest was the relationship between information and the WTP for a vaccine against the H1N1 virus in Greece, Portugal and Spain. For this, an appropriate set of questions using the CV method was adopted. Intuition makes us think that if individuals are confronted with the same information, then they will show similar behaviour. However, this is not the case, because the way in which the individual processes information is more important than the quantity of information at his disposal.

We estimated the same model using two different approaches. First, a Tobit and then a Two-part model, the results were quite different. In spite of the differences, we feel confident assuring that the public's exposure to messages, topics, and information about health care differs depending on which media source the individual turns to most often. This in turn then influences the individual's WTP for preventive health care.

We must stress the impact of the media on the willingness to pay, students that used TV as their primary channel of information had more WTP than those who used the internet or read newspapers. We think the reason for this could lie in the individual's subjective memory. The information received by the formal education sector (course in which the students are enrolled) improves consumer decision-making and contributes to market efficiency.

In the third chapter the main objective was to appraise the effect of preventive health care on the illness absenteeism rate of Portuguese workers. The idea being that those who consume more preventive care will miss work less (for health reasons). The outcomes are not clear, maybe because those who suffer from chronic diseases need to miss more, not only because of their condition but because prevention is needed as a way of preventing or reducing the incidence of others diseases.

Our main findings are: 1) health variables seem to be an important determinant of Portuguese illness absenteeism, especially chronic diseases; 2) socioeconomic variables like age and income are also important determinants; 3) health behaviour could acts as preventive behaviour for absenteeism; 4) social security system also helps to explain absenteeism.

We have not found clear evidence showing that the direct measures of preventive care used affect the decision of being absent for health reasons. Nevertheless, the complementary

measures of preventive behaviour show us the importance of prevention on illness absenteeism.

Finally, and taking into account the actual context of our country, it is extremely relevant to understand if this socio-economic environment could affect the demand for preventive health care. We model the decision to do primary and secondary prevention by estimating a bivariate probit model, which assures the robustness of the results since control for the endogeneity of the variables

The results are somewhat cofounding regarding the literature and what we expected in terms of health behaviour. This could be for reasons other than cultural, or maybe because the data is from 2005. We believe that if the survey presented more actual data, the results could be slightly different.

We found different profiles for those who do primary prevention and those who do secondary prevention. For this reason, it is important to focus policy on the different types of identified profiles.

5.2 Limitations and future research

The main limitation of the study is related to the difficulty of obtaining quality data. This is because the existent databases in this field are usually cross sectional and have a broad goal. In some cases making impossible to focus our analysis on the target group's preview. We tried to get beyond this limitation by designing an instrument for collecting our own data. However, it was impossible to do this for all the studies given that it is time consuming and, more importantly, highly expensive.

Future research is needed; for instance, it is important to use more accurate methodologies and some questions also deserve a more profound study.

We think that more research in the field of health behaviour is needed. Because in our opinion, new behavioural strategies, which are more, focused on education and health promotion, are needed for the future. Some work has already been done, but without scientific support and lacking focus. Of course police makers can impose behavioural changes, but the best solution is increasing individuals' awareness regarding the need to change lifestyle

behaviour. It is important that policy makers design policies that take into account individuals' accountability.

5.3 Final note

Despite some limitations, we feel we have contributed to Prevention research in Portugal with theoretical and practical implications.

We were able to give some insight into individuals' behaviour in a public health crisis (2009-2010), as well as into the role of the main vehicle of information. This will enable more accurate policies that focus on the quality of the information delivered to some groups (which have the potential to be outbreakers). We added information on the changes in perception after the health crises and pointed out some of the causes for this response.

We looked at the mediating role of prevention on the illness absenteeism rate in Portugal, while pointing out that prevention could be very useful for employees and employers. Some programmes devoted to promoting health in the work place already exist and we believe our findings could improve the efficiency and efficacy of these programmes (by focusing more on changes in lifestyle).

Finally, our results shed light on some macroeconomic aspects and allow policy concerning simultaneous decisions of primary and secondary prevention to be put in place.

REFERENCES

Åhs, A. M. H. and. Westerling, R. (2006), "Health care utilization among persons who are unemployed or outside the labour force." *Health Policy*, Vol. 78, pp. 178-193.

Akan, H. , Gurol, Y., Izbirak, G., Ozdatlı, S., Yilmaz, G., Vitrinel, A. and Hayran, O. (2010), "Knowledge and attitudes of university students toward pandemic influenza: a cross-sectional study from Turkey", *BMC Public Health*, Vol. 10, pp. 413.

Allebeck, P. and Mastekaasa, A. (2004), "Chapter 3. Causes of sickness absence: research approache and explanatory models", *Scandinavian Journal of Public Health*, Vol. 32, Suppl. 63, pp. 36-43.

Allen, S. G. (1981), "An empirical model of work attendance", *The Review of Economics and Statistics*, Vol. 63, No. 1, p. 77.

Allen, S. G. (1984), "Trade unions, absenteeism and exit voice", *Industrial and Labor Relations Review*, Vol. 37, pp. 331-345.

Anker, A. E., Reinhart, A., Feeley, T. (2011), "Health information seeking: A review of measures and methods", *Patient Education and Counseling*, Vol. 82, pp. 346–354.

Arinaminpathy N. and Dye C. (2010), "Health in financial crises: economic recession and tuberculosis in Central and Eastern Europe.", *Journal of the Royal Society Interface*, Vol. 7, No. 52, pp. 1559-1569.

Arrow, K. J. (1963), "Uncertainty and the welfare economics of medical care", *The American Economic Review*, Vol. 53, No. 5, pp. 941-973.

Bago d'Uva, T. (2005), "Latent class models for utilisation of health care", *HEDGWorking paper 05/01*. University of York.

Baker, D. B. (1985), "The study of stress at work", *Annual Review of Public Health*, Vol. 6, pp. 367–381.

Ballard-Reisch DS. (1990), "A model of participative decision making for physician–patient interaction", *Health Commun*, Vol.2, pp. 91–4.

Barham, C. and Begum, N. (2005), "Sickness absence from work in the UK", *National Statistcs Feature*, Office for National Statistics, Labour Market Trends, April, pp. 149-158.

Barmby, T. A. and Larguem, M. (2009), "Coughs and sneezes spread diseases: An empirical study of absenteeism and infectious illness", *Journal of Health Economics*, Vol. 28, No. 5, pp. 1012-1017.

Barmby, T. A. and Stephan, G. (2000), "Worker absenteeism: Why firm size may matter", *The Manchester School*, Vol. 68. No. 5, pp. 568-577.

Barmby, T. A., Orme, C. D. and Treble, J. (1991), "Worker absenteeism: an analysis using microdata", *The Economic Journal*, Vol. 101, No.405, pp. 214-229.

Basch, EM., Thaler ,HT., Shi, W., Yakren, S. and Schrag, D. (2004), "Use of information resources by patients with cancer and their companions", *Cancer*, Vol. 100, No 11, pp 2476–83.

Basu, A. and Rathouz, P. (2005), "Estimating marginal and incremental effects on health outcomes using flexible link and variance function models", *Biostatistics*, Vol. 6, No. 1, pp. 93–109.

Basu, A., Manning, W. G. and Mullahy, J. (2004), "Comparing alternative model: log vs Cox proportional hazard?", *Health Economics*, Vol. 13, No. 8, pp. 749–765.

Blough, D. K., Madden, C. W. and Hornbrook, M. C. (1999), "Modeling risk using generalized linear models", *Journal of Health Economics*, Vol. 18, pp. 153–171.

Bradley, S., Green, C. and Leeves, G. (2007), "Employment contracts and effort: Why do temporary workers take less absence?", Working paper 2007/026, Lancaster University Management School.

Brenner, M. H. and Mooney, A. (1983), "Unemployment and health in the context of economic change", *Social Science Medicine*, Vol. 17, No. 16, pp. 1125–1238.

Brodie, M., Brady, L. A. and Altman, D. E. (1998), “Media coverage of managed care: Is there a negative bias?”, *Health Affairs*, Vol. 17, pp. 9–25.

Brown, S. and Sessions, J. G. (1996), “The economics of absence: theory and evidence”, *Journal of Economic Surveys*, Vol. 10, No. 1, p. 23-53.

Cameron, A. C. and Trivedi, P. K. (1996), "Count data models for financial data.", *Handbook of Statistics.* G. S. Maddala and C. R. Rao, pp. 363-391.

Cameron, A. C. and Trivedi, P. K. (1998), *Regression analysis of Count data.*, New York, NY, Cambridge University Press.

Cameron, A. C. and Trivedi, P. K. (2005), *Microeconometrics : methods and applications.* New York, NY, Cambridge University Press.

Cameron, A. C. and Triverdi, P.K. (2009), *Microeconometrics Using Stata*, Stata Press, College Station, TX.

Cantoni, Eva, and Ronchetti, Elvezio. (2006),“A robust approach for skewed and heavy-tailed outcomes in the analysis of health care expenditures”. *Journal of Health Economics,* Vol.25, No. 2, pp. 198-213.

Carney, N., Greenlick, M. R., Austin, D. F., Nygren, P., Hibbard, J. H., Helfand, M., Davies, P., Davis-O’Reilly, C., Hacket, P. A., Kunkel, L. and Krages, K.P. (2000), *Assessment of Self-care Manuals*, Oregon Health Sciences University, Portland, OR.

Carrieri, V. and Bilger, M. (2013). "Preventive care: underused even when free. Is there something else at work?", *Applied Economics*, Vol. 45, pp. 239-253.

Catalano, R. and Dooley, D. (1983), “Health effects of economic instability: a test of the economic stress hypothesis”, *Journal of Health and Social Behavior,* Vol. 24, March, pp. 46–60.

Centre for Disease Control and Prevention (CDC) in the United States. http://www.cdc.gov/

Chatterji, M. and Tilley, C. J. (2002), “Sickness, absenteeism, presenteeism and sick pay”, *Oxford Economic Papers*, Vol. 54, No. 4, pp. 669-87.

Chen, K. and Lange, F. (2008), "Education, information, and improved health: Evidence from breast cancer screening", IZA Discussion Papers 3548, Institute for the Study of Labor (IZA).

Chestnut, L. G., Keller, L. R., Lambert, W. E. and Rowe, R. D. (1996), "Measuring heart patients' willingness to pay for changes in angina symptoms", *Medical Decision Making*, Vol. 16, pp. 65–77.

Chiu, L., Tang, K. Y., Shyu, W. C. and Chang T. P. (1999), "The willingness of families caring for victims of stroke to pay for in-home respite care – results of a pilot study in Taiwan", *Health Policy*, Vol. 46, pp. 239–254.

Chor, J. S., Ngai, K. L., Goggins, W.B., Wong, M.C., Wong, S.Y., Lee, N. et al. (2009), "Willingness of Hong Kong healthcare workers to accept pre-pandemic influenza vaccination at different WHO alert levels: two questionnaire surveys", *British Medical Journal*, Vol. 339, p. b3391.

Cline, R. J. W. and Haynes, K. M. (2001), "Consumer health information seeking on the Internet: The state of the art." *Health Education Research*, Vol.16, No 6, pp.671–692.

Compact Oxford Dictionary, http://www.askoxford.com.

Courtemanche, C. (2009), "Longer hours and larger waistlines? The relationship between work hours and obesity", *Forum of Health Economics and Policy*, Vol. 12, No. 2, p.2.

Cragg, J. G. (1971). "Some Statistical Models for Limited Dependent Variables with Application to Demand for Durable Goods", *Econometrica*, Vol.39, No.5, pp. 829-844.

Currie, J. and Madrian, B. (1999), *A Health, Health Insurance, and the Labor Market*, Handbook of Labor Economics, Volume 3, North-Holland, Amsterdam.

Damman, O. C., Hendriks, M., Rademakers,J. , Delnoij, D. M. and Groenewegen, P. P. (2009), "How do healthcare consumers process and evaluate comparative healthcare information? A qualitative study using cognitive interviews", *BMC Public Health*, Vol. 9, pp. 423.

Deb, P. and Trivedi, P. K. (1997), “Demand for medical care by the elderly: A finite mixture approach”, *Journal of Applied Econometrics* Vol.12, No 3, pp. 313-336.

Deb, P. and Trivedi, P. K. (2002), “The structure of demand for health care: latent class versus two-part models”, *Journal of Health Economics* Vol. 21, No 4, pp. 601-625.

Deb, P. and Trivedi, P. K. (2009), "Provider networks and primary-care signups: do they restrict the use of medical services?," *Health Economics*, vol. 18(12), pp. 1361-1380. John Wiley & Sons, Ltd.

Dehejia, R. and Lleras-Muney, A. (2004), “Booms, busts, and babies’ health”, *Quarterly Journal of Economics*, Vol. 119, No. 3, pp. 1091-1130.

Delgado, M. A. and Kniesner, T. J. (1997), “Count data models with variance of unknown form: An application to a hedonic model of worker absenteeism”, *Review of Economics and Statistics*, Vol. 79, No.1, p. 41-49.

Diener, A., O’Brien, B. J. and Gafni, A. (1998), “Health care contingent valuation studies: A review and classification of the literature”, *Health Economics*, Vol. 7, pp. 313–326.

Dionne, G. and Dostie, B. (2007), “New Evidence on the determinants of absenteeism using linked employer-employee data”, *Industrial and Labor Relations Review*, Vol. 61, No.1, pp. 108-120.

Doebbeling, B., Edmond, M. B., Davis, C. S., Woodin, J. R. and Zeitler, R. R. (1997), “Influenza vaccination of health care workers: evaluation of factors that are Important in acceptance”, *Preventive Medicine*, Vol. 26, pp. 68–77.

Donaldson, C., Farrar, S., Mapp, T., Walker, A. and Macphee, S. (1997), "Assessing community values in health care: is the ‘willingness to pay’ method feasible ?", *Health Care Analysi*, VOL. 5, pp. 7–29.

Duan, N. (1983), “Smearing estimate: A nonparametric retransformation method”, *Journal of the American Statistical Association*, Vol. 78, pp. 605–610.

Duan, N., Manning, W. G., Morris, C and Newhouse, J. (1983), “A comparison of alternative models for the demand for medical care”, *Journal of Business & Economic Statistics*, Vol.1, No. 2, pp. 115-126.

Dunn, L. F., and Youngblood, S. A. (1986), "Absenteeism as a mechanism for approaching an optimal labor market equilibrium: an empirical study", *Review of Economics and Statistics*, Vol. 68, No.4, pp. 668–74.

Ehrlich, I. and Becker, G. (1972), "Market insurance, self-insurance and self-protection", *Journal of Political Economy*, Vol. 80, No. 4, pp. 623-648.

Ehrlich, I. and Chuma, H. (1990), "A model of the demand for longevity and the value of life extension", *Journal of Political Economy*, Vol. 98, No. 4, pp. 761-782.

Emanuel, E. J., Emanuel, L. L. (1991), "Four models of the physician–patient relationship", *Journal of the American Medical Association*, Vol. 267, No. 16, pp. 2221–6.

Epstein, R. M., Alper, B. S., Quill, T. E. (2004), "Communicating evidence for participatory decision making", *Journal of the American Medical Association*, Vol. 291, pp. 2359–66.

Ercolani, M. (2006), "UK Employees Sickness Absence: 1984-2005." Health at Work Project Working Paper, Discussion Papers 06-02, Department of Economics, University of Birmingham.

European Foundation for the Improvement of Living and Working Conditions (2005), The Fourth European Survey on Working Conditions.

Evans, W. and Graham, J. D. (1988), "Traffic safety and the business cycle", *Alcohol, Drugs, and Driving*, Vol. 4, No. 1, pp. 31-38.

Fast, J., Vosburgh, R. E. and Frisbee, W. R. (1989), "The effects of consumer education on consumer search", *Journal of Consumer Affairs*, Vol. 23, pp. 65–90.

Fenn, P. and Ashby, S. (2004), "Workplace risk, establishment size and union density", *British Journal of Industrial Relations*, Vol. 43, No. 3, pp. 461-480.

Fenwick, R. and Tausig, M. (1994), "The macroeconomic context of job stress", *Journal of Health and Social Behavior*, Vol. 35, No. 3, pp. 266–282.

Fox, S. and Rainie, L. (2002). Vital decisions: How Internet users decide what information to trust when they or their loved ones are sick. Technical Report from the Pew Internet and American Life Project: Washington, DC: Pew Research Center.

Frank, R. (2004), "Behavioral economics and health economics", paper presented at the Yrjö Jahnsson Foundation 50th Anniversary Conference on Economic Institutions and Behavioural Economics, June 22-24, 2004, available at: http://www.xn--yrjjahnssoninsti-8nba34apa.fi/tiedotefiles/Frank.pdf (accessed 10 December 2012).

Freeman, D. G. (1999), "A note on 'economic conditions and alcohol problems'", *Journal of Health Economics*, Vol. 18, No. 5, pp. 661-670.

Frick, B. and Malo, M. A. (2008), "Labor Market Institutions and Individual Absenteeism in the European Union: The Relative Importance of Sickness Benefit Systems and Employment Protection Legislation.", *Industrial Relations: A Journal of Economy and Society*, Vol. 47, Issue 4, pp. 505-529.

Galarce, E. M., Minsky, S. and Viswanath, K. (2011), "Socioeconomic status, demographics, beliefs and A(H1N1) vaccine uptake in the United States", *Vaccine*, Vol. 29, No. 32, pp. 5284-5289.

Gilleskie, D. (2010),"Work Absences and Doctor Visits during an Illness Episode: The Differential Role of Preferences, Production, and Policies among Men and Women." Journal of Econometrics 156(1), 2010, p. 148-163.

Gourieroux, C. and M. Visser (1997), "A count data model with unobserved heterogeneity", *Journal of Econometrics*, Vol.79, Nº2, pp. 247-268.

Greene, W. H. (2003), *Econometric analysis*. Upper Saddle River, N.J., [Great Britain], Prentice Hall.

Grossman, M. (1972), "On the concept of health capital and the demand for health", *Journal of Political Economy*, Vol. 82, pp. 223-255.

Grossman, M. (2000), "The human capital model", in Culyer, A.J. and Newhouse, J. P. (Eds.), *Handbook of Health Economics*, volume 1A, Elsevier, New York.

Gurmu, S. and Trivedi, P. K. (1992), "Overdispersion Tests for Truncated Poisson Regression-Models.", *Journal of Econometrics*, Vol. 54, pp. 347-370.

Gurmu, S. and. Trivedi, P. K (1996), "Excess zeros in count models for recreational trips.", *Journal of Business & Economic Statistics*, Vol.14, No4, pp. 469-477.

Hadi, A. S. (1992), "A modification of a method for the detection of outliers in multivariate samples", *Journal of the Royal Statistical Society Series (B),* Vol. 56, pp. 393–396.

Hadi, A. S. (1994), "Identifying multiple outliers in multivariate data", *Journal of the Royal Statistical Society Series (B*), Vol. 54, No. 3, pp. 761–771.

Hammitt, J. K. and Zhou, Y. (2006), "The economic value of air-pollution-related health risks in China: A contingent valuation study", *Environmental and Resource Economics*, Vol. 33, No. 3, pp. 399–423.

Hardyman, R., Hardy, P., Brodie, J. and Stephens, R.(2005), "It's good to talk: comparison of a telephone helpline and website for cancer information", *Patient Educ Couns*, Vol.57, No.3, pp 315–320.

Hernanz, V. and Toharia, L. (2006), "Do temporary contracts increase work accidents? A microeconometric comparison between Italy and Spain", *Labour*, Vol. 20, No. 3, pp. 475-504.

Hiatt, R., Pasick, R.J., Stewart, S., Bloom, J., Davis, P., Gardiner, P., Johnston, M., Luce, J., Schorr, K., Brunner, W. and Stroud, F. (2001), "Community-Based Cancer Screening for Underserved Women: Design and Baseline Findings from the Breast and Cervical Cancer Intervention Study", *Preventive Medicine*, Vol. 33, No3, pp.190-203.

Hibbard, J. H. and Peters, E. (2003), "Supporting informed consumer health care decisions: Data presentation approaches that facilitate the use of information in Choice", *Annual Review of Public Health*, 24, 413–33.

Hsieh, C.-R. and Lin, S.-J. (1997), "Health information and the demand for preventive care among the elderly in Taiwan", *Journal of Human Resources*, Vol. 32, pp. 308–333.

Huver, B., Richard, S., Vaneecloo, N., Delclite, T. and Bierla, I. (2012), "Sick but at Work. An Econometric Approach to Presenteeism", 15th IZA European Summer School in Labor Economics Buch am Ammersee – April 23-29.

Instituto Nacional de Estatístia [INE] (2011), "Desemprego, por género e grupo etário [Unemployment by gender and age]", *National Statistics Institute* (available at Pordata), Lisbon.

Jadad, A. R. and Gagliardi, A. (1998), "Rating health information on the Internet: Navigating to knowledge or to Babel?", *The Journal of the American Medical Association*, Vol. 279, pp. 611–614.

Johannesson, M. and Jonsson, B. (1991), "Economic evaluation in health care: Is there a role for cost–benefit analysis?", *Health Policy*, Vol. 17, pp. 1–23.

Johansson, P. and Palme M. (1993), "The effect of economic incentives on worker absenteeism: An empirical study using Swedish micro data", *Umefi Economic Studies 314*, University of Ume.

Johansson, P. and Palme, M. (1996), "Do economic incentives affect work absence? Empirical evidence using Swedish micro data", *Journal of Public Economics*, Vol. 59, No. 2, pp. 195-218.

Johansson, P. and Palme, M. (2005), "Moral hazard and sickness insurance", *Journal of Public Economics*, Vol. 89, No. 9-10, pp. 1879-1890.

Jones, A. M. (2000), *Health econometrics. Handbook of Health Economics.* A. J. Culyer and J. P. Newhouse. New York, Elsevier, Vol. 1, pp. 265-344.

Jones, A. M. and O'Donnell, O. (2002), *Econometric analysis of health data*. Chichester, Wiley.

Karasek, R. A. and Theorell, T. (1990), *Healthy Work: Stress, Productivity and the Reconstruction of Working Life*, Basic Books, New York.

Keeler, E. B., Manning, W. G., Wells, K. B. (1988), "The Demand for Episodes of Mental-Health Services", *Journal of Health Economics*, Vol. 7, No. 4, pp. 369-392.

Kenkel, D. S. (1990), “Consumer health information and the demand for medical care”, *Review of Economics & Statistics*, Vol. 72, pp. 587–595.

Kenkel, D. S. (1991), “Health behavior, health knowledge, and schooling”, *Journal of Political Economy*, Vol. 99, pp. 287–305.

Kenkel, D. S. (1994), “The demand for preventive medical care”, *Apllied Economics*, Vol. 26(4), pp. 313–325.

Kenkel, D. S. (2000), “Prevention”, in Culyer, A. J. and Newhouse, J.P. (Eds.), *Handbook of Health Economics*, Vol. 1, North Holland, pp. 1675-1720.

Kogut, B. and Zander, U. (1992), “Knowledge of the firm. Combinative capabilities and the replication of technology”, *Organization Science*, Vol.3, pp. 383-397.

Lairson, D. R., Chan, W, Newmark, Georgina, R. (2005), "Determinants of the demand for breast cancer screening among women veterans in the United States," *Social Science & Medicine* Vol. 61, No 7, pp. 1618-1627.

Lambert, D. (1992), “Zero-inflated poisson regression, with an application to defects in manufacturing”, *Technometrics* Vol. 34, No1, pp. 1–14.

Lange, F. (2011), “The role of education in complex health decisions: Evidence from cancer screening”, *Journal of Health Economics*, Vol. 30, pp. 43–54.

Lau, J. T., Yeung, N. C., Choi, K. C., Cheng, M. Y., Tsui, H. Y. and Griffiths, S. (2009), “Acceptability of A/H1N1 vaccination during pandemic phase of influenza A/H1N1 in Hong Kong: Population based cross sectional survey”, *British Medical Journal*, Vol. 339, p. b4164.

Leigh, J. P. (1981), “The effects of union membership on absence from work due to illness” *Journal of Labor Research*, Vol. II, No. 2, pp. 329-336.

Leigh, J. P. (1991), “Employee and job attributes as predictors of absenteeism in a national sample of workers: The importance of health and dangerous working conditions”, *Social Science and Medicine*, Vol. 33, No. 2, pp. 127-137.

Leigh, J.P. (1989), "Specific illnesses, injuries, and job hazards associated with absenteeism." *Journal of Occupational Medicine*, Vol. 31, pp. 792–792.

Lurås, H. (2001), "A healthy lifestyle: the product of opportunities and preferences, an explorative study", Working Paper 2011:11, University of Oslo Health Economics Research Programme.

Manning, W. G. (1998), "The logged dependent variable, heteroscedasticity, and the retransformation problem", *Journal of Health Economics*, Vol. 17, pp. 283–295.

Manning, W. G. and Marquis, M. S. (1996), "Health insurance: The tradeoff between risk pooling and moral hazard", *Journal of Health Economics*, Vol. 15, No. 5, pp. 609-639.

Manning, W. G. and Mullahy, J. (2001), "Estimating log models: to transform or not to transform?", *Journal of Health Economics*, Vol. 20, No 4, pp. 461–494.

Manning, W. G.(2010) materiais fornecidos no curso "Econometric Analysis in Healthcare Demand", no âmbito do Programa International Doctoral Courses in Health Economics and Policy - Swiss School of Public Health - que decorreu de 22 a 26 de Março de 2010 na Universidade de Genebra.

Manning, W. G., Basu, A. and Mullahy, J. (2005), "Generalized modeling approaches to risk adjustment of skewed outcomes data", *Journal of Health Economics*, Vol. 24, No 3, pp. 465–488.

Manning, W. G., Newhouse, J., Duan, N., Keeler, E., Leibowitz, A. and Marques, S. (1987), "Health insurance and the demand for medical care: Evidence from a randomized experiment", RAND Corporation (Pub. no. R-3476-HHS), Santa Monica. Abridged version published in *The American Economic Review*, Vol. 77, pp. 251-277.

Maurer, J., Uscher-Pines, L. and Harris, K. (2010), "Perceived safety and value of seasonal and 2009 influenza A(H1N1) vaccinations: First results from a nationally representative survey", *Preventive Medicine*, Vol. 51, No.2, pp. 185-187.

Ministry of Solidarity, Employment and Social Security (2007). "Social Report 2007" Portugal.

Mitchell, R. C. and Carson, R. T. (1989) *Using Surveys to value public goods: the contingent valuation method*, Resources for the Future, USA.

Montgomery, H. and Svenson, O. (1989), *Process and Structure in Human Decision Making*, John Wiley & Sons, Oxford, England.

Mullahy, J. (1997), "Heterogeneity, excess zeros, and the structure of count data models." *Journal of Applied Econometrics*, Vol.12, pp. 337-350.

Mullahy, J. (1998), "Much ado about two: reconsidering retransformation and the two-part model in health econometrics", *Journal of Health Economics*, Vol. 17, pp. 247–281.

Murray, E., Lo, B., Pollack, L., Donelan, K., Catania, J., White, M., Zapert, K., and Turner, R. (2003), "The impact of health information on the Internet on the physician-patient relationship: Patient perceptions.", *Archives of Internal Medicine*, Vol.163, pp. 1727–1734.

Naing, C. and Tan, R. Y. P. (2011), "Knowledge about the pandemic influenza A (H1N1) and willingness to accept vaccination: a cross-sectional survey", *Journal of Public Health*, Vol. 19, No. 6, pp. 511–516.

Nelson, D., Kreps, G. L., Hesse, B. W., Croyle, R. T., Willis, G., Arora, N.K.(2004), "The Health Information National Trends Survey (HINTS): Development, Design, and Dissemination", *Journal of Health Communication*, Vol. 9, pp. 443–460.

Neuhauser, Linda and Kreps, G.L. (2003), "Rethinking Communication in the E-health Era", *Journal of Health Psychology*, Vol. 8, No1, pp 7–23.

Neumann, P. J., Cohen, J. T., Hammitt, J. K., Concannon, T. W., Auerbach, H., Fang, C. and Kent, D. (2012), "Willingness-to-pay for predictive tests with no immediate treatment implications: a survey of US residents", *Health Economics*, Vol. 21, pp. 238–251.

Niederdeppe, J., Hornik, R., Kelly, B., Frosch, D., Romantan, A., Stevens, R., Barg, F., Weiner, J., and Schwarz, S. (2007), "Examining the dimensions of cancer-related information scanning and seeking behavior.", *Health Communication*, Vol. 22, pp. 153–167.

Nonaka, I. and Takeuchi, H. (1995), *The Knowledge-creating Company*, Oxford University Press, New York: NY.

Nordin, M., Hiebert, R., Pietrek, M., Alexander, M., Crane, M. and Lewis, S. (2002), "Association of Comorbidity and Outcome in Episodes of Nonspecific Low Back Pain in Occupational Populations," *Journal of Occupational and Environmental Medicine*, Vol. 44, No. 7, pp. 677-684.

OECD. (2004), "OECD GLOSSARY OF STATISTICAL TERMS." Retrieved February 2004, 2004.

Olsen, J. A., Kidholm, K., Donaldson, C. and Shackley, P. (2004), "Willingness to pay for public health care: a comparison of two approaches", *Health Policy*, Vol. 70, pp. 217–228.

Pareek, M., Clark, T., Dillon, H., Kumar, R., Stephenson, I. (2008), "Acceptance of pre-pandemic vaccination against influenza in healthcare workers", Working paper.

Parente, S. T., Salkever, D., DaVanzo, J. (2003), "The role of consumer knowledge of insurance benefits in the demand for preventative health care among the elderly", NBER Working Paper, W9912, National Bureau of Economic Research, Cambridge, MA.

Park, J.-H., Cheong, H.-K., Son, D.-Y., Kim, S.-U. and Ha, C.-M. (2010), "Perceptions and behaviors related to hand hygiene for the prevention of H1N1 influenza transmission among Korean university students during the peak pandemic period", *BMC Infectious Diseases*, Vol. 10, p. 222.

Pedersen, L. B., Gyrd-Hansen, D. and Kjaer, T. (2011), "The influence of information and private versus public provision on preferences for screening for prostate cancer: A willingness-to-pay study", *Health Policy*, Vol. 101, pp. 277– 289.

Poland, G. A. (2010), "The 2009–2010 influenza pandemic: effects on pandemic and seasonal vaccine uptake and lessons learned for seasonal vaccination campaigns", *Vaccine*, Vol. 28, Supl. 4, pp. D3–D13.

Redelings, M. D., Piron, J., Smith, L. V., Chan, A., Heinzerling, J., Sanchez, K. M., Bedair, D., Ponce, M. and Kuo, T. (2012), "Knowledge, attitudes, and beliefs about seasonal influenza and H1N1 vaccinations in a low-income, public health clinic population", *Vaccine*, Vol. 30, pp. 454–458.

Ruhm, C. J. (1995), "Economic conditions and alcohol problems", *Journal of Health Economics*, Vol. 14, No. 5, pp. 583–603.

Ruhm, C. J. (2000), "Are recessions good for your health?", *Quarterly Journal of Economics*, Vol. 115, No. 2, pp. 617–650.

Ruhm, C. J. (2003), "Good times make you sick", *Journal of Health Economics*, Vol. 22, No. 4, pp. 637–658.

Ruhm, C. J. (2005), "Healthy living in hard times", *Journal of Health Economics,* Vol. 24**,** No. 2, pp. 341-363.

Ruhm, C. J. (2007), "A healthy economy can break your heart", *Demography*, Vol. 44, pp. 829–848.

Ruhm, C. J. and Black, W. E. (2002), "Does drinking really decrease in bad times?", Journal of Health Economics, Vol. 21, No. 4, pp. 659–678.

Ruiz, F., Amaya, L., Venegas, S. (2007), "Progressive segmented health insurance: Colombian health reform and access to health services", *Health Economics*, Vol. 16, No.1, pp. 3-18.

Schön, D. (1983), *The Reflective Practitioner*, Basic Books, New York, NY, pp.267.

Scoppa, V. (2008), "Worker absenteeism and incentives: Evidence from Italy", Unpublished.

Shureiqi, I., Cooksley, C. D., Morris, J., Soliman, A. S., Levin, B. and Lippman, S. M. (2001), "Effect of Age on Risk of Second Primary Colorectal Cancer," *Journal of the National Cancer Institute*, Vol. 93, No. 16, pp. 1264-1266.

Sin, C. Y. and White, H. (1996), "Information criteria for selecting possibly misspecified parametric models", *Journal of Econometrics*, Vol. 71, pp. 207-225.

Smith, R. D. (2003), "Construction of the contingent valuation market in healthcare: A critical assessment", *Health Economics*, Vol. 12, pp. 609–628.

Smith, R. D. (2008), "Contingent valuation in health care: does it matter how the 'good' is described?", *Health Economics*, Vol. 17, No. 5, pp. 607–617.

Sokejima, S. and Kagamimori, S. (1998), "Working hours as a risk factor for acute myocardial infarction in Japan: a case-control study", *The British Medical Journal*, Vol. 317, September, pp. 775–780.

Staub, K. E. and Winkelmann, R. (2011), "Consistent estimation of zero-infated count models", *University of Zurich Socioeconomic Institutes Working Paper* SOI 0908.

Tavares, A. I. and Pita Barros, P. (2011), "Ex-ante moral hazard and primary prevention, evidence from Portugal", Working Papers in Economics E/no 56/2011, Área Científica de Economia, Universidade de Aveiro Departamento de Economia, Gestão e Engenharia Industrial.

Tefft, N. and Kageleiry, A. (2014), "State- level unemployment and the utilization of preventive medical services", *Health Services Research*, Vol. 49, No 1, pp.186-205.

The European Foundation for the Improvement of Living and Working Conditions (1997), "Preventing absenteeism at the workplace".

Türp, J. C., Kowalski, C. J. and Stohler, C. S. (2000), "Generic pain intensity scores are affected by painful comorbidity," *Journal of Orofacial Pain*, Vol. 14, No. 1, pp. 47-51.

USA - National Heart, Lung, and Blood Institute, 1998.

Van, D., McLaws, M.-L., Crimmins, J., MacIntyre, C. R. and Seale, H. (2010), "University life and pandemic influenza: Attitudes and intended behaviors of staff and students towards pandemic (H1N1) 2009", *BMC Public Health*, Vol. 10, p. 130.

Vírseda, S., Restrepo, M. A., Arranz, E., Magán-Tapia, P., Fernández-Ruiz, M., de la Cámara, A. G. et al. (2010), "Seasonal and pandemic A (H1N1) 2009 influenza vaccination coverage and attitudes among health-care workers in a Spanish University Hospital", *Vaccine*, Vol. 28, pp. 4751–4757.

Viscusi, W. K. (1990), "Do smokers underestimate risk?", *Journal of Political Economy*, Vol. 98, No. 6, pp. 1253-1269.

Vistnes, J. P. (1997), "Gender differences in days lost from work due to illness", *Industrial and Labor Relations Review*, Vol. 50, No. 2, pp. 304-323.

Von Krogh, G., Ichijo, K. and Nonaka, I. (2000), *Enabling Knowledge Creation: How to Unlock the Mystery of Tacit Knowledge and Release the Power of Innovation*, Oxford University Press, New York.

Wagner, T. H., Hibbard, J. H., Greenlick, M. R. and Kunkel, L. (2001), "Does providing consumer health information affect self-reported medical utilization? Evidence from the healthwise communities project", *Medical Care*, Vol. 39, pp. 836–847.

Winkelmann, R. (1999), "Wages, firm size and absenteeism", *Applied Economics Letters*, Vol. 6, No.6, pp. 337-341.

Winkelmann, R. (2003), *Econometric analysis of count data*. Berlin ; New York, Springer.

World Health Organization [WHO] (2005), *Preventing Chronic Diseases a vital investment*, WHO, Geneve.

Wu, S. (2003), Sickness and Preventive Medical Behavior, Working-paper Hamilton College.

Xu, X. and Kaester. R. (2010), "The business cycle and health behaviors", *National Bureau of Economic Research Working Paper*, No. 15737, February.

Zethraeus, Niklas (1998), "Willingness to pay for hormone replacement therapy", *Health Economics*, Vol.7, No 1, pp. 31-38.

Zubarik, R., Eisen, G., Zubarik, J., Teal, C., Benjamin, S., Glaser, M. and Jack, M. (2000), "Education improves colorectal cancer screening by flexible sigmoidoscopy in an inner city population", *The American Journal of Gastroenterology*, Vol. 95, No. 2, pp. 509-512.

Zweifel, P. (1981), "Supplier-Induced Demand in a Model of Physician Behavior", J. Van der Gaag and M. Perlman, *Health, Economics, and Health Economics*, North-Holland: pp. 245-267.

APPENDIXES

Appendix 1: Chapter II

This appendix contains the questionnaires applied in the first wave of the research and the follow-up questionnaire.

QUESTIONNAIRE

The aim of this questionnaire is to know the opinion and information from the population regarding H1N1 virus (known as Influenza A). The data will be processed anonymously (your name is not requested anywhere) only for investigation purposes. Your opinion is very important for us. Please read the following questions carefully and answer them as accurately as possible.

1) Do you know if you are eligible to be vaccinated against H1N1 virus according to the future vaccination campaign from the Portuguese government? (Please choose only one answer)

Yes, for sure	Very likely	Probably not	Sure not	not
☐	☐	☐	☐	☐

1/Vaccinate: 1 (yes, for sure) ; 2 (Very likely) ; 3 (Probably not) ; 4 (Sure not); 5(not)

2) Do you think you belong to the "high risk" group of the government vaccination programme?

Yes. ☐ Which group? _____ No. ☐

2(a)/Risk: 1 (yes) ; 2 (no)

2(b)/Group: 1 (health) ; 2 (occupation) ; 3 (age)

3) In case you are not eligible for free vaccination according to the government programme. What is the maximum price you are willing to pay to get the vaccine? Please complete the sentence:

I would not pay more than _____€ to buy the vaccine.

3/willing: amount

4) If the waiting time to be vaccinated is too high, will you choose not to be vaccinated?

If it is more than 1h, I would rather not be vaccinated	If it is more than 2h, I would rather not be vaccinated	If it is more than 4h, I would rather not be vaccinated	I will wait as long as it takes
☐	☐	☐	☐

4/queue: 1 (If it is more than 1h, I would rather not be vaccinated) ; 2 (If it is more than 2h, I would rather not be vaccinated) ; 3 (If it is more than 4h, I would rather not be vaccinated) ; 4 (I will wait as long as it takes)

5) Do you usually get vaccinated against seasonal influenza?

Every year ☐ Sometimes ☐ Never ☐

5/seasonal: 1 (every year) ; 2 (sometimes) ; 3 (never)

6) Did you start washing your hands more often every day since you became aware of the H1N1 virus? Yes ☐ No ☐

6/wash_hands: 1 (yes) ; 2 (no)

7) How much time do you spend each day with the media searching for information in general and not only about influenza A?

_____hours and _____minutes.

7/min_info: time in minutes

8) Did you change the time you spend each day searching for information after H1N1 virus appeared?

Increased a lot	Increased a little	Kept the same	Decreased a little	Decreased a lot
☐	☐	☐	☐	☐

8/info_flu: 1 (increased a lot) ; 2 (increased a little) ; 3 (kept the same) ; 4 (decreased a little) ; 5 (decreased a lot)

9) How do you keep yourself informed? Please rank your information sources (Radio, TV, Newspapers, Internet) according to the regularity with which you use them. 1 is the one you use the most and 4 is the one you use the least (Please do not mention the ones you do not use at all).

1) _____

2) _____

3) _____

4) _____

9.1/media : 1 (tv) ; 2 (radio) ; 3 (internet) ; 4 (newspapers)

9.2/media: 1 (tv) ; 2 (radio) ; 3 (internet) ; 4 (newspapers)

9.3/media: 1 (tv) ; 2 (radio) ; 3 (internet) ; 4 (newspapers)

9.4/media: 1 (tv) ; 2 (radio) ; 3 (internet) ; 4 (newspapers)

10) Do you know how many deaths have been registered in Portugal until now due to the virus? Complete the following sentence: there were around _____ deaths registered in Portugal.

10/deaths: number

11) Comparing with the previous answer you expect the number of deaths next year to

Increase a lot	Increase a little	Be approximately the same	Decrease a little	Decrease a lot
☐	☐	☐	☐	☐

11/exp_deaths: 1 (increase a lot); 2 (increase a little) ; 3 (be approximately the same) ; 4 (decrease a little) ; 5 (decrease a lot)

12) Do you know how many people were infected by Influenza A virus in Portugal? Complete the following sentence: there are around _____ people infected in Portugal.

12/affected: number

13) Comparing with the previous answer you expect the number of people infected by the virus next year to:

Increase a lot	Increase a little	Be approximately the same	Decrease a little	Decrease a lot
☐	☐	☐	☐	☐

13/exp_affected: 1 (increase a lot); 2 (increase a little) ; 3 (be approximately the same) ; 4 (decrease a little) ; 5 (decrease a lot)

14) "If you're not eligible for the State's free vaccination program, what would be the maximum price you would be willing to pay for this vaccine? Please complete the sentence:

I would pay up to _____€ for the vaccine."

14/price: price

15) Comparing with the previous answer within a year you expect the price of the vaccine to:

Increase a lot	Increase a little	Be approximately the same	Decrease a little	Decrease a lot
☐	☐	☐	☐	☐

15/exp_price: 1 (increase a lot); 2 (increase a little) ; 3 (be approximately the same) ; 4 (decrease a little) ; 5 (decrease a lot)

16) Please answer the following questions regarding your socio-economic situation:

16.1) Gender: male ☐ female ☐

16.1/male: 1 (male) ; 2 (female)

16.2) Age: _____

16.2/age

16.3) Occupation: _____

16.3/pers_occup: 1 (student) ; 2 (student worker) ; two digit categorization

16.4) Net monthly income (please choose only one answer):

Less than 350€	350 € to 900 €	901 € to 2000 €	2001 € to 3500 €	More than 3500 €
☐	☐	☐	☐	☐

16.4/pers_salary: 1 (less than 350) ; 2 (350 to 900) ; 3 (901 to 2000) ; 4 (2001 to 3500) ; 5 (more than 3500)

In case you are a student who depends on the household income answer the following question, if not ignore this question:

16.5) Number of people from the household: _____

16.5/family: number

16.6) Father's occupation: _____

16.6/fath_occup: 0 (unemployed) ; 1(retired) ; 2(deceased) ; two digit categorization

16.7) Net monthly income of the father (please choose only one answer)

Less than 350€	350 € to 900 €	901 € to 2000 €	2001 € to 3500 €	More than 3500 €
☐	☐	☐	☐	☐

16.7/fath_salary: 1 (less than 350) ; 2 (350 to 900) ; 3 (901 to 2000) ; 4 (2001 to 3500) ; 5 (more than 3500)

16.8) Mother's occupation: _____

16.8/moth_occup: 0 (unemployed) ; 1(retired) ; 2(deceased) ; 3 (housemaid) ; two digit categorization

16.9) Net monthly income of the mother (please choose only one answer)

More than 350€	350 € to 900 €	901 € to 2000 €	2001 € to 3500 €	More than 3500 €
☐	☐	☐	☐	☐

16.9/ moth_salary: 1 (less 350) ; 2 (350 to 900) ; 3 (901 to 2000) ; 4 (2001 to 3500) ; 5 (more than 3500)

17) Did you find this questionnaire easy to answer? (Please choose only one answer):

Very easy	Easy	Neither easy nor difficult	Difficult	Very difficult
☐	☐	☐	☐	☐

17/difficulty: 1 (very easy) ; 2 (easy) ; 3 (neither easy nor difficult) ; 4 (difficult) ; 5 (very difficult)

We would like to thank you both for your time and for the valuable information you gave us. Our investigation will proceed following the progress of the information related to H1N1 virus. This is why we would like to ask you to tell us in a few months if you were vaccinated or not. If you agree on giving us that information in the future, please give us your e-mail address or your mobile phone number so that we can contact you. Once again your personal data are protected by law and will be used only for investigation purposes.

E-mail: _____

Mobile phone number: _____

Responsible investigators:

Nikolaos Georgantzis (University of Granada)

Pedro Pita Barros (Universidade Nova de Lisboa)

Antonios Proestakis (University of Granada)

Aurora Garcia Gallego (UJI and University of Granada)

Questionário

O objectivo deste questionário é conseguir saber a opinião e informação da população quanto ao vírus H1N1 (conhecido como a gripe a). os dados irão ser processados de modo anónimo (pelo que o seu nome não é pedido em nenhum momento) apenas para fins de investigação. a sua opinião é muito importante para nós. **por favor leia as perguntas cuidadosamente e responda da forma mais precisa possível.**

1) Sabe se é elegível para ser vacinado contra o vírus H1N1 de acordo com a futura campanha de vacinas do Governo português? (Por favor assinale apenas uma resposta)

Sim, de certeza	Muito provavelmente	Provavelmente não	De certeza que não
☐	☐	☐	☐

1/Vacinate: 1 (sim, de certeza) ; 2 (Muito provavelmente) ; 3 (Provavelmente não) ; 4 (de certeza que não)

2) Considera pertencer ao grupo "elevado risco" do programa de vacinas do estado?

Sim. ☐ Qual Grupo? _____ Não. ☐

2(a)/Risk: 1 (sim) ; 2 (não)

2(b)/Group: 1 (saúde) ; 2 (profissão) ; 3 (idade)

3) No caso de não ser elegível para vacinação gratuita de acordo com o programa do Estado. Qual é o preço máximo que está disposto a pagar para receber a vacina? Por favor complete a afirmação:

Não pagaria mais de _____€ para comprar a vacina.

3/willing: quantidade

4) Se o tempo de espera para tomar a vacina for muito elevado, optará por não ser vacinado?

Se exceder 1h, prefiro não tomar a vacina	Se exceder 2h, prefiro não tomar a vacina	Se exceder 4h, prefiro não tomar a vacina	Esperarei o que for necessário
☐	☐	☐	☐

4/queue: 1 (Se exceder 1h, prefiro não tomar a vacina) ; 2 (Se exceder 2h, prefiro não tomar a vacina) ; 3 (Se exceder 4h, prefiro não tomar a vacina) ; 4 (Esperarei o que for necessário)

5) Costuma tomar a vacina contra a gripe sazonal?

Todos os anos ☐ Algumas vezes ☐ Nunca ☐

5/seasonal: 1 (todos os anos) ; 2 (algumas vezes) ; 3 (nunca)

6) Aumentou o número de vezes, por dia, que lava as mãos desde que sabe da existência do vírus H1N1?

Sim ☐ Não ☐

6/wash_hands: 1 (sim) ; 2 (não)

7) Quanto tempo gasta por dia a obter informação, em geral e não apenas sobre a gripe A, nos media?

_____horas e _____minutos.

7/min_info: tempo em minutos

8) Alterou o tempo que passa por dia a obter informação depois de surgir o vírus H1N1?

Aumentei muito	Aumentei um pouco	Mantive	Diminui um pouco	Diminui muito
☐	☐	☐	☐	☐

8/info_gripe: 1 (aumentei muito) ; 2 (aumentei pouco) ; 3 (mantive) ; 4 (diminui um pouco) ; 5 (diminui muito)

9) Como se mantém informado? Por favor ordene as suas fontes de informação (Radio, TV, Jornais, Internet) conforme a regularidade com que os usa. Sendo 1 o que usa mais e 4 o que usa menos. (Por favor não mencione aqueles que não usa de todo).

1) _____

2) _____

3) _____

4) _____

9.1/media : 1 (tv) ; 2 (radio) ; 3 (internet) ; 4 (jornais)

9.2/media: 1 (tv) ; 2 (radio) ; 3 (internet) ; 4 (jornais)

9.3/media: 1 (tv) ; 2 (radio) ; 3 (internet) ; 4 (jornais)

9.4/media: 1 (tv) ; 2 (radio) ; 3 (internet) ; 4 (jornais)

10) Sabe quantas mortes foram, até ao momento, registadas devido ao vírus em Portugal? Complete a frase seguinte: houve por volta de _____ mortes registadas em Portugal.

10/deaths: número

11) Comparando com a resposta na pergunta anterior espera que, no próximo ano, o número de mortes,

Aumente muito	Aumente um pouco	Se mantenha mais ou menos o mesmo	Diminua um pouco	Diminua muito
☐	☐	☐	☐	☐

11/exp_deaths: 1 (aumente muito) ; 2 (aumente um pouco) ; 3 (se mantenha mais ou menos o mesmo) ; 4 (diminua um pouco) ; 5 (diminua muito)

12) Sabe quantas pessoas foram infectadas pelo vírus da Gripe A em Portugal? Complete a frase seguinte: há por volta de _____ pessoas infectadas em Portugal.

12/affected: número

13) Comparando com a resposta na pergunta anterior, espera que, no próximo ano, o número de pessoas infectadas pelo vírus:

Aumente muito	Aumente um pouco	Se mantenha mais ou menos o mesmo	Diminua um pouco	Diminua muito
☐	☐	☐	☐	☐

13/exp_affected: 1 (aumente muito) ; 2 (aumente um pouco) ; 3 (se mantenha mais ou menos o mesmo) ; 4 (diminua um pouco) ; 5 (diminua muito)

14) Qual a sua expectativa para o preço da vacina uma vez lançada no mercado?

Calculo que seja _____€ / não sabe ☐

14/price: preço

15) Comparando com a sua resposta na pergunta anterior espera que o preço da vacina daqui a um ano:

Aumente muito	Aumente um pouco	Se mantenha mais ou menos o mesmo	Diminua um pouco	Diminua muito
☐	☐	☐	☐	☐

15/exp_price: 1 (aumente muito) ; 2 (aumente um pouco) ; 3 (se mantenha mais ou menos o mesmo) ; 4 (diminua um pouco) ; 5 (diminua muito)

16) Por favor responda às seguintes questões quanto à sua situação sócio-económica:

16.1) Sexo: masculino ☐ feminino ☐

16.1/male: 1 (male) ; 2 (female)

16.2) Idade: _____

16.2/age

16.3) Ocupação: _____

16.3/pers_occup: 1 (student) ; 2 (student worker) ; categorização de dois dígitos

16.4) Rendimento mensal líquido (por favor assinale apenas uma resposta):

Menos de 350€	350 € a 900 €	901 € a 2000 €	2001 € a 3500 €	Mais de 3500 €
☐	☐	☐	☐	☐

16.4/pers_salary: 1 (menos de 350) ; 2 (350 a 900) ; 3 (901 a 2000) ; 4 (2001 a 3500) ; 5 (mais de 3500)

Caso seja um estudante que depende do rendimento do agregado familiar responda à seguinte questão, caso contrário ignore esta questão:

16.5) Número total de membroa do agregado familiar: _____

16.5/family: número

16.6) Profissão do pai: _____

16.6/fath_occup: 0 (desempregado) ; 1(reformado) ; 2(falecido) ; categorização de dois dígitos

16.7) Rendimento mensal líquido do pai (por favor assinale apenas uma resposta)

Menos de 350€	350 € a 900 €	901 € a 2000 €	2001 € a 3500 €	Mais de 3500 €
☐	☐	☐	☐	☐

16.7/fath_salary: 1 (menos de 350) ; 2 (350 a 900) ; 3 (901 a 2000) ; 4 (2001 a 3500) ; 5 (mais de 3500)

16.8) Profissão da mãe: _____

16.8/moth_occup: 0 (desempregada) ; 1(reformada) ; 2(falecida) ; 3 (doméstica) ; categorização de dois dígitos

16.9) Rendimento mensal líquido da Mãe (por favor assinale apenas uma resposta)

Menos de 350€	350 € a 900 €	901 € a 2000 €	2001 € a 3500 €	Mais de 3500 €
☐	☐	☐	☐	☐

16.9/ moth_salary: 1 (menos de 350) ; 2 (350 a 900) ; 3 (901 a 2000) ; 4 (2001 a 3500) ; 5 (mais de 3500)

17) Achou fácil responder a este questionário? (Por favor assinale apenas uma resposta):

Muito fácil	Fácil	Nem fácil nem difícil	Difícil	Muito difícil
☐	☐	☐	☐	☐

17/difficulty: 1 (muito fácil) ; 2 (fácil) ; 3 (nem fácil nem difícil) ; 4 (difícil) ; 5 (muito difícil)

Gostaríamos de lhe agradecer tanto pelo seu tempo como pela informação valiosa que nos disponibilizou. A nossa investigação vai continuar seguindo o progresso da informação relacionada ao vírus H1N1. E é por isso que gostaríamos de pedir-lhe para daqui a uns meses nos dizer se foi vacinado(a) ou não. Se concordar em dar-nos essa informação no futuro, por favor providencie-nos com o seu endereço de E-mail ou o seu número de telemóvel para que consigamos contactá-lo. Mais uma vez, os seus dados pessoais são protegidos pela lei e serão apenas usados para fins de investigação.

E-mail: _____

Número de telemóvel: _____

Investigadores Responsáveis:

Nikolaos Georgantzis (Universidade de Granada)

Pedro Pita Barros (Universidade Nova de Lisboa)

Antonios Proestakis (Universidade de Granada)

Aurora Garcia Gallego (UJI e Universidade de Granada)

FOLLOW-UP QUESTIONNAIRE

The aim of this questionnaire is to know the opinion and information of the population regarding H1N1 virus (known as Influenza A) The data will be processed anonymously (your name is not requested anywhere) only for investigation purposes. Your opinion is very important for us.

Please read the questions carefully and answer them as accurately as possible.

1 Did you get vaccinated on Autumn/Winter 2009/10 against H1N1 virus?

YES □ NO □

If you answered NO in the previous question, qhich were the reasons? (rank the factors by importance over the decision taking, being 1 the most important and 4 the least important)

I DID NOT BELONG TO THE ELIGIBLE GROUPS □ VERY HIGH WAITING TIME □ INFORMATION - MEDIA □ ADVERSE MEDICAL OPINION □

2 Did you have Influenza A during Autumn/Winter 2009/10?

YES □ NO □

3 Was someone from your household or socio-professional group H1N1 virus positive during Autumn/Winter 2009/10?

YES □ NO □

If you answered YES in the previous question, choose your relation(s) with the carrier(s):

PARTNER □ FATHER/MOTHER □ SON/DAUGHTER □ GRANDSON/GRANDDAUGHTER □

BROTHER/SISTER □ SECOND RELATIVE □ FRIEND □ COLLEAGUE □

4 Did you start washing your hands more often every day since you became aware of H1N1 virus?

YES □ NO □

5 Did you get vaccinated against seasonal influenza during Autumn/Winter 2009/10?

YES NO

□ □

6 Did you get vaccinated against seasonal influenza during Autumn/Winter 2010/11?

YES	NO
□	□

7 Do you know the new vaccine agaisnt seasonal influenza (Autumn/Winter 2010/11) already inlcudes a strain from H1N1 virus?

YES	NO
□	□

If you answered YES in the previous question, how did you know it?

MEDIA	FAMILY DOCTOR	FAMILY AND FRIENDS	OTHER
□	□	□	□

8 How much time do you spend each day with the media searching for information in general and not only about influenza A?

minutes/hours (eliminate what does not matter)

How do you keep yourself informed? Please rank your information sources (Radio, TV, Newspapers, Internet) according to the regularity with which you use them. 1 is the one you use the most and 4 is the one you use the least (don not mention the ones you do not use at all).

TV	RADIO	NEWSPAPERS	INTERNET
□	□	□	□

9 Do you know how many deaths have been registered in Portugal due to the H1n1 virus in the current influenza period (Autumn/Winter 2010/11)?

YES	NO
□	□

If you answered YES in the previous question, complete the following sentence:

THERE WERE AROUND _____ DEATHS REGISTERED IN PORTUGAL.

10 Do you know how many people were infected by the H1N1 virus in Portugal in the current influenza period (Autumn/Winter 2010/11)?

YES	NO
□	□

If you answered YES in the previous question, complete the following sentence:

THERE ARE AROUND _____ PEOPLE INFECTED IN PORTUGAL.

11 Would you be willing to pay for a vaccine against Influenza A?

YES	NO
□	□

How much in Euro would you be willing to pay?

€

12 Please answer the following questions regarding your socio-economic situation:

MALE	FEMALE
□	□

25 YEARS OR LESS	26-45 YEARS	46-65 YEARS	66 YEARS OR MORE
□	□	□	□

SMOKER	NON-SMOKER
□	□

SECONDARY EDUCATION OR LOWER	BACHELOR'S OR UNIVERSITY DEGREE	MASTER'S DEGREE	DOCTORAL DEGREE
□	□	□	□

STUDENT	EMPLOYED	SELF-EMPLOYED	OTHER
□	□	□	□

Specify your answer to the previous question: if student, mention the course; if employed, mention the occupation; if OTHER, answer why.

Net monthly income of the houshold (please choose only one answer):

< 350 €	350-900 €	901-2000 €	2001-3500 €	> 3501 €	DOES NOT KNOW
□	□	□	□	□	□

IN CASE YOU ARE A STUDENT WHO DEPENDS ON THE HOUSEHOLD INCOME, ANSWER THE FOLLOWING QUESTIONS, IF NOT, IGNORE THEM:

Number of people from the household:

Father's occupation:

Net monthly income of the father (please choose only one answer):

< 350 €	350-900 €	901-2000 €	2001-3500 €	> 3501 €	DOES NOT KNOW
□	□	□	□	□	□

Mother's occupation:

Net monthly income of the mother (please choose only one answer):

< 350 €	350-900 €	901-2000 €	2001-3500 €	> 3501 €	DOES NOT KNOW
□	□	□	□	□	□

We would like to thank you both for your time and for the valuable information you gave us. Once again your data are protected by law and will be used only for investigation purposes.

Appendix 2: Chapter II

This appendix contains the results of Mann-whitney test by country, faculty and first mean of information used.

Non-parametric test on willingness to pay by country (if will<501)

```
Ho: will(country==portugal) = will(country==greece)
              z =  -6.603
    Prob > |z| =   0.0000


Ho: will(country==spain) = will(country==greece)
              z = -16.504
    Prob > |z| =   0.0000


Ho: will(country==portugal) = will(country==spain)
              z =  10.884
    Prob > |z| =   0.0000
```

Country effect on willingness

Mann-Whitney by faculty

```
Ho: will(fac==economics) = will(fac==law)
              z =  -0.930
     Prob > |z| =   0.3526

Ho: will(fac==economics) = will(fac==medicine)
              z =   7.088
     Prob > |z| =   0.0000

Ho: will(fac==economics) = will(fac==pharmacy)
              z =   4.016
     Prob > |z| =   0.0001

Ho: will(fac==law) = will(fac==medicine)
              z =   7.661
     Prob > |z| =   0.0000

Ho: will(fac==law) = will(fac==pharmacy)
              z =   4.715
     Prob > |z| =   0.0000

Ho: will(fac==medicine) = will(fac==pharmacy)
              z =  -3.299
     Prob > |z| =   0.0010
```

Mann-Whitney by med1
(only significant differences are reported)

```
Ho: will(med1==1) = will(med1==3)
              z =   4.339
     Prob > |z| =   0.0000

Ho: will(med1==1) = will(med1==4)
              z =   3.719
     Prob > |z| =   0.0002

Ho: will(med1==2) = will(med1==3)
              z =   2.104
     Prob > |z| =   0.0353

Ho: will(med1==2) = will(med1==4)
              z =   2.486
     Prob > |z| =   0.0129
```

Appendix 3: Chapter II

This appendix contains the results of the estimation of pooled specification using a linear regression, the results of Breusch- Pagan Test and the roust estimation of the follow-up survey.

```
reg  will gender age fsal_1 fsal_2 fsal_3 fsal_4  phar law eco   med1net med1rad med1paper
vacyes seasyes grinmore e_dead1 dano


      Source |       SS       df       MS              Number of obs =     122
-------------+------------------------------           F( 17,   104) =    2.44
       Model |  136552.151    17  8032.47945           Prob > F      =  0.0031
    Residual |  342839.358   104  3296.53228           R-squared     =  0.2848
-------------+------------------------------           Adj R-squared =  0.1679
       Total |  479391.508   121  3961.91329           Root MSE      =  57.415


------------------------------------------------------------------------------
        will |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      gender |  -7.668902   12.16923    -0.63   0.530    -31.80094    16.46314
         age |  -.2625224   3.001289    -0.09   0.930     -6.21419    5.689146
      fsal_1 |   8.934833   43.49252     0.21   0.838    -77.31246    95.18213
      fsal_2 |   13.30502   16.10528     0.83   0.411    -18.63237     45.2424
      fsal_3 |  -20.07306   14.08517    -1.43   0.157    -48.00447    7.858355
      fsal_4 |  -36.08952   22.52186    -1.60   0.112    -80.75123    8.572179
        phar |  -17.95128   16.51573    -1.09   0.280    -50.70258    14.80003
         law |  -31.18474    20.9186    -1.49   0.139     -72.6671    10.29763
         eco |  -3.373758   15.22439    -0.22   0.825    -33.56429    26.81678
     med1net |   .3849776   13.64845     0.03   0.978    -26.68042    27.45037
```

```
     med1rad |   53.71054   22.56466     2.38   0.019     8.963967    98.45712
   med1paper |    -6.2161   18.15458    -0.34   0.733    -42.21731    29.78511
      vacyes |     -22.36   16.88756    -1.32   0.188    -55.84866    11.12866
     seasyes |    53.1973   15.94815     3.34   0.001     21.57153    84.82307
    grinmore |   .2174737   .1029864     2.11   0.037     .0132478    .4216996
     e_dead1 |   7.854847    19.6417     0.40   0.690    -31.09538    46.80507
        dano |  -28.45754    15.9674    -1.78   0.078    -60.12149    3.206415
       _cons |   59.37253   62.25561     0.95   0.342    -64.08267    182.8277
------------------------------------------------------------------------------


. estat hettest

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: fitted values of will

         chi2(1)      =   216.23
         Prob > chi2  =   0.0000



. reg  will gender age fsal_1 fsal_2 fsal_3 fsal_4  phar  law eco   med1net med1rad med1paper
vacyes seasyes grinmore e_dead1 dano, vce(robust)


Linear regression                               Number of obs =     122
                                                F( 17,   104) =    1.44
                                                Prob > F      =  0.1326
                                                R-squared     =  0.2848
                                                Root MSE      =  57.415
```

	Robust				
will	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]
gender	-7.668902	10.29801	-0.74	0.458	-28.09024	12.75243
age	-.2625224	2.406687	-0.11	0.913	-5.035072	4.510028
fsal_1	8.934833	28.62644	0.31	0.756	-47.83247	65.70213
fsal_2	13.30502	24.63349	0.54	0.590	-35.54413	62.15416
fsal_3	-20.07306	14.51671	-1.38	0.170	-48.86023	8.714112
fsal_4	-36.08952	17.50214	-2.06	0.042	-70.79692	-1.382131
phar	-17.95128	14.02758	-1.28	0.203	-45.7685	9.865949
law	-31.18474	22.18558	-1.41	0.163	-75.17957	12.8101
eco	-3.373758	16.45216	-0.21	0.838	-35.99901	29.25149
med1net	.3849776	10.14283	0.04	0.970	-19.72864	20.49859
med1rad	53.71054	47.32419	1.13	0.259	-40.13511	147.5562
med1paper	-6.2161	20.05275	-0.31	0.757	-45.98146	33.54926
vacyes	-22.36	19.40817	-1.15	0.252	-60.84713	16.12713
seasyes	53.1973	27.34772	1.95	0.054	-1.034259	107.4289
grinmore	.2174737	.058767	3.70	0.000	.1009367	.3340108
e_dead1	7.854847	20.69909	0.38	0.705	-33.19222	48.90191
dano	-28.45754	10.67527	-2.67	0.009	-49.627	-7.288081
_cons	59.37253	50.19701	1.18	0.240	-40.17002	158.9151

Appendix 4 – Chapter III

This appendix contains the results of the estimation of an Ordered Probit for SAH.

	Ordered Probit – SAH		
Dependent Variable SAH	**Coef.**	**Std. Err**	**Z**
Health variables			
Rheumatism	-.3535113	.0245337	-14.41
Depressive disorder	-.3491712	.0296267	-11.79
High blood pressure	-.2566851	.0216698	-11.85
Chronic pain	-.3738215	.0231508	-16.15
Diabetes	-.3269082	.0341236	-9.58
Obesity	-.323763	.0386181	-8.38
Asthma	-.1752661	.0384865	-4.55
kidney stones	-.2209145	.0350425	-6.30
Renal failure	-.3801665	.0753171	-5.05
Cancer	-.4378426	.0625572	-7.00
Cerebral haemorrhage	-.2928317	.0746952	-3.92
Emphysema	-.2558219	.0513316	-4.98
Stroke	-.3987003	.0790234	-5.05
Health Behaviours			
Smoker	.0532262	.0193609	2.75
Drugs prescription	-.1976738	.0185638	-10.65
Meals	.1963264	.027765	7.07
Occupational variables			
Working hours in a week	.0025252	.0007998	3.16
Stress	-.117157	.0299421	-3.91
Other industry	.3926419	.0290562	13.51
Construction industry	.4679673	.0360588	12.98
Manufacturing industry	.2581595	.0341823	7.55

Health Protection System			
NHS	.0013797	.0304423	0.05
Socioeconomic			
Married	.2186147	.0389066	5.62
single	.5658109	.0428948	13.19
widow	.1919431	.0511021	3.76
Age2	-.1669497	.0231897	-7.20
Age3	-.3395477	.0264154	-12.85
Children	.0244026	.0108986	2.24
Male	.3247193	.0183736	17.67
Edumax1	-.0161531	.0234229	-0.69
Edumax2	.1091954	.0292421	3.73
Income	.0002354	.0000238	9.87
Geographic variables			
Norte	-.2961989	.0394926	-7.50
Centro	-.3082295	.0393073	-7.84
Lisboa	-.1951666	.0392978	-4.97
Alentejo	-.1420123	.0393699	-3.61
Algarve	-.2965488	.0395477	-7.50
Açores	-.4507329	.0314447	-14.33
Preventive Variables			
Blood pressure_ tests	-.0104746	.0186652	-0.56
/cut1	-1.417581	.0697199	
/cut2	-.085134	.0690699	
/cut3	.7815385	.0691554	

Appendix 5 – Chapter III

This appendix contains the results of the estimation of the first part of Hurdle Model for Absenteeism.

	Logit MODEL			**Marginal Effects Logit MODEL**		
Dependent variable Missing days - absence from work	**Coefficient**	**Std. Devition**	**Z**	**Coefficient**	**Std. Devition**	**Z**
Health variables						
Rheumatism	0.060442	0.233009	0.26	0.060442	0.23301	0.26
Depressive disorder	0.492318***	0.252986	1.95	0.492318***	0.25299	1.95
High blood pressure	-0.33192	0.208545	-1.59	-0.33192	0.20854	-1.59
Chronic pain	0.119857	0.216571	0.55	0.119857	0.21657	0.55
Diabetes	-0.23394	0.317009	-0.74	-0.23394	0.31701	-0.74
Obesity	0.224299	0.355736	0.63	0.224299	0.35574	0.63
Asthma	0.573012**	0.28808	1.99	0.573012**	0.28808	1.99
kidney stones	0.300111	0.299238	1	0.300111	0.29924	1
Renal failure	-0.49522	0.740818	-0.67	-0.49522	0.74082	-0.67
Cancer	-0.53348	0.73668	-0.72	-0.53348	0.73668	-0.72
Cerebral haemorrhage	-0.31808	0.677159	-0.47	-0.31808	0.67716	-0.47
Emphysema	-0.16282	0.422814	-0.39	-0.16282	0.42281	-0.39
Stroke	0.91203***	0.532235	1.71	0.91203***	0.53224	1.71
Health Behaviours						
Smoker	0.257941	0.213539	1.21	0.257941	0.21354	1.21
Drugs prescription	1.937676*	0.294922	6.57	1.937676*	0.29492	6.57
Meals	0.114146	0.312754	0.36	0.114146	0.31275	0.36
Occupational variables						
Stress	0.699223**	0.232108	3.01	0.699223*	0.23211	3.01
Working hours in a week	0.002288	0.007895	0.29	0.002288	0.0079	0.29
Other industry	-0.13795	0.302706	-0.46	-0.13795	0.30271	-0.46

Construction industry	0.60906***	0.365335	1.67	0.60906***	0.36534	1.67
Manufacturing industry	-0.4551	0.376997	-1.21	-0.4551	0.377	-1.21
Social Protection System						
NHS	-0.05254	0.352143	-0.15	-0.05254	0.35214	-0.15
Socioeconomic variables						
Married	-0.40564	0.374015	-1.08	-0.40564	0.37402	-1.08
Single	-0.67815	0.420919	-1.61	-0.67815	0.42092	-1.61
Divorce	0.1632	0.492901	0.33	0.1632	0.4929	0.33
Age2	-0.35146	0.231147	-1.52	-0.35146	0.23115	-1.52
Age3	-0.78175*	0.263807	-2.96	-0.78175*	0.26381	-2.96
Children	0.036966	0.116701	0.32	0.036966	0.1167	0.32
Male	0.409995**	0.209347	1.96	0.409995**	0.20935	1.96
Edumax1	-0.18916	0.244893	-0.77	-0.18916	0.24489	-0.77
Edumax2	-0.20684	0.318506	-0.65	-0.20684	0.31851	-0.65
Income	-0.00033	0.000257	-1.28	-0.00033	0.00026	-1.28
Geographic variables						
Norte	0.387096	0.29505	1.31	0.387096	0.29505	1.31
Centro	0.247443	0.298414	0.83	0.247443	0.29841	0.83
Madeira	-0.67364	0.531733	-1.27	-0.67364	0.53173	-1.27
Alentejo	-0.10533	0.335877	-0.31	-0.10533	0.33588	-0.31
Algarve	0.207012	0.314558	0.66	0.207012	0.31456	0.66
Açores	0.056064	0.401233	0.14	0.056064	0.40123	0.14
Preventive variables						
Blood pressure_ tests	-0.0864	0.381492	-0.23	-0.0864	0.38149	-0.23
Constant	-4.64597	0.812749	-5.72			

Notes: Coefficients market with *, ** and*** are significant at 1, 5 and 10 per cent level, respectively.

Appendix 6 – Chapter III

This appendix contains the results of the estimation of the second part of Hurdle Model for Absenteeism.

	NB2 MODEL			Marginal Effects NB2 MODEL		
Dependent variable Missing days - absence from work	**Coefficient**	**Std. Devition**	**Z**	**Coefficient**	**Std. Devition**	**Z**
Health variables						
Rheumatism	0.167788	0.279766	0.6	0.167788	0.27977	0.6
Depressive disorder	0.277989	0.261937	1.06	0.277989	0.26194	1.06
High blood pressure	-0.26481	0.300449	-0.88	-0.26481	0.30045	-0.88
Chronic pain	-0.27492	0.295155	-0.93	-0.27492	0.29516	-0.93
Diabetes	0.435335	0.333176	1.31	0.435335	0.33318	1.31
Obesity	-0.75668***	0.455428	-1.66	-0.75668***	0.45543	-1.66
Asthma	1.1452**	0.510548	2.24	1.1452**	0.51055	2.24
kidney stones	-0.26585	0.321215	-0.83	-0.26585	0.32121	-0.83
Renal failure	-0.51373	0.526399	-0.98	-0.51373	0.5264	-0.98
Cancer	-0.14076	0.467245	-0.3	-0.14076	0.46724	-0.3
Cerebral haemorrhage	-1.48767***	0.811526	-1.83	-1.48767***	0.81153	-1.83
Emphysema	-0.76628	0.565048	-1.36	-0.76628	0.56505	-1.36
Stroke	-2.63216*	0.821635	-3.2	-2.63216*	0.82164	-3.2
Health Behaviours						
Smoker	0.366521***	0.200102	1.83	0.366521***	0.2001	1.83
Drugs prescription	0.760774**	0.370349	2.05	0.760774**	0.37035	2.05
Meals	-0.07623	0.331637	-0.23	-0.07623	0.33164	-0.23
Occupational variables						
Stress	-0.75857*	0.217088	-3.49	-0.75857*	0.21709	-3.49
Working hours in a week	-0.00615	0.008146	-0.76	-0.00615	0.00815	-0.76
Other industry	0.416236	0.396167	1.05	0.416236	0.39617	1.05
Construction industry	0.400485	0.451869	0.89	0.400485	0.45187	0.89
Manufacturing industry	0.770515	0.495845	1.55	0.770515	0.49585	1.55

Social Protection System						
NHS	-0.70171***	0.393055	-1.79	-0.70171***	0.39306	-1.79
Socioeconomic variables						
Married	0.993284**	0.45054	2.2	0.993284**	0.45054	2.2
Single	0.657754	0.544407	1.21	0.657754	0.54441	1.21
Divorce	1.4351**	0.576891	2.49	1.4351**	0.57689	2.49
Age2	0.597407**	0.283059	2.11	0.597407**	0.28306	2.11
Age3	0.453184	0.285988	1.58	0.453184	0.28599	1.58
Children	-0.04385	0.135825	-0.32	-0.04385	0.13583	-0.32
Male	0.220745	0.252444	0.87	0.220745	0.25244	0.87
Edumax1	0.428834	0.269541	1.59	0.428834	0.26954	1.59
Edumax2	0.379182	0.391176	0.97	0.379182	0.39118	0.97
Income	-0.00031	0.000359	-0.86	-0.00031	0.00036	-0.86
Geographic variables						
Norte	0.225505	0.326303	0.69	0.225505	0.3263	0.69
Centro	-0.03281	0.270612	-0.12	-0.03281	0.27061	-0.12
Madeira	-0.45486	0.488179	-0.93	-0.45486	0.48818	-0.93
Alentejo	-0.05019	0.394689	-0.13	-0.05019	0.39469	-0.13
Algarve	-0.39366	0.384479	-1.02	-0.39366	0.38448	-1.02
Açores	-0.86855***	0.465349	-1.87	-0.86855***	0.46535	-1.87
Preventive variables						
Blood pressure_ tests	-0.3282	0.699329	-0.47	-0.3282	0.69933	-0.47
Constant	-0.00631	1.155949	-0.01			

Notes: Coefficients market with *, ** and*** are significant at 1, 5 and 10 per cent level, respectively.