A microeconometric analysis of health care utilization in Europe

Majo, M.C.

Maria Cristina Majo

A Microeconometric Analysis of Health Care Utilization in Europe

A Microeconometric Analysis of Health Care Utilization in Europe

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Universiteit van Tilburg, op gezag van de rector magnificus, prof. dr. Ph. Eijlander, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie aan het Einaudi Instituut voor Economie en Financiën (EIEF) van de Universiteit Tor Vergata te Rome (Italië) op dinsdag 21 december 2010 om 11.00 uur door

Maria Cristina Majo

geboren op 13 januari 1979 te Napels, Italië.

Promotor:
prof. dr. A.H.O. van Soest

Commissieleden:
prof. dr. V. Atella
prof. dr. D. Fabbri
prof. dr. M. Padula
prof. dr. F. Peracchi

A Microeconometric Analysis of Health Care Utilization in Europe

TESI

per il conseguimento del titolo di Dottore di Ricerca all'Università di Tilburg (Paesi Bassi), con l'autorità del Magnifico Rettore, Prof. Ph. Eijlander, alla presenza della commissione d'esame designata dal Collegio dottorale nella discussione pubblica tenuta presso l'Istituto Einaudi per l'economia e la finanza (EIEF), Università di Roma Tor Vergata (Italia), in data martedì 21 dicembre 2011, alle ore 11.00, da

Maria Cristina Majo

nata il 13 gennaio 1979 a Napoli, Italia.

Acknowledgements

"The teacher who is indeed wise does not bid you to enter the house of his wisdom but rather leads you to the threshold of your mind."
K. Gibran, The Prophet (1923).

There are many people I would like to thank for contributing, in one way or another, to the completion of this thesis.

First of all I would like to express my sincere gratitude to Prof. Arthur van Soest for being my thesis supervisor. I am especially grateful to him for the way he welcomed me at Tilburg University. He has been my guide in the doctoral process, he was a source of continuous inspiration and challenge. I learned so much from him, both scientifically and personally. His knowledge and his being always respectful and open-minded about my ideas and research output impressed me tremendously. I thank him for the infinite number of hours he spent reviewing my work. Without his energy and commitment I would have not completed this dissertation. Thank you!

Furthermore, I would like to thank the scholarship from Tor Vergata University that granted me the opportunity to start the doctoral research, and Prof. Franco Peracchi who encouraged me to go abroad for a research visit which represented a crucial step in my research. I owe this thesis to the Economics of Ageing in Europe (AGE) RTN European Program (HPRN–CT–2002–00235) and to the EU-project SHARELIFE. They provided me financial support and made my research visit at Tilburg University possible. My visit at the Department of Econometrics at Tilburg University has been extremely inspiring and educational, and provided unique opportunities for a young researcher like me.

I also wish to express my gratitude to my colleagues at the Public Mental Health Department and at the Methodology and Statistics Section at the Trimbos Institute in Utrecht. The way they welcomed me when I started working there was impressively generous. They have been patient during my attempts to learn Dutch, and provided a great and *gezellig* working atmosphere. Thank you for always being supportive and allowing me enough time to finalize my dissertation.

A PhD is a complex path, you face many challenges in improving your knowledge and your research skills, but above all it represents an important personal challenge. The support of who believes in you, and being able to keep your private life active and joyful, are the keys to keep you motivated and finish your thesis. I would like to thank my friends, without their help this thesis wouldn't exist and probably neither would my mental sanity :-). I will mention a few of them (as to list everybody would require an entire book on itself): Alerk, Amar, Andrey, Bart, Benjamin, Bianca, Cecilia, Chris, Claudio, Corrie, Dagma, DJ, Domenico, Edwin, Flaminia, Giuseppe, Harrie, Heejung, Jan-Willem, Karen, Katie, Khoa, King, Linde, Marcel, Maria, Melissa, Michael, Michele, Norma, Oktay, Olha, Owen, Rafiq, Renata, Rick, Tom, Valeria, Yan, Yvette, Zhamilya. A special thanks to Andrea, Ben, Corrado, Emiliya, and Gema for making these years in Tilburg an indelible memory; and to Ilaria and Silvia for being my longtime (now also longdistance) friends. I offer my sincerest gratitude and thanks to them all, those who are mentioned here, and those who I met on my way and gave their contribution to improve myself both personally and scientifically.

Finally, I would like to thank those whose contribution was most intangible, yet so important: my parents, Paola and Piero, and my exceptional brother, Gabriele, for always showing an interest in what I was doing in Tilburg, even though this meant that I had to leave Rome and be far away. I am grateful to my parents who shaped my belief in the importance of a good education, and laid the foundations for my interest in academic research. *Grazie per aver creduto in me e per avermi, da sempre, incoraggiato e sostenuto.*

Last but not least, I would like to thank Willem, for supporting me while I was writing my PhD-thesis and for sharing with me an important part of my life. *Ik waardeer enorm je steun en vertrouwen gedurende de afgelopen jaren.* Thanks Willem, for everything... and more yet to come ;-)!

's-Hertogenbosch, November 2010

Contents

List of Figures

List of Tables

Introduction

The share of the total European population older than 65 is set to increase – from 16.1% in 2000 to 22% by 2025 and 27.5% by 2050 (European Commission 2001). These numbers will certainly pose big challenges to existing health care systems. Economic health care policies should be aimed at reducing the burden of aging populations on society and at the same time ensuring the availability of health and social services for older persons. This would promote their continued participation in a socially and economically productive life.

Moreover, in recent years there has been increasing interest in health promotion and disease prevention activities. The aging of the population encourages innovations in preventive care, as future reduction of morbidity and mortality is linked to the diffuse adoption of preventive practices. Preventive care therefore plays a key role in population health. Links between access to health care, utilization of care services, and socio-economic position are well established (see, for example, WHO 2005). Existing literature (Syme 1998) suggests that socio-economic status is relevant to both morbidity and mortality of diseases, and is therefore an important factor to take into consideration for evaluating and managing prevention.

More specifically, by analyzing the relationship between socio-economic status, health, and health care use for a variety of developed countries (with a main focus on Europe), this thesis attempts to address several questions:

- What are the socio-economic factors driving the use of health care services: income, wealth and/or education?
- Does the relationship between socio-economic factors and health care use vary with different types of health care services, such as primary care, specialist care, or in- and outpatient care in a hospital?
- How is preventive clinical service utilization related to socio-economic status in the population aged 50 and over?
- Are there different socio-economic factors driving the use of preventive care services than those driving usual care?
- How should the empirical analysis be modified when dealing with count data with excess zeros?

The data that are used in the three chapters of the thesis are drawn from the Survey of Health, Ageing and Retirement in Europe (SHARE)[1]. SHARE provides crucial information for the evaluation of health systems, including harmonized information on a variety of dimensions such as health, health care use, and socio-economic conditions at the individual level.

This thesis consists of three separately readable chapters that were independently written[2].

Chapter 1 addresses the question how income affects health care utilization by the population aged 50 and over in the United States and a number of European countries with varying health care systems. The probabilities that individuals receive several medical services (visits to general practitioner, specialist, dentist, inpatient, or outpatient services) are analyzed separately using probit models. In addition to controls for income and demographic characteristics, controls for health status (both subjective and objective measures of health) are used. We analyze how the relationship between income and health care utilization varies across countries and relate these cross country differences to characteristics of the health care system, i.e., per capita total and public expenditure on health care, gate-keeping for specialist care, and co-payments.

In Chapter 2 we deal with the question how preventive clinical service utilization by the population aged 50 and over is related to socio-economic status in a number of European countries with varying health care systems. The probabilities that individuals receive preventive clinical services (influenza vaccination, blood check, colonoscopy, blood stool test, eye exam, and mammogram for women) are analyzed separately using probit models. In addition to controls for education and demographic characteristics, controls for economic factors and health status (both subjective and objective measures of health) are used. The analysis of education first, and then of all three indicators of socio economic status – education, income, and work status – suggests that economic and social resources are associated with whether respondents use preventive services. The main result is that education level emerges as a very important determinant for the uptake of preventive care.

Chapter 3 is devoted to the analysis of response variables that are scored as counts and that present a large number of zeros, which often arises in quantitative health care analysis. A zero-inflated Poisson model with fixed-effects is defined to identify respondent- and health-related characteristics associated with health care demand. This is a new model that is proposed to model count measures of health care utilization and account for the panel structure of the data. Parameter estimation is achieved by conditional maximum likelihood. An application of the new model is implemented using SHARE data from the 2004–2006 waves, and compared to existing panel data models for count data. Results

[1]Chapter 1 also uses data from the Health and Retirement Survey (HRS) for the United States.

[2]More details are provided at the beginning of each chapter.

show that separately controlling for whether outcomes are zero or positive in one of the two years does make a difference for counts with a larger number of zeros.

Chapter 1

Income and Health Care Utilization*

1.1 Introduction

Ensuring socio-economic equity and reactivity of health care systems is often considered a high priority in health care policy (Van Doorslaer et al. 2006). In the United Kingdom for example, equitable access to health care is an explicit goal of government policy (Deaton 2002). The ministers of health from Chile, Germany, Greece, New Zealand, Slovenia, Sweden, and the United Kingdom have formed an international forum on matters relating to access to health care services, to sustain the goal of equitable access to good quality health care (Oliver and Mossialos 2008). Policy makers should have insight in the inequality changing effects of various health care systems, as lack of access and quality may cause or at least reinforce the positive association between socio-economic status (SES) and health, the so-called SES gradient in health (Deaton 2002).

SES is a comprehensive concept based on income, education, occupation, and sometimes wealth. Income is a commonly used measure of SES because it is relatively easy to report for most individuals and easier to compare across countries than, for example, education level. For this reason we choose income as the measure for SES, and refer to income and SES interchangeably, in spite of the broader meaning that SES entails. In this study we compare the relationship between SES and health care utilization in countries with very different health care policies, exploiting the large cross-country variation in health care systems to analyze which policies are effective to make the utilization of health care more equitable.

The share of the total European population older than 65 is set to increase – from 16.1% in 2000 to 22% by 2025 and 27.5% by 2050 (European Commission 2001). People 65+ represented 12.4% of the United States (US) population in the year 2000 but are expected

*This chapter is based on "Income and Health Care Utilization Among the 50+ in Europe and the US", M.C. Majo, and A. van Soest. *Acknowledgments:* I am particularly grateful to Arthur van Soest, Katherine Grace Carman, Tullio Jappelli, Franco Peracchi, Frederic Vermeulen, Luc Bissonnette, seminar participants at Tor Vergata and Tilburg University, for extensive and very useful comments on earlier versions.

to grow to be 20% of the population by 2030[1]. These numbers will certainly pose big challenges to existing health care systems, asking for economic health care policies aimed at reducing the burden of aging populations on society and at the same time ensuring the availability of health and social services for older persons, promoting their continued participation in a socially and economically productive life. Aging may not be the main factor in driving up rising health-care costs over the coming decades: the demographic shift is also accompanied by a changing health profile, with an increasing incidence of chronic diseases among older persons. This asks for policies aimed at containing the prevalence of chronic diseases associated with population aging, and at dedicating more resources to preventive measures (such as, for example those aimed at reducing smoking and excessive alcohol consumption). There is ample evidence that mortality and morbidity, the relative incidence of a disease or condition that alters health and the quality of life, are inversely related to SES correlates such as income, education, or wealth (Deaton 2002). Moreover recent literature has emphasized the positive relationship between health conditions and SES, the "health-SES gradient" (Marmot 1999; Smith 1999; Banks et al. 2007), and the stylized fact that wealthier individuals (who also tend to have higher income) live longer.

Although most OECD countries aim at ensuring equitable access to health care and offer basic health care to the complete population irrespective of their SES, the utilization of many health care services is associated with SES, and the nature of this association varies across countries with widely varying arrangements in terms of co-payments and deductibles for services and prescribed drug treatments, private health insurance and private health facilities, quality differences across hospitals and other health care facilities, private and public insurance for specific treatments such as dental care, policies for promoting preventive health care, etc.

Most likely the relationships between SES and health care use and the various types of health care services are different. For example, it is likely that the higher the SES, the better one can find one's way in the health care system, obtain a surgical treatment when needed, and the easier it is to obtain a referral to a specialist. On the other hand, general practitioners (GP) are usually more accessible to all individuals, irrespective of their SES. Disproportionate use of specialist care among the higher socio-economic status groups can be due to the association between education and health knowledge, making the higher SES groups better informed about access to and usefulness of care. Health itself also plays a role here, since the fact that low SES is associated with poor health implies that the needs of health care are higher for the low SES groups. Social policy initiatives are needed to provide access to health care on the basis of need and in order to gain control over escalating health care costs.

By analyzing the relationship between SES, health, and health care use for a variety of developed countries, this paper addresses several questions: What is the nature of the relationship between SES and health care use among the 50+ population? Does the relationship vary with different types of health care services, such as primary care, specialist care, or in- and outpatient care in a hospital? What are the socio-economic factors driving

[1]US Department of Health and Human Services, Administration on Aging, USA (2000), www.dhhs.gov.

the use of health care services: income, wealth and/or education?

While the policy relevance of the relationship between SES and health care utilization seems obvious and is emphasized in the existing literature on the debate on “health equity” (cf., e.g., Oliver and Mossialos 2008), it should be mentioned that there is an ongoing debate on the theoretical and operational targets. Sen (2002) discusses health equity in the broader framework of social justice, and argues that since health is central to not only quality of life but also the ability to do what one has reason to do, health equity is crucial for social justice and equitable access to health care is more important than, for example, equitable access to luxury consumption.

Although there seems to be general consensus about its importance, there is an open debate on what health equity means and what the targets are that should be aimed at. Oliver and Mossialos (2008) mention three principles of equity in health and health care: equal access to health care for those in equal needs; equal utilization of health care for those in equal need; and equal (or, rather, equitable) health outcomes. They conclude that only the former is a reasonable policy target, but what is meant by equal access and equal need is not well-defined. Moreover, access to health care is hard to measure, which is why the focus is often on equal utilization of health care services as an observable proxy. Differences in preferences may well imply that equitable access does not lead to equitable utilization.

Health and health care equity has often been seen as in conflict with health care efficiency. Culyer (2006) argues that there is not necessarily a conflict between the two. He uses the concept of an efficiency frontier in health production – health care efficiency implies that health care production must reach the Pareto frontier such that it is impossible to improve health care services for one group without harming another group. Health care equity has implications for which Pareto efficient allocation is attained. The debate agrees upon the fact that this point must imply some basic level of health care utilization for everyone who needs it irrespective of their SES, but not on what this basic level exactly is.

The contribution of our paper is empirical and determined by the nature and quality of our data, in the spirit of earlier studies by, for example, Van Doorslaer et al. (2006), who also focus on the relationship between SES and health care utilization keeping the need for health care constant. We consider health care utilization as a proxy of health care access, since we have data on the former and not on the latter. We investigate the mechanisms that lead to a relationship between SES and health care utilization and often interpret differences in utilization as differences in access.

The remainder of this paper is organized as follows. In the next section, we provide a conceptual framework and the empirical strategy it implies. In Section 1.3 we discuss the main features of the health care systems in the countries we study. Section 1.4 describes the data and Section 1.5 presents the associations between SES (measured by income) and health care utilization by country. Section 1.6 links the findings in Section 1.4 to those in Section 1.5 to analyze the implications of health care policy for the the association between SES and health care utilization. Section 1.7 concludes.

1.2 Framework

The relevant framework is the model of Grossman (1972) and its extensions; see, e.g., Grossman (2000). While the original study presents a precisely defined model in which theoretical predictions are possible, we focus on the extended framework which adds empirically relevant realistic features, though at the cost of reducing its value for using the theory to predict empirical relationships. Individuals maximize lifetime utility, where utility in a given period depends upon consumption and the stock of health. Health has the nature of a capital good, which deteriorates over time but can be increased by investments, requiring health inputs. The main inputs are health care (preventive or curative) and health related behaviors ((not) smoking, (not) drinking, exercising, healthy diet, etc.). The marginal return on investment in health care depends upon the current status of health, which is why most people seek health care if they have a health problem.

The demand for health care can therefore be seen as an input demand function. It will depend on the (effective monetary) consumer price of health care and, if this is nonzero, on the available income, since the individual has to trade off investing in health against consumption. The effective price depends on co-payments and may be low in case the individual has health insurance. There are also other, non-monetary, costs involved with seeking health care, particularly the time needed to acquire health care (opportunity costs of time, which will be particularly relevant for workers, but also the disutility of spending time in waiting rooms). Thus even if the effective monetary price for the consumer is zero due to health insurance, seeking health care comes with a cost. In addition, the demand for health care will depend upon (and probably decline with) health, since the marginal return will depend on (and decline with) health. In principle the marginal return of health care investments may also depend on other inputs such as (not) smoking or exercising, but keeping health constant it is not so clear whether this effect should play a big role or what sign it should have. Finally, the demand for health care will depend upon access to information (though this may be particularly important more for preventive forms of health care that are not commonly known if people are not automatically referred to them by their GP, than for the health care services that are considered here. See, for example, Avitabile et al. 2008).

In this framework, the health care system and health care policy affect the use of health care services by low and high SES groups through several mechanisms. The effective (monetary) costs will be more important for low income than for high income groups. Non-monetary costs such as waiting times may play a larger role for those with high opportunity costs of time (workers, and particularly workers with high SES). Access to information on health care availability will depend on education and social networks. All these features of the health care system can be influenced by health policy, and better understanding these mechanisms can help to adjust the health care system so that it better accommodates health care needs rather than willingness or ability to pay.

What does this theoretical framework imply for our analysis of how SES impacts health care use by adults aged 50+ across countries? First take current health, information access, and insurance status as given. Now consider the effective (out of pocket) price the consumer has to pay, accounting for co-payments. Excluding the unlikely case of a Giffen good, we can predict that demand falls if the price rises, keeping other factors constant. Since prices directly depend on whether co-payments apply and how large they are, this also leads to the prediction that demand is lower in countries with higher co-payments, ceteris paribus.

This, however, does not say much about the relationship with the most important SES index in our analysis, which is (household) income. The question is how the income effect varies with price. If the effective price is zero, the use of health services is determined by non-monetary factors only, and we expect that the income effect is close to zero. If prices are positive, however, the sign of the income effect is theoretically undetermined without making assumptions on the form of the utility function, and empirical evidence is needed. To the best of our knowledge there is no direct empirical evidence on this, but we expect the income effect to be positive if the user price is positive, and larger if the effective price is higher. This implies that the income sensitivity of the demand for health care is larger for higher effective prices. This leads to the empirical prediction that the income gradient is larger for services and in countries where co-payments are substantial. Moreover, we expect that average effective prices are lower in countries where the health care system is to a large extent publicly funded. This leads to the prediction that there is a negative relationship between the SES gradient and the share of public health spending in gross domestic product (GDP).

On the other hand, health care services may also be costly in terms of time. In particular, for workers, the opportunity cost of time spent in, e.g., waiting rooms will increase with the hourly wage rate. This effect might dominate if the effective monetary price is low. That is, demand for health care might actually fall with SES, particularly among the younger part of the 50+ population who are often doing paid work, and in countries where waiting times in hospitals, emergency rooms, or doctor's offices are long. In any case, the compensating effect of the opportunity cost of time leads to the prediction that the SES gradient will be lower for workers than for non-workers.

Other supply side factors may also affect the use of health care and its income gradient. In particular, the way in which general physicians and specialists are remunerated differs across countries. Sometimes they get a fee for each service, sometimes for each patient, and sometimes a fixed salary. This may influence their advice to patients, and patients in different socio-economic groups may cope with this in different ways (see, for example, Fabbri and Monfardini 2002). For example, it seems plausible that higher socio-economic groups are better able to force doctors to make judgements on the basis of medical grounds rather than their own financial interest, implying that the effect of the remuneration will vary with socio-economic status, or, equivalently, the effect of socio-economic status will vary with the remuneration system.

Public or private health insurance may obviously matter a lot for the effective price of health care. If everyone is fully insured for all health care services and co-payments are always zero, the effective price is zero. While if co-payments are substantial and many services are not covered by the insurance, the effective price can be quite high. In most European countries and in the US, the system is somewhere in between these two. Moreover, costs may be different for different types of care (GP care, specialist care, hospital visits, etc.), and this is one of the reasons why we model each type of care separately.

We encounter a complication when examining the health stock itself. Health is positively correlated with SES. Since health is likely to negatively affect the demand for health care, analyzing the relationship between health care demand and SES without controlling for health will lead to lower (more negative or less positive) estimates of the effect of SES on health care use than if health is controlled for – the lower SES groups demand more care because they need it more (or in terms of the theoretical model, because its marginal return is higher), and not because of their lower SES as such. It therefore seems better to control for health in the analysis. This is also in line with what we want to measure: health care equity refers to equitable access to health care for *those in equal need*, i.e., for those with the same health condition. But this of course raises the issue that health can be affected by past health care (and health behavior) choices – health is quasi-fixed in the short run, but depends on the individual's choices in the long run.

What are the implications of the theoretical framework for the empirical strategy? We run probit regressions explaining health care utilization from SES indicators (income, in the benchmark model), and the SES measure interacted with country dummies, to examine whether the hypotheses formulated above are supported or not. Complications arise because we want to control for various factors: health behavior, information about health care services, and health[2].

As argued above, it is not a priori clear whether variation in health behavior would affect our findings, and if so, in which direction. We therefore do not incorporate health behavior in our main estimations. As a robustness check, however, we also estimate a version of the model that includes controls for health behavior (which are available in our data). This ignores the fact that health behavior may be potentially endogenous because it is a choice of the individual. We lack the appropriate instruments to take that into account.

Information access is difficult to measure. In our main model, we do not incorporate it in the regression but keep it in mind when interpreting the results. For example, if we find a positive relationship between health care use and SES, one potential explanation is that high SES groups have more access to information.

We also do not have the data to account for the endogeneity of health. But since controlling for health (i.e. health care needs) is crucial in our context, we control for health in the main analysis and thereby account for the potential endogeneity problems in interpreting the results as in Maurer (2007). Following Van Doorslaer et al. (2006), we compare results that control for current health with results that do not. As an intermediate strategy, we also consider specifications that only control for a limited set of health variables that are plausibly exogenous (such as whether the doctor has ever told the respondent he or she has cancer, arthritis, etc.).

[2]We do not incorporate voluntary health insurance (VHI). The theoretical framework applies to the price of health care conditional on the insurance that an individual has. But VHI is often the own choice of the individual, and this choice may be related to the individual's preferences for health or health care (cf., e.g., Jones et al. 2006). It may be fixed in the short run but not in the long run.

1.3 Health Care Systems in Europe and the US

Important cross-country differences exist with respect to the financing and delivery systems of health care. There is no generally accepted classification of health care systems: they are usually categorized according to their financing, but this is only one aspect of a health care system. The characteristics summarized in Tables 1.1 and 1.2 give some insight on institutional differences which may have an impact on cross-country differences in health care utilization by income level and can be of relevance when interpreting our results presented in Section 1.6.

Table 1.1 summarizes some of the characteristics of health care systems in the US and the European countries which we analyze (Austria (AT), Belgium (BE), Denmark (DK), France (FR), Germany (DE), Greece (GR), Italy (IT), the Netherlands (NL), Spain (ES), Sweden (SE), and Switzerland (CH))[3]. Table 1.2 shows the type of financing and scope of the health care system in the United States in more detail. We can broadly divide countries in groups according to the organization of their health care system. The first group includes countries (Denmark, Greece, Italy, Spain, and Sweden) characterized by public health care systems (National Health System – NHS) mainly financed by taxes and providing for almost universal coverage (Beveridgean systems). In the second group are countries (Austria, Belgium, France, Germany, the Netherlands) whose health care systems are mainly financed by social contributions (Social Health Insurance – SHI) based on individual income level and which are based on coverage by social security or sickness funds (Bismarckian systems). Switzerland has a "Private mandatory insurance" system (since 1996) financed through premiums; it guarantees universal coverage by compulsory (and publicly subsidized) private health insurance. The insurance premium varies by region but is independent of income and risk.

The US is the only OECD country where voluntary health insurance is the main system for most of the population. This country has a considerable share of the population without insurance coverage: according to the Census Bureau's 2005 Current Population Survey (CPS), there were 45.8 million uninsured individuals in 2004, or 15.7% of the civilian non-institutionalized population. On the other hand, almost the complete US population of ages 65 and over automatically has access to Medicare so that this part of the population is covered by a universal public health care system. In the other countries considered in this study, some population groups buy private health coverage because either they are not eligible to public coverage or they can choose to opt out of it. This is the case, for example, for the Netherlands, where a third of the population is not eligible to public health insurance coverage, and Germany, where employees with annual earnings over €45,900 and their dependants can choose to opt out of the statutory health insurance scheme. In Belgium and France, the insured have to pay different co-payments depending on the type of service, while in other countries visits to public sector doctors are free at the point of delivery (Denmark, Germany, Greece, Italy, and Spain).

Secondary care rules vary from country to country: a gate-keeping system that requires

[3]The tables refer to 2004, the year in which our micro data were collected.

the authorization of referrals to specialists by a designated primary care provider is active in some countries. However, in some countries gate-keeping can be sometimes bypassed through emergency departments of hospitals (like in Spain), whereas in other countries it is often not enforced (like in Italy and Greece). In the US there is no gate-keeping system for those aged 65+.

General practitioners are paid by capitation in Denmark, Italy, and the Netherlands; by salary in Greece, Spain and Sweden, and on fee-for-service basis in the other countries (OECD 2004). Under a capitation system, doctors are paid a fee for each patient registered with them; under a fee-for-service system, doctors are paid on the basis of the service provided; and under a salary system, doctors are employed by the state or the insurer with a salary that does not directly depend on the number of treatments or the number of patients. Remuneration of specialists is differentiated across types of specialization, but the data that we have do not allow distinguishing among these types of specialist. Specialists working in public hospitals in the European countries in the Survey of Health, Ageing and Retirement in Europe (SHARE) are mostly salaried, whereas in the US they are paid on a fee-for-service basis.

Specialist consultation requires some co-payments in most of the countries considered. In Italy a flat rate payment is required for public consultations and outpatient visits; in Spain specialist consultations are free at point of delivery. In Greece consultations are paid out-of-pocket, since private financing is very high. In the US co-payments do not apply to those aged 65+, who are covered by Medicare.

Unlike GP and specialists services, dental care is not publicly provided: dental visits are usually financed out-of-pocket, being paid the full cost in Italy, the Netherlands, Spain and Sweden, and financed through co-payments or co-insurance in the other countries.

1.4 Data

Van Doorslaer et al. (2000) compare the SES gradient in several countries using nationally representative country specific datasets. They acknowledge the potential drawback that measures of health care use, SES, health or other controls may not be comparable across countries, and emphasize the usefulness of having harmonized international data sets to avoid these potential comparability problems. For a selected set of European countries in the European Community Household Panel (ECHP), Van Doorslaer et al. (2006) analyze the relationship between the use of primary and specialist care and SES, controlling for health. Their analysis covers the complete adult population. They find that health care use increases with SES if health is controlled for, particularly specialist care.

The first wave of the SHARE data that have become available for eleven European countries, in combination with the Health and Retirement Study (HRS) data for the US[4], offers a unique opportunity for a richer analysis of the population aged 50 and over. First,

[4]See Appendix A for more information on these data sources. A similar source of data (ELSA) exists for England, but this unfortunately does not contain the information on utilization of health care services that we analyze in this paper.

these data sets provide detailed information on health care use, including specialist visits, dental care, and in- and outpatient treatment in hospitals. Second, they contain extensive information on SES, with harmonized data on education, income, and wealth components. Third, they allow controlling for a rich set of objective and subjective health variables. Therefore SHARE and HRS represent unique data sets for the analysis of the relationship between human capital and SES on the one hand, and the use of health care facilities on the other hand, accounting for the health-SES gradient by controlling for health.

This paper uses data from 2004[5]: wave 1 of SHARE (release 2.0.1) for Europe, and wave 7 of the HRS for the US. We use data from the eleven countries that contributed to the 2004 baseline study in SHARE[6]: Austria, Germany, Denmark, Spain, France, Greece, Italy, the Netherlands, Sweden, Switzerland and Belgium. The study sample is restricted to adults aged 50 and older and we dropped observations with incomplete information on background variables[7]. Our final sample counts 26,563 individuals for SHARE and 19,084 individuals for HRS.

1.4.1 Utilization of Health Services

Health service use is measured by the following questions: "During the last twelve months[8], about how many times in total have you seen or talked to a medical doctor about your health?"; "How many of these contacts were with a GP or with a doctor at your health care centre?"; "During the last twelve months, have you consulted any of the specialists mentioned on card 12?"; "During the last twelve months, have you seen a dentist or a dental hygienist?". Similar questions were asked for inpatient and outpatient care. In this paper, we focus on the binary variables of using a given type of service at least once (variable coded as 1) or not at all (variable coded as 0) during the past 12 months[9].

Figure 1.1 shows a cross country comparison of the use of health care services by income class, based upon our samples from SHARE and HRS. HRS does not distinguish between GP and specialist visits, and only provides information on "doctor visits" (which includes GP, specialist, and outpatient visits). Therefore, GP and specialist use by income class are provided for the SHARE countries only.

Figure 1.1 shows highly differentiated pictures of health service utilization rates across countries and across health services, irrespective of income class. The fraction of the 50+ population visiting a GP at least once varies across SHARE countries from hardly more than 60% in Greece to almost 90% in Belgium and France, three countries that all have almost complete coverage of their population by the public health care system. Differences

[5]Income is collected as gross income in SHARE wave 1 and as net income in wave 2. We focus on wave 1 to make income directly comparable with the HRS, which also includes gross income.

[6]See Börsch-Supan et al. (2005), Börsch-Supan and Jürges (2005) and www.share-project.org for details on the SHARE data.

[7]The sample design implies that individuals younger than 50 years with a partner of 50 years or older are also interviewed. These respondents are not included in our analysis.

[8]In the HRS, the questions refer to the last two years instead of the past twelve months.

[9]Two years in the US. This difference is not corrected for in the descriptive statistics but is captured by the US dummy in the regressions.

for other services are even larger. The use of specialist services seems exceptionally low in Denmark, being less than 20%, and quite high in Belgium and Germany, although coverage by public health care is less complete in Germany than in many other countries. Inpatient and outpatient services seem particularly popular in the US. It must be kept in mind, however, that the US question refers to a two year period while the SHARE question refers to the past 12 months. This may explain the difference for inpatient services but cannot explain the difference in outpatient services, where the US utilization rate is more than twice as large as the utilization rate in any of the SHARE countries. Particularly in outpatient services, there is also large dispersion within Europe. Such dispersion is also found in dentist care, which is much less common in the southern European countries than in the US and the rest of Europe. Denmark and Sweden have the highest proportion of dental care users.

There is also substantial variation in the income gradients across health services as well as countries. The use of doctor, inpatient, and outpatient care does not increase with income in most countries, in accordance with the fact that for basic health services most countries have achieved close to universal coverage of their population at relatively low and sometimes zero financial cost. In fact, the association between income and inpatient or GP care seems negative, which is probably due to the fact that the low income groups are less healthy and more in need of health care. This finding is in line with earlier studies like Van Doorslaer et al. (2000, 2006). For specialist and outpatient care, no clear positive or negative association is found. The only exception here is dental care – its use clearly rises with income in all SHARE countries and in the US.

1.4.2 Demographics and Health Variables

In this section we define the explanatory variables that we include in the model. Tables 1.3 and 1.4 show descriptive statistics of our working sample. The demographic variables included in the analysis are age, gender and marital status. Age is grouped into 5-year bands: 50–54; 55–59; 60–64; 65–69; 70–74; 75–79, and 80+. Marital status is categorized as married or not married (which includes "living with a partner" and "living as a single").

SES is included in the model as household income, adjusted for household size (that is, divided by the square root of the number of household members). Income is measured as the log of gross annual household income for 2003 and is derived from disaggregated income sources including labour and non-labour income, transfer income, investment income, benefit income and pension income (gross total individual income of each respondent, sum of the gross incomes of other household members and other benefits, capital assets income, excluding rent payments received and imputed rents). All amounts are in thousands of PPP-adjusted dollars[10].

This paper focuses on the SES gradient in terms of log income, considered a short term indicator for SES. In a sensitivity analysis, we also look at other SES indexes which can

[10]PPP exchange rates are taken from the OECD web-site:
http://www.oecd.org/document/47/0,3343,en_2649_34357_36202863_1_1_1_1,00.html#historicalppp.

be seen as long-term indicators of SES, in particular education level and household wealth. Education level is defined according to the ISCED-97 harmonized coding for international comparisons[11], with the following three categories: non-advanced qualification, high school qualification and advanced qualification. Wealth is defined as household net worth in thousands of PPP-adjusted dollars, adjusted for household size.

Health care equity is often defined as equal access for those with equal need. The need for health care services is incorporated through several indicators of the respondent's health. We control for self-reported health status (SPHS, coded as 0 "very good/excellent" and 1 "less than very good"), and more objective measures of health. The variables "limitations with activity of daily living" (ADL) (such as dressing, bathing, or getting in and out of bed) and "mobility limitation" (MOBILIT) indicate the extent to which individuals consider themselves physically handicapped. Both variables are reclassified into two categories: no limitations with ADL (or MOBILIT) and one or more limitations with ADL (or MOBILIT). In addition, we include a variable indicating whether or not the respondent has two or more chronic diseases (CHRONIC), based upon questions that ask whether the respondent suffers from a number of chronic diseases[12]. Finally, we control for three dummies related to weight and height: underweight, overweight, and obese; the benchmark group is those of normal weight. These dummies are based upon the body mass index (BMI): weight (in kilograms) divided by height (in cm) squared. BMI categories are as follows: BMI $\leq$ 18.5 (underweight); 18.5 $<$ BMI $<$ 25 (normal weight); 25 $\leq$ BMI $<$ 30 (overweight); BMI $\geq$ 30 (obese).

1.5 The Income Gradient of Health Care Use

In this Section, we describe the income gradient of health care use using probit models explaining the yes/no answer to the questions whether respondents have used the type of health care service at least once in the past twelve months (two years in the US). In each probit model, the independent variable of interest is log household income. The models are estimated separately for each type of care and for each country. We present the income slopes as a descriptive tool.

We distinguish three models in each case, differing in the additional factors that we control for. The first model does not control for any additional factors, the second controls for basic demographics (age, gender, marital status); the third specification adds the controls for health. Tables 1.5 – 1.10 present the country specific estimates of the coefficient on log income for each type of health care that we consider for each of the three models.

As a sensitivity analysis we check what happens when we also control for education level in the third specification of the model. The estimated effect of education on health care use is significantly positive for specialist and dentist visits, and the estimates of the

[11]See for details on ISCED coding: www.uis.unesco.org/ev.php?ID=3813_201&ID2=DO_TOPIC.

[12]The number of chronic diseases is a count of the following diseases an individual might have: heart problems, high blood pressure, high cholesterol, cerebral vascular disease, diabetes, lung diseases, asthma, arthritis, osteoporosis, cancer, stomach ulcer, Parkinson disease, cataracts, hip fracture or femoral fracture.

coefficient of log income hardly change. Regarding doctor visits, middle and high education coefficients are positive and significant for SE, IT, GR, and US (for FR and DK only for high education). Regarding the other health care services, the education controls are generally not significant, except for the US, where they are always positive and significant. Overall, the education effects are usually in line with the log income effects but significance levels sometimes differ (for example, for doctor visits, education level coefficients are significant whereas log income is not).

In the same way, we estimate the third specification of the model adding controls for wealth (assets). This has no effect on the income coefficients and the coefficients on the wealth variables are not significant.

Similarly we test the robustness of the results with other SES measures (education level and assets) to support the choice of log household income as the measure of SES. We estimate each probit model first with assets[13], then with educational qualification[14] as independent variables reflecting SES (instead of log income). Whenever the coefficients on assets or on education qualifications are significant, the sign is the same as for log income, leading to results that are qualitative similar to those obtained for log household income. Therefore the main conclusions remain unchanged when log income is replaced by another measure of SES or when more than one SES measure is used.

Tables 1.5 – 1.10 present the estimated marginal effects at the country specific means. They can thus be interpreted as 100 times the number of percentage points the probability of using the service would increase if income increased by 1%, keeping constant all other explanatory factors included in the model.
The general picture of Tables 1.5 – 1.10 is that the SES gradients are very heterogeneous across health care services and across countries, but less across model specifications. Once basic demographics are controlled for, controlling for health often raises the income coefficient (from negative to zero, or from zero to positive, etc.), in line with the notion that lower income groups have more health problems, and health problems obviously increase the use of health care.

Table 1.5 presents the results for doctor visits, combining GP, specialist and outpatient visits[15]. Particularly when health is controlled for, the income slope is positive in six countries, including the US, but there is large variation in size and significance levels across countries, and in some countries the income slope is essentially zero. To understand these differences, it seems better to look at the more disaggregate level where GP services, specialist services, and outpatient services are distinguished, presented in Tables 1.6, 1.7 and 1.8, respectively.

For GP use (Table 1.6) the sign of the income effect is negative or insignificant for the majority of the countries if health conditions and demographics are not controlled for. Controlling for health conditions changes the picture, with insignificant income effects in all countries except SE where, surprisingly, the income slope becomes significantly positive

[13]In PPP-adjusted dollars measured at household level, corrected for household size.
[14]As defined in the previous section.
[15]GP and specialist visits are distinguished in the SHARE data only, nevertheless we are able to analyze outpatient service use separately both in SHARE and HRS.

and quite large, whereas in a country like DK, with a health care system which is in many respects similar to that in SE, the slope is zero. Part of the explanation suggested by the theoretical framework might be that DK has no co-payments while in SE very modest co-payments exist (Docteur and Oxley 2003, pp. 54–55). Other possible explanations for the differences might be differences in the extent to which health care is publicly funded and whether the GP acts as a gate-keeper to other forms of care.

The picture for specialist use is quite different (Table 1.7): the income gradient is positive and significant in most SHARE countries, with or without controls for demographics and health conditions. Particularly in CH, the income gradient of specialist access seems very large, in line with what we saw in Figure 1.1. In SE, DK, and ES, the income effect is insignificant but still positive once demographic characteristics and health conditions are controlled for.

For outpatient use (Table 1.8) we find significant positive income effects for the US and SE. In the US, outpatient care is more important (both in absolute terms and compared to inpatient care) than in the European countries (see Figure 1.1) and it seems that particularly the richer groups make much use of this. An explanation for this may be that co-payments on typical outpatient hospital treatments like X-rays and pathology are higher in the US than in Europe (Docteur and Oxley 2003, Table 7). Co-payments cannot explain the strong positive income effect in Sweden; perhaps this is because outpatient care can substitute specialist care in this country, since Sweden is one of the few countries where we find no SES gradient in specialist care (see Table 1.7).

The results for inpatient care are presented in Table 1.9. Without controls for demographic characteristics and health conditions, income effects vary from significantly negative in the US, DE, and SE, to insignificantly positive in the other European countries. Once all the controls are added, the income effect is usually small and positive (with a few exceptions) and never significant at the 5% level. According to Docteur and Oxley (2003), most countries have no or a modest co-payment for every day spent in the hospital, except in the US where co-payments can be substantial. Possible explanations for a positive effect of income might be that hospitals get higher fees for treatments of higher income groups covered by different type of insurance (cf. Van Doorslaer et al. 2000) or that access barriers (such as information acquisition or an appointment with a specialist) mainly hamper the lower income groups[16].

The strongest effect of income is in dentist and dental care use (Table 1.10): we find a positive effect of income for all countries (even when controlling for health, except GR), and the effect is significant at the 5% level in nine of the twelve countries. The costs of dental care are often not covered by basic insurance in most of the countries we considered. Higher income apparently leads to easier access and better chances to purchase an adequate and affordable level of private coverage. It is interesting to compare the ranking of the income gradients here with the ranking of the costs of a standard treatment - dental fillings across

[16]Stargardt (2008) compares the costs of a hip replacement, a common operation for the elderly, across selected countries. He finds that Spain is much cheaper than other countries, whereas Italy is quite expensive. There is only a weak correlation between these costs and the income effects in the same countries, which probably should be expected since patients hardly ever end up paying this themselves.

countries, given by Tan et al. (2008). They find the highest costs of treatment in England, Italy and Spain, and much lower costs in Germany, the Netherlands and, particularly, Denmark and France (unfortunately they provide no information on the other SHARE countries). If higher costs of treatment lead to higher prices for health care consumers (in the form of co-payments or because treatment is not covered) one would expect a positive relationship between the income effect and the cost. This is not what we find for ES, which has rather low income effects compared to the other countries considered, though it is one of the most expensive countries for dental care. IT has a higher income effect than all the countries mentioned in the study by Tan et al. (2008), except DK.

1.6 Health Care Use and Health Policy

In the previous section we found substantial differences in the relationship between income and health care utilization across countries. In this Section we analyze the cross-country correlation between the income gradient that we estimated in the previous section and differences in health care policy across countries. Table 1.11 presents the characteristics of the health care systems. These are the policy instruments that can affect the income gradient of health care services. The variables per capita total expenditures on health care and per capita public health expenditure[17] (here defined as percentage of total expenditure on health) are measures of health care funding. How this affects the income gradient obviously depends on how the funding is allocated. More public health expenditures can benefit the poor if they increase access to basic services, but they may also be used for less basic services that are more than proportionally used by the higher income groups. Per capita health expenditures per year vary from slightly less than US$ 2,000 in GR to more than US$ 6,000 in the United States. They are much lower in the Southern European countries than in the rest of Europe and much higher in the US than in any of the European countries – the difference between the US and Switzerland, the European country where these expenditures are highest, is still more than 50%[18].

The third macro-variable reflecting differences in health care policy is a dummy for whether the general physician acts as a gate-keeper (GK) for access to other types of health care such as specialist care (excluding dentists). We expect that general physicians do not base their referral decisions on income and therefore may reduce the importance of other determinants of using specialist care, such as its price. Since visiting a GP does not substantially depend on income, gate-keeping may also reduce the gradient due to information access: the information on specialist services provided by the GP will be less related to the patient's SES than information collected by the patients themselves. On the other hand, those who are more informed may push their GP harder to refer them to a specialist. Moreover, it seems plausible that gate-keeping increases the time effort needed to obtain specialist care, making it less attractive for individuals with high opportunity costs, e.g. higher wage earners. All these scenarios lead to the hypothesis that gate-keeping

[17]Expressed in US$ using purchasing power parity (OECD Health Data, OECD 2007).

[18]Similar results are obtained if the ratio of health care expenditures and GDP is used.

reduces the income gradient of specialist care and other types of care to which gate-keeping applies, such as many types of inpatient care which often start with referral to a specialist.

The relationship of gate-keeping with outpatient care is not so clear; some outpatient care requires referral but other types do not (particularly emergency care). We expect that gate-keeping increases utilization of GP services, and to the extent that higher SES groups want more specialist services, that gate-keeping also has the indirect effect of increasing demand for (referrals through) GP visits.

Table 1.11 also shows the more common type of remuneration for doctors in every country[19]: fee-for-service (F) where doctors are paid on the basis of the service provided, capitation (C) where doctors are paid a fee for each patient registered with them, and salary (S) where doctors are employed by the state or the insurer with a salary that does not directly depend on the number of treatments or the number of patients. In countries with a fee-for-service payment scheme, doctors may tend to lengthen the duration of the treatments, which makes visits to a specialist more likely than in countries where other types of remuneration apply.

As discussed in Section 1.2, under plausible assumptions about underlying preferences, co-payments are expected to increase the SES health care utilization gradients since they increase the effective price of the services. Co-payments vary across services, sometimes refer to amounts, and sometimes are a percentage of the total cost of a specific service. As a consequence, specifying a co-payment amount for each broad type of health services in our analysis is not possible. We therefore only work with a dummy variable on whether co-payments apply. Table 1.11 shows that co-payments for GP care are common in five out of twelve countries considered. In all these countries except GR, co-payments also apply to specialist and in- or outpatient services, while there are several countries where co-payments apply to some of these services but not to GP care. Co-payments are very common for dentist services – DE and NL are the only countries where they do not apply.

For the empirical analysis, we ran similar probit models as in the previous section, pooling all countries and interacting log income with the five policy indexes discussed above defined at the country level[20]. Furthermore we included only one or two macro-variables at a time. The identifying assumption in these models is that the cross-country differences in *income slopes* are exclusively driven by the macro-variables included in that regression, while differences in the *levels* of health care utilization can also be due to the other macro-variables and other factors (economic, institutional, or cultural). Unfortunately,

[19]We use the same remuneration types as Jimenez-Martin et al. (2004), where the types are defined for doctors and GPs.

[20]Per capita total health expenditure (PCPHE), per capita public health expenditure (PCPUBHE), doctors type of remuneration (CAP, SAL, with FFS being the base category), a dummy for whether the GP acts as a gate-keeper (GK; gate-keeping refers to a system where the primary care provider coordinates patient care and refers patients to specialists, hospitals and other medical services), and a dummy for whether the health service requires co-payments (COPAYS). We also included country dummies, but no interactions between log income and country dummies. The country dummies also accounted for the difference between the US and the SHARE questions: the former asks about using the health care services in the past two years, the latter about the last 12 months. While this would affect the levels, we assumed it had no effect on the income slopes.

the number of countries appeared not to be large enough to disentangle the effect of each macro-variable on the income gradient separately, neither in a multivariate regression context nor when including one macro variable at the time – we tried both specifications but results were inaccurate and insignificant (details are available upon request).

Instead, we follow a more descriptive approach, showing how the income slopes relate to the different macro-variables described above. Figure 1.2 shows the results. It should be kept in mind here that the correlations are based upon 11 or 12 points (11 or 12 countries, depending on whether the US is included or not) only, and can be driven by a few of these countries. The most salient finding is a positive association between aggregate health care expenditures and the income gradient of the use of health care services. Positive association is found for doctor visits, specialist services, outpatient services, and dental care, irrespective of the measure for public health expenditures that is used. This suggests that the extra services provided in countries with relatively large health expenditures mainly benefit the richer part of the (older) population. For GP visits, the sign of the association depends on which measure of health care expenditures is used. For inpatient services, we find a negative but very weak association. For these services, the fact that larger health care expenditure may increase access for the poor could compensate the effect of providing extra services mainly used by the richer part of the population.

Gate-keeping is positively associated with the income gradient in doctor visits, GP visits, and outpatient services, but negatively with specialist visits. The latter effect is as expected, since the need of referral through a GP may make a specialist visit more dependent on medical need and less on other factors such as income or access to information networks. The positive associations with GP visits are in line with the fact that their greater demand for specialist services induces high income groups to visit their GP if they need a referral. The positive association with outpatient services may (again) be explained by substitution of specialist visits by outpatient hospital treatment.

The association between co-payments and income is largely as expected. It is positive for doctor visits, specialist visits, outpatient services, and dental care. It is zero or even negative for GP visits and inpatient services. Like the associations with the level of public health expenditures, this is consistent with the notion that specialist, outpatient, and dental care services contain more non-basic "luxury" services where the patients have a choice and make a trade off between costs and benefits. Higher (monetary) costs induced by co-payments are more often an impediment for low income groups than for higher income groups.

1.7 Conclusions

We have analyzed the relationship between income as a measure of SES and the use of several health care services for the 50+ population in the US and a number of European countries. Using a health production framework, we have analyzed the potential income effects and how they vary with prices and other institutional features. This leads to predictions for empirical work – for example, the association between the consumer price and the income effect is expected to be positive, while the effect is predicted to be negatively correlated to quality aspects such as waiting times. Health policies that change the effective price of health care services, or change other factors that make the services less or more accessible to low or high SES groups, are therefore expected to influence the relationship between the use of the health care service and socio-economic status. Since equal access to health care services for people with equal health problems is an explicit policy target in many countries, it is important to analyze which aspects of health policy lead to such a gradient.

We find clear evidence of a positive income gradient for several health care services, particularly for specialist visits, outpatient services, and dental care. These are also the services for which we find the clearest positive association between the income gradient and public expenditure on health care at the aggregate (country) level. These services probably contain more non-basic services than the other types of health care use that we consider, implying that whether or not to use them is a choice of the consumer. For low income groups, the cost may weigh more heavily and limited access to information about available health care possibilities may play a role as well. In any case, our results suggest that countries with higher public health expenditures do not automatically get closer to the policy goal of health care equity, i.e. equal access for those with the same needs. On the contrary, our results suggest that the extra services that the extra money can buy disproportionally benefit the richer part of the (older) population.

Validating the theoretical predictions requires more detailed insight in the prices and characteristics of various types of health care services than is currently available. There is interesting recent work on price indicators based upon specific treatments (Busse et al. 2008) but this covers only a limited set of countries and focuses more on the production costs and reimbursements to doctors and hospitals than on the prices for the patients. Additionally future research on what is covered by which insurance is needed.

Table 1.1: Characteristics of Health Care Systems in SHARE Countries and US (2004)

	COVERAGE TYPE	ELIGIBILITY FOR PUBLIC COVERAGE	%THE ON PHI	%THE ON OOP	%POP WITH PUB/MAND	%POP WITH VPHI	TYPES OF PRIVATE COVERAGE
AT	Social Insurance	Almost all labor force participants and retirees are covered by a compulsory statutory health insurance. 1% are without coverage.	7.3	17.5	99.9	0.1; 31.8	Primary (Substitute); Complementary, Supplementary
BE	Social Insurance	Compulsory statutory health insurance includes one scheme for salaried workers and one scheme for self-employed. The latter excludes coverage of 'minor risks' such as outpatient care, physiotherapy, dental care, and minor operations.	n.a.	19.7	99	57.5	Primary (Substitute); Complementary, Supplementary
DK	Public Tax Financed	All population is eligible to public coverage financed by State, County and Municipal taxation.	1.6	15.3	100.0	28.0 (1998)	Complementary, Supplementary
FR	Social Insurance	The social security system provides coverage to all residents. 1% of the population is covered through Couverture Maladie Universal (CMU).	12.7	9.8	99.9	92.0	Complementary, Supplementary
DE	Social Insurance	All employed people (not self-employed) are covered by statutory health insurance coverage. Employers with an income above a threshold can opt out of the social sickness found system.	12.6	10.4	90.9	9.1	Primary (Substitute); Complementary, Supplementary
GR	Public Tax Financed	All population is eligible to public coverage financed by a combination of taxation and social health insurance contributions.	n.a.	n.a.	100.0	10.0	Duplicate, Supplementary

Continued on next page

Table 1.1: continued from previous page

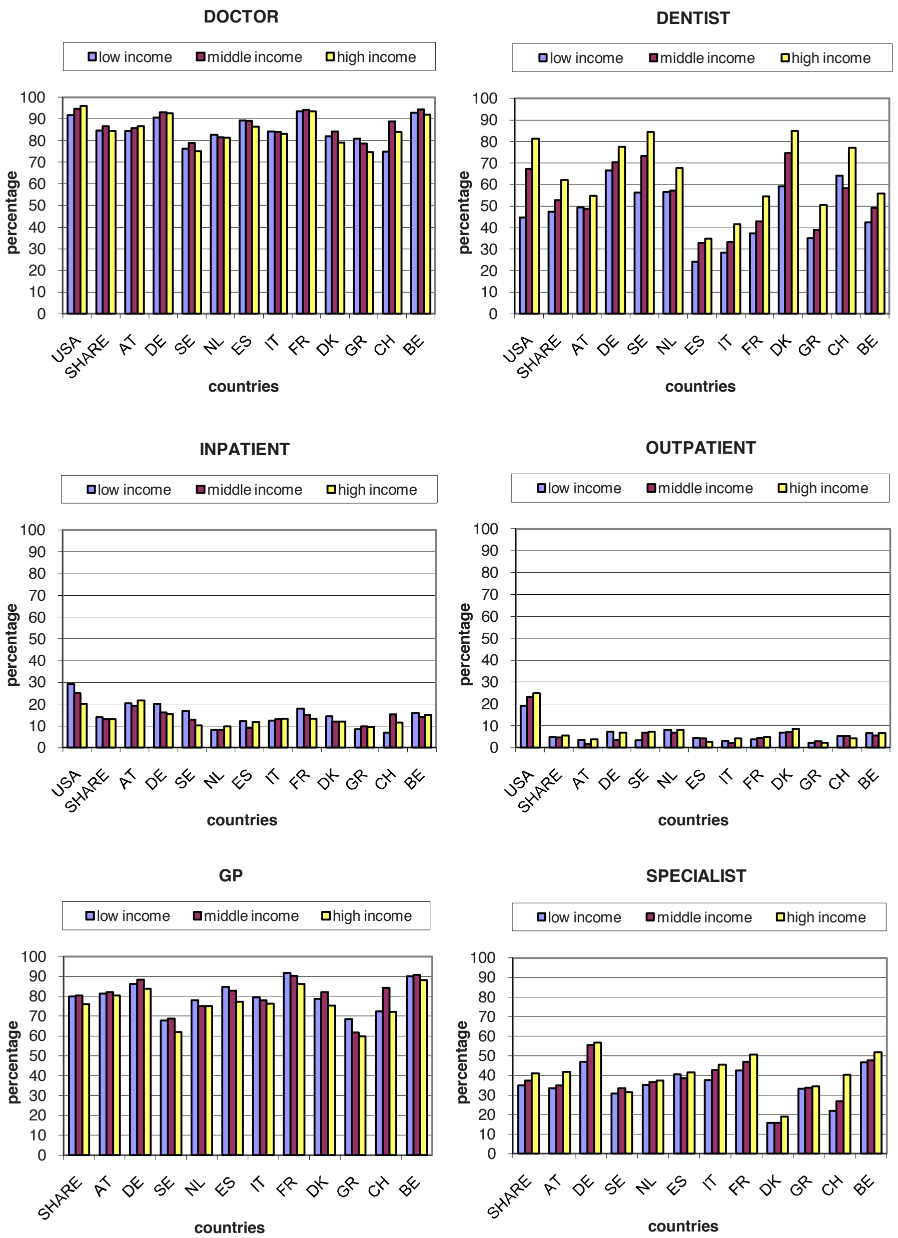
	COVERAGE TYPE	ELIGIBILITY FOR PUBLIC COVERAGE	%THE ON PHI	OOP	%POP WITH PUB/MAND	VPHI	TYPES OF PRIVATE COVERAGE
IT	Public Tax Financed	All population is covered by the National Health Service system, financed by general taxation.	0.9	22.3	100.0	15.6 (1998)	Duplicate, Complementary, Supplementary
NL	Social Insurance	Eligibility to statutory health insurance is determined by income. Individuals above a threshold are not covered (28.9% in 2000).	15.2	10.1	72.0	28.0; 64*	Primary (Principal); Supplementary
ES	Public Tax Financed	Almost all the population is covered by the National Health Service system, financed by general taxation. A minor group of self-employed liberal professionals and employers are uncovered.	3.9	23.6	97.3	2.7; 10.3	Primary (Substitute, Principal); Duplicate, Supplementary
SE	Public Tax Financed	All population is covered by a statutory social health insurance system, financed by local taxes and state grants.	n.a.	n.a.	100.0	negl.	Complementary, Supplementary
CH	Private Mandatory	All permanent residents are mandated to purchase basic health insurance.	10.5	31.5	100.0	80.0	Supplementary
US	Private Voluntary	Individuals eligible to public programs include the above 65 and several disabled (Medicare), poor or near poor (Medicaid) and poor children (SCHIP). Eligibility thresholds to Medicare are set by state.	35.1	13.3	24.7	71.9	Primary (Principal); Complementary, Supplementary

Source: OECD (2004). *Notes:* 'negl.' indicates a proportion covered of less than 1%; 'n.a.' indicates not available; * estimated.
THE: total health expenditures; (V)PHI: voluntary private health insurance; OOP: out of pocket expenditures; POP: population; PUB/MAND: public/mandatory.

Table 1.2: United States: Type of Health Financing and Scope

VOLUNTARY HEALTH INSURANCE	Private Health Insurance schemes financed through employers' and employees' premiums, but about 40% of all employers pays the full premium for their employees; Predominantly middle-class and higher class population.
MEDICARE	Federal health insurance program, financed through taxes (75%) and contributions (25%) paid into Social Security; People aged 65+, people with disabilities, people with End-Stage Renal Disease, also middle-class population.
MEDIGAP	Medicare supplemental health insurance policy sold by private insurance.
MEDICAID	Join federal and state program; People with low income (11%) or with no insurance (15%).

Source: WHO (2004), Country Profiles, www.who.int.



Notes: Weighted statistics based on 2004 SHARE and HRS data.

Figure 1.1: Health Care Use by Income

Table 1.3: Income by Country

Country	Mean	Std. Dev.	Variance	p25	p50	p75
AT	33349.96	32660.75	1.07e+09	14016.92	23764.62	40091.64
DE	37770.06	37923.73	1.44e+09	15074.41	26333.00	47071.98
SE	38381.26	27373.97	7.49e+08	20920.06	31029.11	47179.50
NL	41507.49	38940.71	1.52e+09	16867.77	30720.52	53432.91
ES	20488.52	29284.59	8.58e+08	6,334.11	11786.35	23682.49
IT	22402.76	24733.75	6.12e+08	8,790.07	15341.92	27475.69
FR	37575.93	46804.90	2.19e+09	13412.86	22966.36	41115.70
DK	38521.71	32883.28	1.08e+09	16510.29	30612.65	48494.54
GR	18075.29	17192.18	2.96e+08	8,329.86	13319.43	23183.83
CH	47892.35	44345.83	1.97e+09	16627.12	35185.15	64482.54
BE	38751.76	54288.86	2.95e+09	12768.93	21519.57	43268.27
US	40839.25	70786.93	5.01e+09	13056.00	24878.85	46325.39
Total	36678.92	54398.61	2.96e+09	12662.87	23489.34	43314.55

Notes: Income is measured as gross annual household income for 2003, adjusted for household size. All amounts are in thousands of PPP-adjusted dollars.

Table 1.4: Descriptive Statistics of the Working Sample

Variable	Mean	Std. Dev.	Min.	Max.
logincome	9.962	1.345	0	15.08
assets	288314.8	924719.9	-2719208	5.46E+07
mid edu	0.306	0.461	0	1
high edu	0.296	0.457	0	1
55–59	0.162	0.369	0	1
60–64	0.166	0.372	0	1
65–69	0.162	0.368	0	1
70–74	0.130	0.336	0	1
75–79	0.098	0.297	0	1
80+	0.118	0.322	0	1
woman	0.554	0.497	0	1
unmarried	0.667	0.471	0	1
sphs	0.660	0.474	0	1
adl (1+)	0.126	0.332	0	1
mobilit (1+)	0.561	0.496	0	1
chronic (2+)	0.382	0.486	0	1
underweight	0.015	0.123	0	1
overweight	0.404	0.491	0	1
obese	0.216	0.412	0	1
N	45647			

Table 1.5: Income Gradient of Health Care Use – Doctor (GP, Specialist, and Outpatient)

Country	N	(1)	(2)	(3)
AT	1789	1.041**	1.006**	0.865*
		(0.487)	(0.493)	(0.488)
DE	2899	0.412	0.42	0.715**
		(0.443)	(0.407)	(0.337)
SE	2933	1.58	2.633**	3.768***
		(0.98)	(1.029)	(1.021)
NL	2806	1.195*	1.248*	1.354**
		(0.642)	(0.64)	(0.603)
ES	2164	0.129	0.03	-0.027
		(0.406)	(0.384)	(0.337)
IT	2440	1.044***	0.826**	0.833**
		(0.395)	(0.388)	(0.36)
FR	2880	0.196	0.141	0.157
		(0.339)	(0.318)	(0.243)
DK	1568	-1.548	-0.687	-0.03
		(1.052)	(1.076)	(0.994)
GR	2608	-0.119	0.222	0.352
		(0.536)	(0.526)	(0.502)
CH	929	0.55	0.655	0.966
		(1.015)	(0.987)	(0.925)
BE	3547	-0.069	-0.014	-0.001
		(0.341)	(0.313)	(0.247)
US	19084	0.959***	0.918***	1.094***
		(0.109)	(0.104)	(0.096)

Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.
Marginal effects: all coefficients are multiplied by 100. Base categories: age 50–54; male; married; bmi 'normal weight'. Variables included: Log income (Column 1); Demographic characteristics (age, gender, marital status) (Column 2); Health controls (self reported health status, adl, mobility, chronic conditions, bmi) (Column 3).

Table 1.6: Income Gradient of Health Care Use – GP

Country	N	(1)	(2)	(3)
AT	1789	0.836	0.887	0.736
		(0.562)	(0.572)	(0.567)
DE	2899	-1.030*	-0.668	0.016
		(0.616)	(0.597)	(0.575)
SE	2933	0.611	2.285*	3.853***
		(1.141)	(1.21)	(1.257)
NL	2806	0.671	0.776	0.975
		(0.726)	(0.729)	(0.717)
ES	2164	-0.541	-0.594	-0.641
		(0.509	(0.497)	(0.474)
IT	2440	0.731	0.484	0.591
		(0.463)	(0.46)	(0.453)
FR	2880	-0.958**	-0.994**	-0.678
		(0.48)	(0.465)	(0.417)
DK	1568	-2.257**	-1.151	-0.286
		(1.135)	(1.163)	(1.107)
GR	2608	-1.263**	-0.851	-0.854
		(0.644)	(0.648)	(0.653)
CH	929	-2.358*	-2.501**	-1.973
		(1.242)	(1.232)	(1.218)
BE	3547	-0.299	-0.23	-0.138
		(0.419)	(0.395)	(0.353)

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Marginal effects: all coefficients are multiplied by 100. Base categories: age 50–54; male; married; bmi 'normal weight'. Variables included: Log income (Column 1); Demographic characteristics (age, gender, marital status) (Column 2); Health controls (self reported health status, adl, mobility, chronic conditions, bmi) (Column 3).

Table 1.7: Income Gradient of Health Care Use – Specialist

Country	N	(1)	(2)	(3)
AT	1789	1.886**	1.844**	1.843**
		(0.771)	(0.793)	(0.805)
DE	2899	2.741***	2.662***	2.840***
		(0.859)	(0.882)	(0.903)
SE	2933	-0.49	-0.154	1.101
		(1.143)	(1.198)	(1.242)
NL	2806	2.094**	1.978**	2.452***
		(0.864)	(0.873)	(0.91)
ES	2164	1.320*	1.12	1.174
		(0.676)	(0.685)	(0.715)
IT	2440	2.344***	2.230***	2.585***
		(0.615)	(0.622)	(0.658)
FR	2880	2.617***	2.502***	2.827***
		(0.768)	(0.786)	(0.808)
DK	1568	1.408	1.248	1.974
		(1.098)	(1.2)	(1.28)
GR	2608	1.181*	1.364**	1.587**
		(0.64)	(0.651)	(0.668)
CH	929	6.951***	7.709***	7.778***
		(1.46)	(1.501)	(1.519)
BE	3547	2.105***	2.212***	2.287***
		(0.668)	(0.673)	(0.684)

Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.
Marginal effects: all coefficients are multiplied by 100. Base categories: age 50–54; male; married; bmi 'normal weight'. Variables included: Log income (Column 1); Demographic characteristics (age, gender, marital status) (Column 2); Health controls (self reported health status, adl, mobility, chronic conditions, bmi) (Column 3).

Table 1.8: Income Gradient of Health Care Use – Outpatient

Country	N	(1)	(2)	(3)
AT	1789	-0.155	-0.22	-0.241
		(0.232)	(0.23)	(0.215)
DE	2899	0.442	0.516	0.437
		(0.417)	(0.42)	(0.416)
SE	2933	1.892***	2.203***	2.390***
		(0.691)	(0.728)	(0.736)
NL	2806	0.372	0.431	0.519
		(0.493)	(0.491)	(0.493)
ES	2164	-0.049	-0.035	-0.044
		(0.26)	(0.261)	(0.24)
IT	2440	0.375	0.349	0.358
		(0.268)	(0.267)	(0.244)
FR	2880	0.068	0.049	0.036
		(0.32)	(0.317)	(0.314)
DK	1568	0.966	0.285	0.322
		(0.812)	(0.832)	(0.846)
GR	2608	0.019	0.031	0.05
		(0.202)	(0.186)	(0.185)
CH	929	-0.16	-0.061	-0.068
		(0.598)	(0.598)	(0.486)
BE	3547	0.511	0.48	0.543
		(0.35)	(0.348)	(0.347)
US	19084	1.861***	1.908***	2.515***
		(0.238)	(0.254)	(0.269)

Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.
Marginal effects: all coefficients are multiplied by 100. Base categories: age 50–54; male; married; bmi 'normal weight'. Variables included: Log income (Column 1); Demographic characteristics (age, gender, marital status) (Column 2); Health controls (self reported health status, adl, mobility, chronic conditions, bmi) (Column 3).

Table 1.9: Income Gradient of Health Care Use – Inpatient

Country	N	(1)	(2)	(3)
AT	1789	0.357	0.256	0.087
		(0.631)	(0.649)	(0.648)
DE	2899	-1.608***	-1.391**	-0.98
		(0.606)	(0.614)	(0.602)
SE	2933	-1.495**	-0.678	0.468
		(0.74)	(0.798)	(0.809)
NL	2806	0.696	0.601	0.717
		(0.529)	(0.53)	(0.509)
ES	2164	0.48	0.442	0.438
		(0.448)	(0.459)	(0.441)
IT	2440	0.11	-0.101	0.024
		(0.396)	(0.394)	(0.393)
FR	2880	-0.498	-0.588	-0.353
		(0.527)	(0.538)	(0.543)
DK	1568	-0.829	1.144	2.149*
		(0.849)	(1.068)	(1.115)
GR	2608	0.646	0.692	0.710*
		(0.422)	(0.432)	(0.41)
CH	929	1.235	1.328	1.653*
		(0.994)	(1.013)	(0.944)
BE	3547	-0.168	-0.065	0.009
		(0.459)	(0.461)	(0.453)
US	19084	-2.962***	-2.187***	0.032
		(0.23)	(0.243)	(0.258)

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.
Marginal effects: all coefficients are multiplied by 100. Base categories: age 50–54; male; married; bmi 'normal weight'. Variables included: Log income (Column 1); Demographic characteristics (age, gender, marital status) (Column 2); Health controls (self reported health status, adl, mobility, chronic conditions, bmi) (Column 3).

Table 1.10: Income Gradient of Health Care Use – Dentist

Country	N	(1)	(2)	(3)
AT	1789	0.697	1.420*	1.442*
		(0.77)	(0.809)	(0.816)
DE	2899	3.024***	2.239***	1.750**
		(0.718)	(0.739)	(0.747)
SE	2933	8.853***	7.051***	6.482***
		(1.017)	(1.051)	(1.051)
NL	2806	1.817**	2.365***	2.141**
		(0.823)	(0.853)	(0.861)
ES	2164	1.024*	1.151*	1.141*
		(0.609)	(0.602)	(0.604)
IT	2440	2.812***	3.051***	3.082***
		(0.615)	(0.616)	(0.624)
FR	2880	3.132***	3.193***	2.924***
		(0.763)	(0.784)	(0.789)
DK	1568	7.405***	4.236***	3.956***
		(1.093)	(1.146)	(1.165)
GR	2608	1.286**	1.136*	1.079
		(0.646)	(0.662)	(0.665)
CH	929	5.651***	5.623***	5.299***
		(1.339)	(1.376)	(1.376)
BE	3547	2.964***	3.015***	2.868***
		(0.668)	(0.682)	(0.685)
US	19084	10.936***	9.964***	8.491***
		(0.298)	(0.311)	(0.316)

Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.
Marginal effects: all coefficients are multiplied by 100. Base categories: age 50–54; male; married; bmi 'normal weight'. Variables included: Log income (Column 1); Demographic characteristics (age, gender, marital status) (Column 2); Health controls (self reported health status, adl, mobility, chronic conditions, bmi) (Column 3).

Table 1.11: Health Care Systems in SHARE Countries and US

Country	Total Health Expenditure (Per Capita US$ PPP)	Public Health Expenditure (% Total Health Exp.)	GP as Gate–keeper (GK)	Doctor's Type of Payment	Co-payments					
					DOC	GP	SPEC	DENT	INPT	OUTPT
AT	3397	0.756	NO	F	NO	NO	NO	YES	YES	NO
DE	3162	0.769	NO	F	YES	NO	YES	NO	YES	YES
SE	2964	0.846	YES	C	YES	YES	YES	YES	YES	YES
NL	3156	0.625	YES	C	NO	NO	NO	NO	NO	NO
ES	2128	0.709	YES	S	NO	NO	NO	YES	NO	NO
IT	2401	0.758	YES	C	YES	NO	YES	YES	NO	YES
FR	3117	0.794	NO	F	YES	YES	YES	YES	YES	YES
DK	3030	0.843	YES	F	NO	NO	NO	YES	NO	NO
GR	1991	0.446	NO	S	YES	YES	YES	YES	YES	YES
CH	3990	0.585	NO	F	YES	YES	YES	YES	YES	YES
BE	3311	0.731	NO	F	YES	YES	YES	YES	YES	YES
US	6014	0.447	YES	F	YES	YES	YES	YES	YES	YES

Source: Van Doorslaer et al. (2006); OECD Health Data (2007); WHO (2004).
Notes: Doctor's Type of Payment: fee-for-service (F), capitation (C), and salary (S).

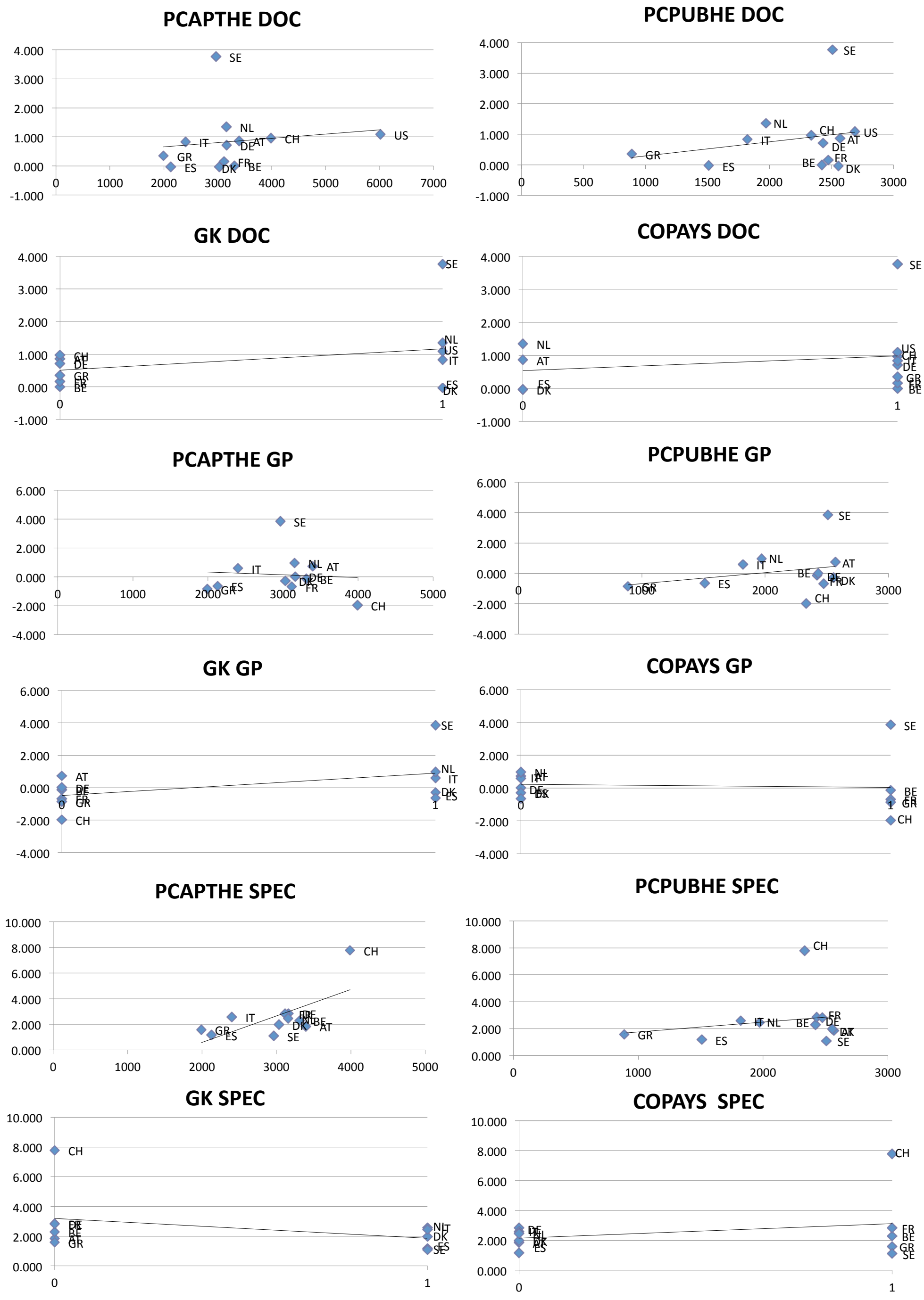


Figure 1.2: Income Gradient and Institutional Variables

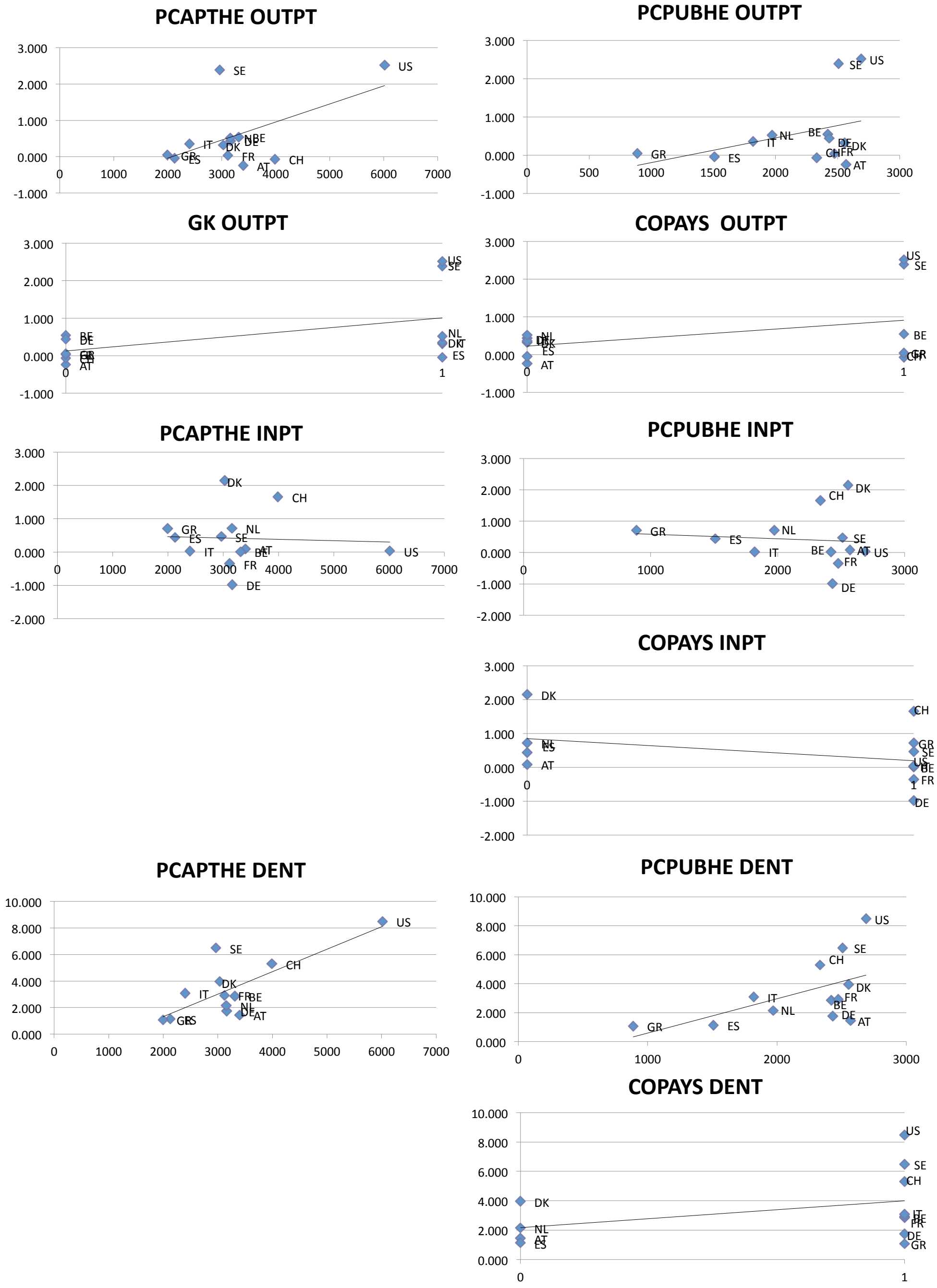


Figure 1.2: continued from previous page

Chapter 2

Microeconometric Determinants of Preventive Health Care[§]

2.1 Introduction

In recent years there has been increasing interest in organizing health promotion and disease prevention activities. The aging of the population encourages innovations in preventive care, as future reduction of morbidity and mortality is linked to the diffuse adoption of preventive practices. Preventive care therefore plays a key role in population health. Existing literature (Syme 1998) suggests that socio-economic status is relevant in both morbidity and mortality of diseases, and is therefore an important factor to take into consideration for evaluating and managing prevention. Identifying critical factors determining preventive care allows for a better programming and resource allocation for specific policies. Preventive care refers to forms of health care that are used to avoid future symptoms, or disease. Preventive care can help reduce morbidity and the probability of death from a particular disease. This paper investigates individual decisions to take preventive health related actions. The objective is to analyze the relation of education and other indexes of socio-economic status with clinical preventive services in Europe.

Links between access to health care, utilization of care services, and socio-economic position are well established (see, for example, WHO 2005). Some studies have also established a positive relationship between socio-economic status and the utilization of preventive health care services. A Taiwan study (Wang and Lin 1996) finds that women with lower socio-economic status (lower income and level of education, or living outside the city) are more likely to underuse pap smear screening (a procedure to test for cervical cancer in women). Moreover, the well known barriers to the utilization of preventive services are reinforced by alienation and social isolation (Bullough 1972), both of which can exacerbate

[§]This chapter is based on "Microeconometric Determinants of Preventive Health Care: An Application to European Countries", M.C. Majo, and A. van Soest. *Acknowledgments:* I am particularly grateful to Arthur van Soest, Franco Peracchi, Marianna Brunetti, and seminar participants at Tor Vergata University for extensive and very useful comments.

existing disadvantaged socio-economic conditions.

To our knowledge this is the first time that preventive care systems are analyzed at the level of a large number of European countries. This paper investigates socio-demographic and health factors in the use of six clinical preventive services, using data from a population-based sample of adults of ages 50 and older in 13 European countries. Prior studies dealing with disparities in the use of preventive services have typically focused on one specific preventive service (in isolation) (see, for example, Witt 2008, and Jonas et al. 2009), or two (Avitabile et al. 2008). Moreover many studies are from the US, or have a narrower geographic focus (see, for example, Carman and Kooreman 2010, and Wang and Lin 1996). Another important aspect is that such studies usually tend to focus on financial access rather than considering a multidimensional concept of access. Here we focus on more than one preventive clinical service, we use a dataset covering thirteen countries in Northern, Western, Southern as well as Central Europe, and we use a multidimensional concept of access in order to understand socio-economic differences in the use of these services. We consider several indexes of socio-economic status and several factors thought to be relevant for preventive care as potential confounders in the analysis. These included socio-demographic factors as age, partnership status, occupation, and country dummies accounting for regional and cultural differences. We focus on screening tests that are strongly recommended to individuals aged 50 or older, regardless of their health history. This avoids the problem of selection bias that arises in samples of individuals that have already been diagnosed with a specific disease. Nevertheless we adjust for illness burden, using available information about self-reported health status, more objective measures of health, such as the number of chronic diseases, and health behavior measures.

The remainder of the paper is organized as follows. In the next section the literature is reviewed. In Section 2.3 we discuss the main features of the preventive care services that we study. Section 2.4 describes the data and Section 2.5 presents the results of the estimated models explaining the use of preventive care services. Section 2.6 concludes.

2.2 Literature Review

The health economics literature presents several theoretical models to analyze prevention. Both the human capital approach (Grossman 1972) and insurance models provide an interesting framework to evaluate an individual's prevention decisions from an economic point of view.

The human capital approach emphasizes the similarity between the decision to invest in health capital and in other forms of human capital. In a first approach, the decision to invest in health capital depends on time, purchased medical care and other goods (Grossman 1972). In subsequent work, Grossman and Rand (1974) distinguish between medical care and prevention based on health capital depreciation rates (allowing the marginal product of medical care to increase with the degree of illness). Primary prevention and medical treatment are considered substitutes. Determining the appropriate value of patients' time can be complicated, however. The human capital method may overestimate both the amount of time spent and its value to the patient (Jonas et al. 2009). On the other hand human capital models are useful to analyze the importance of education level and age in preventive health care demand. Education level is assumed to increase the efficiency in health production, and health capital is assumed to depreciate at a higher rate as people get older (Kenkel 2000).

The existence of insurance gives way to a different approach, where the individual responds to uncertainty by purchasing insurance or by prevention (in expected utility maximization models). In such a setting, primary prevention reduces the probability of illness, while the individual has medical coverage in case illness occurs. Standard economic theory suggests that health insurance coverage may cause a reduction in prevention activities, but to our knowledge empirical studies have not provided much evidence to support this prediction so far. Courbage and Coulon (2004) investigated how different insurance plans affect individual behaviors in terms of prevention activities in the UK. They tested if purchasing private health insurance modifies the probability of taking any preventive actions (defined here as not smoking and frequency of exercise). They found no evidence that private insurance coverage in the UK reduced the prevention related activities. Moreover, there must be other reasons rather than cost barriers that influence the decision of taking preventive actions: there are countries where preventive checks are provided free of charge, yet underuse is observed (see Katz and Hofer 1994). There is some literature that has investigated preventive care in an expected utility framework. Parente et al. (2005) utilized an expected utility model to describe the benefit of two preventive services, influenza vaccinations and mammograms. In their empirical analysis they showed that increased knowledge about Medicare benefits has a substantial positive effect on the use of both preventive services. Picone et al. (2004) assessed the role of risk and time preferences, expected longevity, and education on demand for three measures used for early detection of breast and cervical cancer in an expected utility framework. In the empirical analysis they found that women who did not have a mammogram had lower education, higher rate of time preference, and lower incomes. Witt (2008) investigated the net benefit of mammography in an expected utility framework. In line with previous literature, she showed that older age increases mammography use, and that decreases in time and opportunity costs, and better health behaviors generally have the same effect.

Health promotion and prevention are extensively characterized also in the medical literature. Cherrington et al. (2007) found that individuals' beliefs about the value of periodic health examinations are associated with the likelihood of receiving preventive care services. They also found that younger individuals are less likely to receive preventive services than older age groups and that males less often get preventive services than females.

Carman and Kooreman (2010) analyzed individual perceptions of the effectiveness of preventive interventions in the Netherlands. They compared the decision about using preventive care to other risky investments. They found that individuals have very poor perceptions of the absolute levels of the probabilities of sickness and survival, and tend to ignore or suppress health issues.

Regarding socio-economic status, both economic theory and the medical literature recognize its importance, despite of the large difference between their approaches. From an economic theory perspective socio-economic status is related to the depreciation rate of investing in prevention due to schooling, or to purchasing insurance subject to income. The argument in medical terms is that socio-economic status matters for lifestyle and standard of living and by influencing the incidence and consequences of diseases, it changes the need to take preventive actions. Katz and Hofer (1994) compared the association of income and education with breast and cervical cancer screening in Ontario, Canada and the United States. They found that breast and cervical screening rates were similar between countries, but mammography rates were two to three times higher in the United States than in Canada across all age groups. Universal insurance coverage (in Ontario) is not sufficient to overcome the large disparities in screening across socio-economic status demonstrated in both countries. Low screening rates among poor women have several possible causes rather than lack of insurance.

Beyond epidemiology, there are important economic issues that must be addressed to understand preventive care choices. In accordance with the human capital models and expected utility models, we include educational levels, income, occupation controls, and age as explanatory variables for preventive care use in the econometric analysis. Findings of previous studies dealing with insurance models do not motivate including health insurance coverage (moreover our data do not allow us to measure insurance coverage). On the other hand an extensive literature shows how education increases awareness of and reduces risky health behaviors (Kenkel 1991) and affects the demand for early detection of breast cancer screening and flu vaccination (Kenkel 1994; Mullahy 1999). We use probit models explaining the binary choices whether individuals use a certain type of preventive care or not. Variables like chronic disease and body mass index are also incorporated in our analysis: such variables, by changing the probability of determined disease, also change the individual incentives to do prevention.

2.3 Preventive Health Care

Preventive care refers to forms of health care that are used to avoid future symptoms, disease or death. The standard definition[1] of prevention usually distinguishes two categories, according to the scope and consequences for individual health: primary and secondary prevention (though some preventive actions can be defined in both of those categories). Primary prevention is defined as both the prevention of disease before it occurs and the reduction of its incidence. The aim is essentially to take action in order to prevent the disease itself. This category includes actions like vaccination, healthy dietary habits, personal hygiene, public sanitation, exercise, regular medical consultation and check-ups. Secondary prevention includes interventions that will detect disease in the early stages before symptoms manifest with the aim of reducing the severity of the disease or provide a cure. The aim is to achieve an early diagnosis and treatment. Examples of this category are cancer screening programs (breast, colorectal and cervical cancer, for instance).

[1]U.S. Department of Health and Human Services 2000.

The present study considers six preventive care interventions available in many European countries: influenza vaccination (flu shot), colonoscopy, blood stool test, blood check (which includes blood pressure, blood cholesterol, and blood sugar check), eye exam, and x-ray of the breast (mammogram) for women. Flu shots, blood pressure and cholesterol screening, and eye exams are examples of primary clinical preventive care; colon screening, blood stool test, and mammograms are examples of secondary preventive care.

Flu vaccination in the last 12 months (FLUSH). Flu shots can reduce the probability of dying from influenza or related complications in the coming flu season by 80% (Fiore et al. 2007); in most of the European countries which we consider the health care system provides vaccines to everyone over the age of 65. Even though vaccination can considerably reduce the incidence and severity of influenza, its take-up is often far from complete.

Blood check in the last 12 months (BLOODTEST). It refers to the questions whether a nurse or doctor checked the respondent blood pressure, blood cholesterol, or blood sugar in the last 12 months. Blood pressure checks are for detecting high blood pressure. Having high blood pressure puts a person at more risk for strokes, heart attacks, kidney failure, loss of vision, and atherosclerosis (hardening of the arteries). Blood sugar levels are extremely important to overall health, in particular checks are made to prevent from diabetes. Cholesterol is a type of fat (lipid) made by the body. It is essential for good health and is found in every cell in the body. However, having a high cholesterol level in the blood (hypercholesterolemia) can increase the risk of heart disease and stroke.

Colonoscopy (COLONOS) and Blood stool test (BLOODSTOOL). Respondents are asked whether they ever had colonoscopy or blood stool test. Most guidelines recommend endoscopic examination of colon/sigmoid from the age of 50 in men and women (the frequency varies; here it is asked as “ever had” as a conservative measure). To allow for the early detection and prevention of colorectal cancer, the fecal occult blood test is recommended yearly for everyone starting at age 50. Colonoscopy screening may be done (along with a flexible sigmoidoscopy) every 5 years to check for colorectal polyps or cancer. They have been shown to be effective in reducing colorectal cancer mortality by allowing early detection and removal of colorectal cancers. However, the use of endoscopic screening in the general population remains low (Chao et al. 2004). This might be due to the fact that the screening colonoscopy process requires a considerable amount of time and some discomfort for patients preparing for, having, and recovering from the procedure. Jonas et al. (2009) described the use of willingness-to-pay to value the time required for screening colonoscopy and the discomfort associated with the procedure of patients from an off-campus University endoscopy center in the United States. They found that difficulty of preparation is the main determinant of willingness-to-pay, and that willingness-to-pay is, not surprisingly, sensitive to income.

Eye examination in the past two years (EYE). It is indicated in older persons owing to the risk of glaucoma and the increasing prevalence of diabetes with age. Guidelines indicate that all adults should have eye exams every year once they have reached age 40, and these exams become even more important in case of high blood pressure, diabetes, or a family history of eye disease.

Mammogram in the last two years (MAMMOG). It represents a variable measuring the individual woman's own preventive health care. According to European guidelines[2] women between the ages of 50 and 69 must have the right to attend high-quality mammography screening at two-year intervals in dedicated and certified centers paid for by health insurance schemes. In 2000, the last year for which global data exists, some 400,000 women died from breast cancer, representing 1.6% of all female deaths (WHO 2002). Annual mammograms can reduce the probability of death from breast cancer by 15% by identifying breast cancer early (Gøtzsche and Nielsen 2006). Compared with thinner women, obese women have higher mortality rates for breast and cervical cancer. In addition, obesity leads to adverse social and psychological consequences. A recent study (Wee et al. 2000) suggests that obese women receive preventive services less often than normal-weight women.

All in all the rates at which individuals receive clinical preventive services vary by type of service, with overall rates of nearly all services significantly lower than what is recommended by clinical guidelines (Coffield et al. 2001). Moreover, similar to what happens for standard types of health care, the use of clinical preventive services is patterned by key social variables such as age, ethnicity, gender, education, and income, but also to risky health behaviors. Obesity is associated with higher mortality rates for cardiovascular disease and cancer of the cervix, breast, and colon, therefore barriers to preventive screening and counseling in obese patients can have medical and economic consequences (Wee et al. 2000).

2.4 Data and Methods

The data are drawn from the Survey of Health, Ageing and Retirement in Europe (SHARE). This paper uses data from wave 1 and wave 2 of SHARE (2004 and 2006 respectively). SHARE is the first database that includes quality of care indicators for those aged 50+. It represents crucial information for the evaluation of health systems, including harmonized information on a variety of dimensions such as health, health care use, and socio-economic conditions. A detailed description of the survey methods and descriptive statistics is available elsewhere (see Börsch-Supan et al. 2005; Börsch-Supan and Jürges 2005; and on the Internet at www.share-project.org). Briefly, the study sample is restricted to adults aged 50+. It covers eleven European countries in the first wave (2004): Austria (AT), Germany (DE), Denmark (DK), Spain (ES), France (FR), Greece (GR), Italy (IT),

[2]European guidelines for quality assurance in breast cancer screening and diagnosis (2006).

the Netherlands (NL), Sweden (SE), Switzerland (CH), and Belgium (BE). In the second wave (2006), the Czech Republic (CZ) and Poland (PL) were also added[3]. It is the first European survey that provides cross-nationally harmonized micro-data on the economic, social and health situation of the 50+ population. It offers a unique opportunity for a much richer analysis of the relationship between human capital and socio-economic status on the one hand and the use of health preventive care facilities on the other hand. Indicators of preventive care are extracted from the drop-off questionnaire. Information about a variety of types of health care was collected in the interview, such as immunizations, check-ups, breast cancer screening and eye checks. Indicators of quality of care in SHARE are measured based on self-reports and rely on aspects of medical consultations that are easily recognized by respondents, irrespective of their level of education (see Börsch-Supan et al. 2005, Chapter 3.9). Respondents fill in the drop-off questionnaire only once, so that a longitudinal dimension is not available and only cross-section analysis is possible. The variables used in the analysis are described and summarized in Table 2.1. The working sample includes those who answered at least one of the preventive exam questions.

2.4.1 Preventive Care Measures in SHARE

SHARE respondents were asked if they had received a variety of clinical preventive health services within the past year. Six clinical preventive services were addressed, including 5 that are recommended for both men and women and mammogram testing, which is recommended for women. All preventive exams are available in SHARE in the form of a binary variable indicating preventive service uptake.

Figure 2.1 shows a cross country comparison of the use of the preventive care services, based upon our working sample. Highly differentiated pictures of the different preventive care measures can be seen. The fraction of those aged 50+ getting flu vaccination varies from 47% in Belgium to less than 20% in Poland, the Czech Republic, and Greece. Blood tests are administered to between 80-90% in most countries, with the exception of Denmark and Switzerland where the rates are 60% and 65%, respectively. Colonoscopy exams are done for only 8% of the respondents in Greece, but reach almost 38% in Austria. Similarly blood stool test has the lowest uptake in Greece (5%) and the highest in Austria (63%), showing a similar picture across countries for these two exams. On the other hand, in Belgium and the Netherlands only 9% and 4% of the respondents ever had a blood stool test exam, but 21.5% and 13.5% of respondents had a colonoscopy. Only in a few countries special programs for such screenings are in place (Holland et al. 2006). In Austria everybody aged 50 years or more is invited for regular check-ups and is informed about the risks of colorectal cancer, whereas in other countries there are no special provisions for colorectal cancer screening. Moreover the testing protocol differs across countries: in France, for example, individuals at risk are advised to have colonoscopy only if the fecal blood test is positive. Eye exam prevalence varies from 41% in the Netherlands to 73% in

[3]Israel and Ireland are also in SHARE 2006, however they are not included in the analysis. Ireland has very few observations and the way in which the Israeli health care system deals with preventive care is quite different from the European countries.

France. The picture that we see for mammography shows large variation across countries, with Denmark and Greece having the lowest percentages (24% and 32%, respectively) and Sweden, France, and the Netherlands the highest (68%, 70%, and 73%, respectively). Such differences across countries are likely to be due to the time range between one exam and the other, and to other cultural and institutional differences. In some countries the health care system offers preventive exams free of charge every two years, in other countries every three to four years, whereas in other countries the exam is not offered free of charge.

2.4.2 Independent Variables

As discussed in Section 2.2, the covariates in our models include schooling, economic factors (income, occupation), socio-demographic (age, gender, partnership status) and medical characteristics (self-reported health status, number of chronic conditions, number of mobility limitations, and body mass index).

The demographic variables included in the analysis are age, gender, and partnership status (which also includes being married). Age is measured at the time of the interview, and is grouped into 5-year age bands: 50–54; 55–59; 60–64; 65–69; 70–74; 75–79, and 80+. Gender may affect preventive care take-up: for example women might be found to have higher take-up rates as they are commonly found to be more risk averse than men. Partnership status is categorized as: married or living with a partner (coded as 1) versus single (coded as 0). Here the hypothesis is that partnered individuals have higher take-up rates because of shared health information or positive externalities within the couple. Socio-economic status (SES) characteristics include household income, occupation, and schooling. Household income is defined as equivalent income (adjusted for household size through dividing by the square root of the number of household members). Income is measured as the log of gross annual household income for 2003 for the wave 1 observations and log of net annual household income for 2006 for the observations in wave 2)[4] and it is derived from disaggregated income sources including labor and non-labor income, transfer income, investment income, benefit income and pension income (total individual income of each respondent, sum of the incomes of other household members and other benefits, capital assets income, excluding rent payments received and imputed rents). A dummy variable JOB indicates whether the respondent works or not. According to the literature, this is a proxy for opportunity costs, as time off work needed to take preventive care can be costly.

Education level is categorized as low (LOWEDU), middle (MIDEDU), and high (HIGHEDU) education level according to the ISCED-97 harmonized coding for international comparisons[5]. Educational attainment is likely to capture potential differences in health literacy across education levels. Health literacy is defined as the degree to which individuals have the capacity to obtain, process, and understand basic health information needed to make appropriate health decisions and services needed to prevent or treat illness. As

[4] A year control variable will also take into account the difference between gross and net income in the two waves.

[5] See for details on ISCED coding: www.uis.unesco.org/ev.php?ID=3813_201&ID2=DO_TOPIC.

a consequence we would expect better educated individuals to have higher demand for preventive care. We expect education to have a higher impact on the choice of doing prevention. We expect that the more educated the respondent, the more she will be concerned about her health, and the better she can process information received from the general practitioner. Moreover, a higher socio-economic status is often associated with less social exclusion or alienation, which, as discussed in Section 2, reduces the probability of taking preventive actions.

We control for self-reported health status (SPHS), classified in two categories (very good/excellent and less than very good, with the former being the base category). Individuals who have a worse perception of their health are expected to attach higher value to preventive care services (see Hunt et al. 1981). We also control for the number of limitations with mobility and the number of chronic diseases. The variables "limitations with activity of daily living" (ADL) (such as dressing, bathing, or getting in and out of bed) and "mobility limitation" (MOBILIT) indicate the extent to which individuals consider themselves physically handicapped. Both variables are reclassified into two categories: no limitations with ADL (or MOBILIT) and one or more limitations with ADL (or MOBILIT). In addition, we include a variable indicating whether or not the respondent has two or more chronic diseases (CHRONIC), based upon questions that ask whether the respondent has ever been diagnosed with a number of chronic diseases[6]. Finally, we control for health behavior by including dummies for underweight, overweight, or obese; the benchmark group is those of normal weight. These dummies are based upon body mass index (BMI): weight (in kilograms) divided by height (in cm) squared[7]. BMI is a predictor of risk for obesity-related diseases. It is also a measure for high-risk conditions, as BMI is a reliable indicator of total body fat, which is related to the risk of disease and death. Finally, we include country dummies to account for cultural and regional differences, their interaction with education, and a wave dummy to control and distinguish observations in waves 1 and 2.

We estimate four models in each case, using probit models explaining the yes/no answer to the question whether respondents have received the type of preventive care check at least once in the past twelve months (flu shot, blood test), or in the past two years (mammogram and eye check), or whether they ever had it (colonoscopy, blood stool test). The probit models are estimated separately for each type of check-up differing in additional factors that we control for. The base estimation consists of a probit regression, with independent variables education level qualifications, country dummies and their interactions with education level[8], and demographic characteristics including gender and partnership status. In the second model we also control for income and employment status. The third

[6]The number of chronic diseases is a count of the following diseases an individual might have: heart problems, high blood pressure, high cholesterol, cerebral vascular disease, diabetes, lung diseases, asthma, arthritis, osteoporosis, cancer, stomach ulcer, Parkinson disease, cataracts, hip fracture or femoral fracture.

[7]BMI categories: $\text{BMI} \leq 18.5$ underweight; $18.5 < \text{BMI} < 25$ normal weight; $25 \leq \text{BMI} < 30$ overweight; $\text{BMI} \geq 30$ obese.

[8]The coefficient estimates of the interactions are not shown here, but available upon request (they are often jointly not significant).

specification adds health status controls.

2.5 Results

In this section we describe the results of the probit estimations explaining the yes/no answer to the questions whether respondents have received the various types of preventive care at least once in the past twelve months (flu shot, blood test), at least once in the past two years (mammogram and eye check), or at least once at any time in the past (colonoscopy, blood stool test). The probit models are estimated separately for each type of check up. Tables 2.2 – 2.7 present the estimated marginal effects at the country specific means. The marginal effects of the educational dummies can be interpreted as the change in percentage points in the probability of using the preventive clinical service for respondents with an intermediate or high education level compared to those with low education, keeping constant all other explanatory factors included in the model. The other marginal effects have a similar interpretation.

2.5.1 Flu Shot Vaccination (in the last year)

The probit estimation, which is shown in Table 2.2, includes respondents from both 2004 and 2006, with both CZ and PL included in the model. In accordance with what we expected, higher educated respondents get significantly more flu vaccinations than those with a lower level of education, irrespective of whether or not we control for demographic characteristics, economic variables, and health conditions. Age has a positive and significant effect, and the age effect increases as respondents get older. This is in line with the fact that flu shots are often provided for free to those who have reached a certain age. There seems to be no gender difference in flu vaccination, but being married or living with a partner increases the probability of getting a flu vaccination compared to living as a single.

When controlling for economic factors (income and employment status; column (2)), the effect of higher education remains positive and significant and even increases in magnitude. Income is also positive and significant, even though in most of the European countries the health care system provides vaccines for free to everyone over the age of 65. The estimated marginal effect of 0.806 implies that the probability of using a flu shot would increase by about 0.08 percentage points if income increased by 10%, keeping constant all other explanatory factors included in the model. This effect seems rather small compared to the effect of education. Those who are working are less likely to get flu vaccination than those who are not working or retired. As said previously, this may be due to the fact that it is relatively costly for workers to get preventive screenings during work hours.

When we also control for health status (column (3)) the difference between secondary and low education also becomes significant and the effect of higher education is strengthened further. The reason is that education is positively correlated with health, and people in good health are less likely to get a flu shot than those with poor health. The effects of the health variables show that having a lower perception of own health status makes the respondents more likely to get a flu vaccination. Also, those with a higher number of chronic conditions or mobility limitations are more likely to get a flu vaccination; this is probably because their health status makes them more vulnerable to suffer seriously or even die from the flu when they catch it. The effects of being obese and overweight are also positive and significant (with normal-weight being the base category). Limitations with activity of daily living or being underweight do not seem to have any effect on flu vaccination uptake.

There are significant differences across countries. The negative sign for SE, DK, GR, CZ, and PL shows that respondents in these countries have a lower probability of getting a flu vaccination than in DE (the omitted benchmark country), keeping other factors constant. On the other hand, respondents in NL, ES, IT, FR and BE are more likely to get vaccinated than their DE counterparts. The year dummy is not significant. Using log likelihood values to compare the three models, a significant improvement can be seen when going from the first to the third model (column (1) to column (3)), where the log likelihood value is significantly higher according to a standard likelihood ratio test (LR test with p–value 0.000).

2.5.2 Blood Test Check (in the last year)

The probit estimation, which is shown in Table 2.3, includes only new respondents in 2006, with both CZ and PL included in the model (no observations are available for GR). The effect of high education is significant at the 10% level when we only control for demographic characteristics (column (1)) and at the 1% level when we also control for health status and income and employment status (column (3)). Secondary education is not significantly different from low education in blood test uptakes.

When we control for economic factors (column (2)) income is significant, and those with higher income are more likely to take the test. The estimated marginal effect of income is somewhat higher than for flu shots. Employment status is not significant. There seems to be an increasing uptake with age until age 75–80; those aged 80+ have lower take up than the 75–80 group, keeping other factors constant. Women make significantly more use of preventive blood tests than men. Partnership status is not significant.

When we control for health status (column (3)), the worse the health status (both perceived and objective) and the more overweight, the more likely the respondents have their blood pressure, cholesterol, and sugar levels checked. As we previously mentioned, obesity is associated with higher blood pressure and higher cholesterol and sugar level. Moreover, poor health makes it more important to periodically check the blood owing to the risk of increasing prevalence of diabetes or strokes, or worsening of other chronic conditions. This is why general practitioners usually prescribe such tests in case of chronic conditions such as diabetes. Moreover, in some countries the test is free of charge in case a patient has one or more chronic conditions.

Country dummies are significantly positive for FR and BE and significantly negative for SE, NL, DK, PL, showing substantial differences across European countries considered in the analysis. The log likelihood values for the three models again show an improvement when going from the first to the third model (column (1) to column (3)), where the log likelihood value is significantly higher (LR test with p–value 0.000).

2.5.3 Colonoscopy and Blood Stool Test (ever had)

The probit estimations, which are shown respectively in Table 2.4 and Table 2.5, include respondents from 2004 only (CZ and PL are not included in the model). The education patterns are very clear and stronger than for the preventive services discussed above for both colonoscopy and blood stool tests; in all specifications, both those with secondary and those with higher education level are more likely to take up these tests than respondents with low education, and the magnitude of the differences is also larger than for flu vaccinations or blood tests. Both exams are rather intrusive and it is likely that higher educated people can access information about the usefulness of these types of preventive services better than low educated respondents, implying they have a different attitude towards such exams. When we control for economic factors (column (2)), income is significant for both colonoscopy and blood stool test. The probability of getting colonoscopy would increase of 0.1 percentage points if income increased by 10% (keeping constant all other explanatory factors included in the model). Those who are working are significantly less likely to get a colonoscopy or a blood stool test than those who are not working or retired, which is probably due to differences in the opportunity costs of time. No significant gender effect is found for these tests but partnership matters for the blood stool test, probably because of shared health information within the couple. Colonoscopy is more likely among those aged 75–80, whereas the blood stool test is most prevalent in the age group 65–69. Blood stool tests are less invasive and easier to perform than a colonoscopy and typically prescribed at an earlier age. Geiger et al. (2008) examined the influence of age, race, gender, education, income, media usage, and interactions with health care providers on knowledge, attitudes, and behavior regarding colonoscopy screening for colorectal cancer and found that knowledge of and participation in screening colonoscopy is low in the US population, especially among Hispanics. This is in line with our findings that a higher education level and higher income are associated with higher prevalence of colonoscopy screening.

When we control for health status (column (3)), a worse health status (according to both subjective and objective measures) leads to an increase in both exam uptakes. As for the preventive services discussed above, having a lower perception of their own health status makes the respondents more likely to get preventive care. Surprisingly, however, it seems that obese respondents are less likely to feature both preventive measures than respondents of normal weight, in contrast with general guidelines. People with weight problems might feel more uncomfortable with taking these exams.

Country dummies are significant and negative for almost all the countries, with the exception of FR (not significant for colonoscopy) and AT (positive), showing that the tests are relatively common in DE (the omitted country). Such across countries differences can be due both to different perceptions across national cultures or to differences in information on these types of screenings or institutional differences in guidelines provided by the health care system. When using log likelihood values to compare the three models, an improvement can be seen for both colonoscopy and blood stool test when going from the first to the third model (column (1) to column (3)), where the log likelihood value is significantly higher (LR test with p–value 0.000).

2.5.4 Eye Exam (in the last two years)

The probit estimation, which is shown in Table 2.6, includes respondents from both 2004 and 2006, with both CZ and PL included in the model. Eye exam uptake is strongly associated to education level. The higher the education level, the more likely it is that the respondents get their eyes checked, with a much stronger effect of education than for the services discussed earlier. In particular, the difference between secondary education and primary education is quite large (about 9 percentage points in all three models), and much larger than the difference between tertiary and secondary education (which is between 2 and 3 percentage points).

Controlling for economic factors (column (2)), we find that the higher the income, the more the respondents get their eyes checked, with an increase of 0.19 percentage points if income increases by 10%, keeping other explanatory factors constant. The effect of doing paid work is positive and becomes significant once we also control for health status. The net benefit is ambiguous here: while it is relatively costly for workers to obtain preventive care, workers also have relatively more to lose from having eye problems. Apparently, the latter effect dominates the former, since workers may realize the benefit of early detection to avoid productivity loss due to eye problems.

The elderly age classes show positive and significant coefficients and so do gender and partnership status coefficient estimates. As mentioned above, eye examination is advised for older persons owing to the risk of glaucoma and the increasing prevalence of diabetes with age. When we control for health status (column (3)), we find that having a lower perception of one's own health status does not matter for the likelihood of an eye exam. On the other hand, the number of mobility limitations and having more than two chronic conditions make the respondent more likely to get such an exam. Being underweight has a negative and significant effect, although the existing literature does not report a clear link between body weight and risk of glaucoma.

Significant differences across countries are found, with significantly negative effects in most countries, suggesting that taking an eye exam is more likely in DE and FR than in other countries (keeping other factors constant). The year control is not significant. When using log likelihood values to compare the three models, an improvement can be seen when going from the first to the third model (column (1) to column (3)), where the log likelihood value is significantly higher (the LR tests of model 3 against the other two models both have p–value 0.000).

2.5.5 Mammogram (in the last two years)

The probit estimation, which is shown in Table 2.7, includes female respondents from 2004 and 2006, with both CZ and PL included in the model. In accordance with what we expected, higher educated females significantly get more mammogram exams than those with the lowest level of education, irrespective of whether we control for economic status and health conditions or not – Adding economic status or health controls hardly changes the effect of education. When we control for economic factors (column (2)), income has a positive and significant effect, while occupation is not significant. Mammograms are recommended for women between ages 50 and 70. Our results show significant and positive coefficients for women aged 60− in accordance with the recommendations, as older age (until 70) increases breast cancer risk.

Controlling for health status (column (3)) shows that having more than two chronic conditions increases the probability of taking the exam. Respondents having limitations with activity of daily living are significantly less likely to take the exam, as do underweight and obese respondents. The result for obese women is surprising, since compared to other women, obese women have higher mortality rates for breast and cervical cancer. Our results are in accordance with the study by Wee et al. (2000) who also found that obese women receive preventive services less often than normal-weight women. Here social barriers and alienation feelings might play a role in lowering the uptake.

There are significant differences across countries. In DK and PL, respondents have a lower probability of getting a mammogram compared to DE, whereas several all other countries have significant and positive coefficients. These differences across countries are likely to be due to the differences in price and in the recommended time range between mammography screenings, and to other cultural and institutional differences. In some countries the health care system offers mammogram tests for women free of charge every two years, while in other countries this is every three to four years, and in some other countries mammography is not free of charge at all. The year dummy is not significant. When using log likelihood values to compare the three models, an improvement is found when going from the first to the third model (column (1) to column (3)), where the log likelihood value is significantly higher (LR test with p–value 0.000).

2.5.6 Preventive Screening, Education, and Income

All in all education level emerges as a very important determinant for preventive care screening take-up. Each probit model has also been estimated including interactions between educational qualifications and country dummies in all the three specifications[9]. Results generally remain unchanged and likelihood ratio tests indicate that the interactions are jointly insignificant in most cases. A similar picture is found when estimating preventive care country by country. We do find significant differences across countries, probably since we do not include specific variables that measure institutional differences across countries as specific information for all the countries analyzed was not available (and collecting it would go beyond the scope of the present paper).

When we also control for income and employment status, the effects of education do not change much. Moreover, these models also show that (keeping education, employment

[9]Coefficient estimates are not shown here, but are available upon request

status, and other factors constant), seeking preventive care is positively associated with income. This confirms the strong and robust nature of the socio-economic gradient in getting preventive care. Indirectly, since preventive care is beneficial for health outcomes, it may also explain at least part of the socio-economic gradient in health outcomes that is commonly found in the literature (see, e.g., Smith 1999, Meer et al. 2003, Hurd and Kapteyn 2003, Adams et al. 2003, or Michaud and van Soest 2008).

The effect of education on preventive care remains positive and significant for the majority of the countries analyzed also after controlling for health characteristics. Our results therefore unambiguously suggest a significant role for schooling in explaining all preventive care screenings, which confirms the crucial importance of information and being able to process the information available concerning one's own health. Moreover, those with a better understanding of their own health status are more likely to take preventive action to reduce the risk of health problems. Knowledge of and participation in preventive screening becomes crucial if the aim is to increase the participation rates in preventive care measures. This may indeed represent an important policy tool for increasing their uptake of preventive actions.

2.6 Conclusions

Preventive care plays a key role in population health. Preventive care interventions can improve the quality of life and reduce morbidity. Our analysis of education first, and then of three indicators of SES – education, income, and employment status – suggests that economic and social resources are associated with whether respondents use preventive services. With more education and skills, respondents are more likely to have better jobs with employer-sponsored health benefits (or to have incomes that permit them to afford comprehensive individual coverage). More informed respondents are also likely to better navigate their way through increasingly complex health care systems, with more educated respondents being usually more patient.

Income is similarly related to the use of preventive services. With higher economic status comes greater purchasing power and greater likelihood of medical access. Non-poor respondents have more available income to purchase health insurance or to benefit from access to employment-based insurance. Even considering that some low-income respondents and older respondents are eligible for specific health insurance plans, on average, they are still less likely to access medical care without problems, and to use preventive services. Work status is the least significant SES indicator of preventive services. After controlling for other determinants such as health status, respondents with an employment, on average, use more preventive services than respondents who are not employed, which suggests that workers have less time to seek out and obtain preventive services.

Unfortunately we have little information to measure the role of risk and time preference. SHARE provides a measure of risk aversion in terms of the amount of financial risk that the respondent is willing to take when she saves or invests. The item has only been included in the second wave of SHARE. Therefore we could only run the analysis with a more restricted sample, and results were often not significant and not in line with what expected. How risk and eventually time preference influence the decision of taking preventive actions represents an interesting topic for further research when more detailed information will become available.

Most theoretical models of health insurance predict a negative relationship between health insurance and the level of preventive actions as under full insurance coverage there is no incentive to take preventive actions, creating an ex-ante moral hazard problem (Kenkel 2000). Dave and Kaestner (2006) analyze the impact of universal Medicare coverage (elderly population) in US against four healthy behaviors: tobacco, alcohol, exercise, and weight. They found evidence of ex-ante moral hazard, after modeling for different patterns for insured/uninsured individuals prior to Medicare (younger than 65 years of age) and dividing the insurance effect in ex-ante moral hazard and a physician visit. They conclude that physician visits would increase prevention; however additional investigations would be necessary to make the evidence conclusive. Unfortunately we do not have enough information on health insurance to measure the insurance effect in ex-ante moral hazard. This remains another important topic for future research.

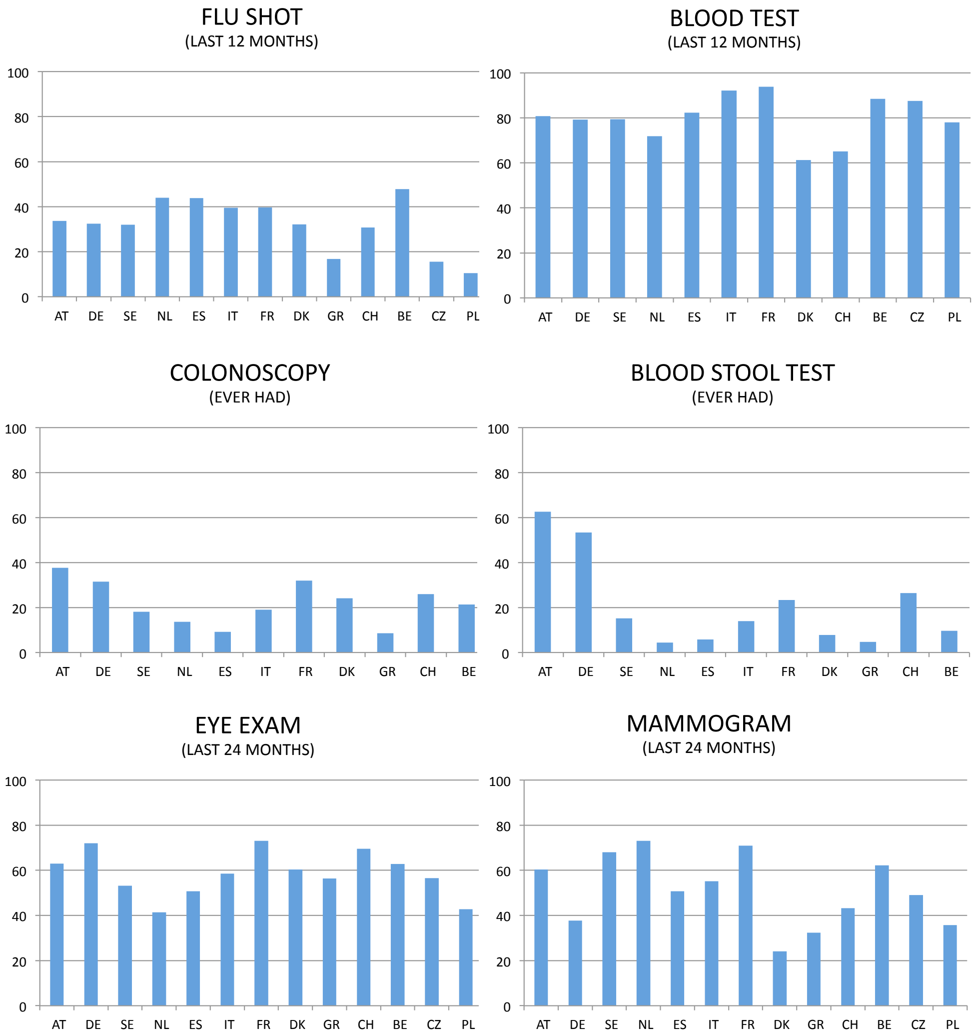
While the direction and nature of the association between SES and health status are difficult to unravel, the implications for individual's health and well-being warrant attention. Chronic illness and poor overall health may lead to lower incomes when respondents are less able or likely to secure employment. On the other hand, poverty-related conditions such as greater likelihood of confronting access barriers, difficulties in obtaining timely primary care, or not using preventive care can increase risk of chronic illness and compromised health status. Policies aimed at controlling the growing costs of health care should look at preventive programs as a key factor. Decision makers and health care providers must design more effective health promotion and disease prevention strategies to raise individual's awareness of risky behaviors, to motivate positive, health-enhancing behavior, and to increase use of preventive screening services.

The effectiveness and efficiency of such polices also depend on how much individuals understand the health risk involved and decide to act accordingly. Better and more accessible information on the medical effectiveness of the intervention and the associated risk of death is important. Therefore it seems that informative materials focused on specific socio-demographic segments need to be developed to encourage screening.

Table 2.1: Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min.	Max.
FLUSH	24479	0.331	0.471	0	1
BLOODTEST	6979	0.829	0.376	0	1
COLONOS	16908	0.213	0.409	0	1
BLOODSTOOL	17330	0.203	0.403	0	1
EYE	24512	0.585	0.493	0	1
MAMMOG	13307	0.555	0.497	0	1
midedu	24743	0.334	0.472	0	1
highedu	24743	0.182	0.386	0	1
logincome	24743	10.285	1.584	0	18.21
job	24743	0.287	0.452	0	1
age	24743	63.995	9.795	50	99
gender	24743	0.538	0.498	0	1
partnership	24743	0.754	0.430	0	1
sphs	24743	0.700	0.458	0	1
adl	24743	0.091	0.287	0	1
mobilit	24743	0.477	0.499	0	1
chronic	24743	0.429	0.495	0	1
underweight	24743	0.012	0.108	0	1
overweight	24743	0.429	0.495	0	1
obese	24743	0.176	0.381	0	1
country	24743	17.662	5.323	11	29
year	24743	2004.572	0.904	2004	2006

Notes: The working sample includes those who answered for at least one of the preventive exams. Calculations based on 2004 and 2006 SHARE data (drop-off questionnaire).



Notes: %. Calculations based on 2004 and 2006 SHARE data (drop-off questionnaire).

Figure 2.1: Prevalence of Preventive Care Use by Country

Table 2.2: Determinants of Preventive Care – Flu Shot Vaccination

	(1)		(2)		(3)	
	ME	SE	ME	SE	ME	SE
MIDEDU	0.901	(0.777)	0.944	(0.782)	1.810**	(0.787)
HIGHEDU	2.755***	(0.913)	3.095***	(0.936)	4.531***	(0.949)
logincome			0.806***	(0.309)	0.840***	(0.312)
job			-5.329***	(0.978)	-3.129***	(0.991)
55–59	6.565***	(1.102)	5.493***	(1.121)	4.354***	(1.126)
60–64	15.856***	(1.091)	13.225***	(1.195)	11.824***	(1.201)
65–69	33.966***	(1.092)	30.452***	(1.270)	28.787***	(1.277)
70–74	40.711***	(1.161)	37.090***	(1.341)	34.803***	(1.350)
75–80	46.185***	(1.279)	42.559***	(1.444)	39.450***	(1.458)
80+	49.221***	(1.376)	45.574***	(1.53)	42.270***	(1.563)
female	0.471	(0.645)	0.023	(0.65)	-0.447	(0.665)
partnership	2.946***	(0.77)	2.597***	(0.778)	3.061***	(0.781)
SPHS					3.152***	(0.793)
ADL					0.406	(1.125)
MOBILIT					1.943***	(0.738)
CHRONIC					10.217***	(0.698)
underweight					-2.748	(2.932)
overweight					1.724**	(0.709)
obese					2.736***	(0.929)
AT	-1.264	(1.517)	-1.561	(1.518)	0.349	(1.524)
SE	-4.522***	(1.417)	-5.926***	(1.624)	-4.690***	(1.639)
NL	11.190***	(1.396)	11.021***	(1.397)	12.665***	(1.407)
ES	7.607***	(1.561)	8.177***	(1.575)	8.160***	(1.581)
IT	5.170***	(1.479)	5.370***	(1.484)	5.941***	(1.490)
FR	5.496***	(1.515)	5.469***	(1.516)	6.355***	(1.522)
DK	-2.873*	(1.670)	-4.434**	(1.821)	-3.396*	(1.841)
GR	-22.280***	(1.676)	-21.534***	(1.696)	-20.977***	(1.710)
CH	-2.179	(1.659)	-2.274	(1.685)	0.53	(1.704)
BE	13.529***	(1.371)	13.405***	(1.373)	14.095***	(1.384)
CZ	-20.641***	(1.877)	-22.209***	(1.980)	-22.610***	(1.985)
PL	-30.642***	(1.992)	-30.739***	(1.995)	-32.397***	(2.004)
year 2004	-0.242	(0.900)	-0.362	(0.902)	0.089	(0.905)
Observations	24479		24479		24479	
Log-likelihood	-13229		-13212		-13022	

Standard errors (SE) in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Marginal Effects (ME): all coefficients are multiplied by 100. Base categories: low education qualification; non working status; age 50–54; man; livingas a single; bmi 'normoweight'; Germany(DE). Column (1) is the base model with education qualification, country dummies, and demographic characteristics; (2) + logincome and occupation; (3) + health controls (sphs, adl, mobility, chronic conditions, BMI).
Calculations based on both 2004 and 2006 SHARE data (drop-off questionnaire).

Table 2.3: Determinants of Preventive Care – Blood Test

	(1)		(2)		(3)	
	ME	SE	ME	SE	ME	SE
MIDEDU	0.048	(1.035)	-0.143	(1.041)	1.095	(0.999)
HIGHEDU	2.370*	(1.420)	2.06	(1.448)	3.909***	(1.388)
logincome			1.023***	(0.388)	0.915**	(0.363)
job			-1.534	(1.251)	1.94	(1.209)
55–59	4.983***	(1.271)	4.558***	(1.308)	2.973**	(1.245)
60–64	10.231***	(1.385)	9.251***	(1.537)	6.487***	(1.476)
65–69	13.607***	(1.543)	12.443***	(1.759)	8.962***	(1.694)
70–74	14.587***	(1.684)	13.320***	(1.899)	8.977***	(1.835)
75–80	16.981***	(1.947)	15.711***	(2.130)	9.663***	(2.052)
80+	13.458***	(2.137)	12.233***	(2.302)	6.384***	(2.245)
female	2.440***	(0.898)	2.228**	(0.909)	1.820**	(0.884)
partnership	0.527	(1.118)	0.111	(1.130)	0.543	(1.084)
SPHS					6.511***	(1.049)
ADL					4.174**	(1.764)
MOBILIT					2.749***	(1.040)
CHRONIC					13.587***	(0.977)
underweight					-6.796*	(3.646)
overweight					3.044***	(0.938)
obese					5.179***	(1.274)
AT	5.164	(5.247)	4.878	(5.241)	6.35	(5.031)
SE	-5.989**	(2.588)	-8.334***	(2.775)	-5.667**	(2.641)
NL	-12.170***	(2.404)	-12.455***	(2.405)	-10.380***	(2.283)
ES	-3.096	(2.756)	-2.583	(2.759)	-0.626	(2.642)
IT	-0.237	(2.370)	0.058	(2.371)	0.657	(2.252)
FR	12.660***	(2.635)	12.460***	(2.636)	13.631***	(2.513)
DK	-9.659***	(2.789)	-12.117***	(2.965)	-9.526***	(2.845)
CH	-0.964	(2.295)	-1.66	(2.328)	3.019	(2.233)
BE	11.643**	(4.560)	11.737**	(4.562)	11.720***	(4.369)
CZ	2.855	(2.110)	0.738	(2.268)	0.353	(2.153)
PL	-6.751***	(2.043)	-6.531***	(2.050)	-9.317***	(1.964)
Observations	6979		6979		6979	
Log-likelihood	-3003		-2999		-2781	

Standard errors (SE) in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Marginal Effects (ME): all coefficients are multiplied by 100. Base categories: low education qualification; non working status; age 50–54; man; living as a single; bmi 'normoweight'; Germany(DE). Column (1) is the base model with education qualification, country dummies, and demographic characteristics; (2) + logincome and occupation; (3) + health controls (sphs, adl, mobility, chronic conditions, BMI).
Calculations based on 2006 SHARE data (drop-off questionnaire). BLOODTEST is not available for Greece (GR).

Table 2.4: Determinants of Preventive Care – Colonoscopy

	(1)		(2)		(3)	
	ME	SE	ME	SE	ME	SE
MIDEDU	2.667***	(0.781)	2.625***	(0.785)	3.462***	(0.786)
HIGHEDU	5.050***	(0.880)	5.172***	(0.902)	6.405***	(0.910)
logincome			1.083***	(0.313)	1.142***	(0.316)
job			-4.961***	(0.955)	-2.996***	(0.961)
55–59	4.190***	(1.073)	3.296***	(1.087)	2.516**	(1.086)
60–64	7.091***	(1.082)	4.744***	(1.177)	4.016***	(1.175)
65–69	10.155***	(1.102)	6.986***	(1.263)	5.850***	(1.263)
70–74	10.870***	(1.180)	7.637***	(1.343)	5.594***	(1.346)
75–80	12.602***	(1.283)	9.337***	(1.433)	6.342***	(1.444)
80+	12.829***	(1.361)	9.515***	(1.504)	5.487***	(1.530)
female	0.881	(0.647)	0.489	(0.651)	-0.098	(0.662)
partnership	0.012	(0.758)	-0.456	(0.7660)	0.009	(0.764)
SPHS					4.978***	(0.779)
ADL					2.983***	(1.111)
MOBILIT					2.300***	(0.728)
CHRONIC					6.350***	(0.695)
underweight					2.249	(2.682)
overweight					-0.15	(0.699)
obese					-2.079**	(0.937)
AT	4.334***	(1.250)	4.054***	(1.251)	5.724***	(1.251)
SE	-12.641***	(1.296)	-14.833***	(1.518)	-13.498***	(1.524)
NL	-16.990***	(1.353)	-17.188***	(1.353)	-15.759***	(1.352)
ES	-21.952***	(1.615)	-21.172***	(1.628)	-21.169***	(1.626)
IT	-10.916***	(1.453)	-10.625***	(1.458)	-10.001***	(1.450)
FR	1.183	(1.411)	1.118	(1.410)	1.493	(1.407)
DK	-7.083***	(1.447)	-9.366***	(1.619)	-8.098***	(1.629)
GR	-24.492***	(1.494)	-23.595***	(1.515)	-22.242***	(1.517)
CH	-3.562**	(1.691)	-3.971**	(1.718)	-1.432	(1.723)
BE	-8.426***	(1.218)	-8.506***	(1.219)	-7.827***	(1.221)
Observations	16908		16908		16908	
Log-likelihood	-8224		-8206		-8084	

Standard errors (SE) in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Marginal Effects (ME): all coefficients are multiplied by 100. Base categories: low education qualification; non working status; age 50–54; man; living as a single; bmi 'normoweight'; Germany(DE). Column (1) is the base model with education qualification, country dummies, and demographic characteristics; (2) + logincome and occupation; (3) + health controls (sphs, adl, mobility, chronic conditions, BMI).
Calculations based on 2004 SHARE data (drop-off questionnaire).

Table 2.5: Determinants of Preventive Care – Blood Stool Test

	(1)		(2)		(3)	
	ME	SE	ME	SE	ME	SE
MIDEDU	2.553***	(0.731)	2.478***	(0.735)	2.911***	(0.734)
HIGHEDU	4.184***	(0.831)	4.225***	(0.852)	4.754***	(0.857)
logincome			1.007***	(0.288)	1.003***	(0.289)
job			-4.216***	(0.889)	-3.007***	(0.894)
55-59	4.608***	(0.996)	3.793***	(1.009)	3.178***	(1.005)
60-64	6.760***	(1.002)	4.732***	(1.093)	4.010***	(1.090)
65-69	7.401***	(1.026)	4.646***	(1.181)	3.648***	(1.179)
70-74	6.642***	(1.121)	3.857***	(1.274)	2.227*	(1.276)
75-80	7.329***	(1.225)	4.502***	(1.366)	2.264*	(1.376)
80+	6.214***	(1.315)	3.320**	(1.448)	0.512	(1.473)
female	1.055*	(0.607)	0.728	(0.611)	0.196	(0.620)
partnership	2.483***	(0.722)	2.045***	(0.729)	2.292***	(0.725)
SPHS					2.669***	(0.729)
ADL					-0.096	(1.077)
MOBILIT					1.299*	(0.684)
CHRONIC					5.376***	(0.658)
underweight					2.922	(2.518)
overweight					-1.072	(0.654)
obese					-2.739***	(0.882)
AT	4.985***	(1.061)	4.774***	(1.062)	5.836***	(1.062)
SE	-27.111***	(1.140)	-29.144***	(1.355)	-28.437***	(1.358)
NL	-43.762***	(1.401)	-43.937***	(1.401)	-43.124***	(1.399)
ES	-39.257***	(1.547)	-38.567***	(1.557)	-38.437***	(1.553)
IT	-28.531***	(1.321)	-28.255***	(1.325)	-27.972***	(1.319)
FR	-19.900***	(1.262)	-19.927***	(1.260)	-19.734***	(1.254)
DK	-37.381***	(1.528)	-39.464***	(1.668)	-38.908***	(1.674)
GR	-41.405***	(1.423)	-40.500***	(1.441)	-39.623***	(1.442)
CH	-17.448***	(1.461)	-17.856***	(1.487)	-16.390***	(1.489)
BE	-34.002***	(1.162)	-34.073***	(1.165)	-33.814***	(1.167)
Observations	17330		17330		17330	
Log-likelihood	-6747		-6731		-6666	

Standard errors (SE) in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Marginal Effects (ME): all coefficients are multiplied by 100. Base categories: low education qualification; non working status; age 50–54; man; living as a single; bmi 'normoweight'; Germany(DE). Column (1) is the base model with education qualification, country dummies, and demographic characteristics; (2) + logincome and occupation; (3) + health controls (sphs, adl, mobility, chronic conditions, BMI). Calculations based on 2004 SHARE data (drop-off questionnaire).

Table 2.6: Determinants of Preventive Care – Eye Exam

	(1)		(2)		(3)	
	ME	SE	ME	SE	ME	SE
MIDEDU	9.322***	(0.790)	8.783***	(0.795)	9.435***	(0.801)
HIGHEDU	11.975***	(0.950)	10.754***	(0.972)	11.672***	(0.985)
logincome			1.933***	(0.302)	1.903***	(0.304)
job			1.206	(0.953)	2.872***	(0.967)
55–59	0.85	(1.028)	1.115	(1.046)	0.125	(1.051)
60–64	1.813*	(1.057)	2.441**	(1.156)	0.96	(1.163)
65–69	3.957***	(1.104)	4.821***	(1.273)	2.846**	(1.282)
70–74	6.218***	(1.189)	7.176***	(1.357)	4.622***	(1.371)
75–80	10.871***	(1.325)	11.796***	(1.477)	8.486***	(1.498)
80+	12.688***	(1.431)	13.578***	(1.572)	10.065***	(1.615)
female	8.948***	(0.662)	9.078***	(0.669)	8.777***	(0.685)
partnership	4.124***	(0.791)	3.415***	(0.800)	3.802***	(0.804)
SPHS					0.958	(0.806)
ADL					-0.882	(1.203)
MOBILIT					1.852**	(0.768)
CHRONIC					10.719***	(0.738)
underweight					-6.783**	(2.977)
overweight					1.227*	(0.731)
obese					-0.406	(0.959)
AT	-9.543***	(1.661)	-9.324***	(1.664)	-8.046***	(1.674)
SE	-17.233***	(1.513)	-22.378***	(1.700)	-21.855***	(1.716)
NL	-28.514***	(1.516)	-28.878***	(1.518)	-27.982***	(1.527)
ES	-18.378***	(1.686)	-17.049***	(1.700)	-17.409***	(1.708)
IT	-12.325***	(1.603)	-11.464***	(1.610)	-11.389***	(1.617)
FR	5.212***	(1.701)	5.114***	(1.703)	5.625***	(1.711)
DK	-12.376***	(1.747)	-17.079***	(1.888)	-16.877***	(1.908)
GR	-12.933***	(1.646)	-11.235***	(1.668)	-10.822***	(1.681)
CH	-0.869	(1.812)	-2.718	(1.833)	-0.553	(1.852)
BE	-8.082***	(1.511)	-7.613***	(1.513)	-7.628***	(1.525)
CZ	-13.765***	(1.863)	-17.729***	(1.963)	-17.925***	(1.972)
PL	-28.440***	(1.876)	-27.483***	(1.884)	-28.870***	(1.898)
year 2004	-0.343	(0.953)	-0.617	(0.955)	-0.091	(0.959)
Observations	24512		24512		24512	
Log-likelihood	-15904		-15881		-15734	

Standard errors (SE) in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Marginal Effects (ME): all coefficients are multiplied by 100. Base categories: low education qualification; non working status; age 50–54; man; living as a single; bmi ‘normoweight’; Germany(DE). Column (1) is the base model with education qualification, country dummies, and demographic characteristics; (2) + logincome and occupation; (3) + health controls (sphs, adl, mobility, chronic conditions, BMI). Calculations based on both 2004 and 2006 SHARE data (drop-off questionnaire).

Table 2.7: Determinants of Preventive Care – Mammogram

	(1)		(2)		(3)	
	ME	SE	ME	SE	ME	SE
MIDEDU	7.718***	(1.151)	7.286***	(1.160)	7.312***	(1.169)
HIGHEDU	9.030***	(1.509)	8.225***	(1.547)	8.327***	(1.566)
logincome			1.814***	(0.430)	1.796***	(0.432)
job			-0.902	(1.448)	-0.78	(1.464)
55-59	4.490***	(1.521)	4.269***	(1.555)	3.759**	(1.562)
60-64	0.752	(1.551)	0.336	(1.689)	-0.499	(1.700)
65-69	-7.868***	(1.615)	-8.430***	(1.819)	-9.760***	(1.836)
70-74	-23.862***	(1.751)	-24.416***	(1.948)	-25.896***	(1.976)
75-80	-40.800***	(1.989)	-41.483***	(2.171)	-43.383***	(2.213)
80+	-56.803***	(2.296)	-57.481***	(2.459)	-58.808***	(2.534)
partnership	7.185***	(1.082)	6.356***	(1.100)	6.384***	(1.104)
SPHS					0.401	(1.227)
ADL					-5.494***	(1.732)
MOBILIT					0.627	(1.128)
CHRONIC					4.956***	(1.098)
underweight					-12.401***	(3.689)
overweight					1.678	(1.097)
obese					-4.203***	(1.380)
AT	25.826***	(2.332)	25.866***	(2.335)	26.576***	(2.349)
SE	38.168***	(2.266)	33.824***	(2.498)	34.060***	(2.515)
NL	43.304***	(2.272)	42.945***	(2.274)	43.406***	(2.286)
ES	18.050***	(2.406)	19.317***	(2.427)	19.347***	(2.437)
IT	19.435***	(2.255)	20.278***	(2.268)	20.243***	(2.276)
FR	42.833***	(2.460)	42.759***	(2.461)	43.219***	(2.474)
DK	-21.311***	(2.586)	-25.604***	(2.783)	-25.356***	(2.812)
GR	-3.667	(2.381)	-2.164	(2.414)	-2.164	(2.425)
CH	5.376**	(2.500)	3.903	(2.529)	4.995*	(2.554)
BE	28.184***	(2.176)	28.565***	(2.181)	28.674***	(2.196)
CZ	7.944***	(2.629)	4.319	(2.766)	4.48	(2.775)
PL	-6.582**	(2.692)	-5.950**	(2.705)	-5.814**	(2.724)
year 2004	-0.693	(1.442)	-0.949	(1.444)	-0.62	(1.448)
Observations	13097		13097		13097	
Log-likelihood	-7258		-7249		-7219	

Standard errors (SE) in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Marginal Effects (ME): all coefficients are multiplied by 100. Base categories: low education qualification; non working status; age 50–54; living as a single; bmi ‘normoweight’; Germany(DE). Column (1) is the base model with education qualification, country dummies, and demographic characteristics; (2) + logincome and occupation; (3) + health controls (sphs, adl, mobility, chronic conditions, BMI). Female respondents only. Calculations based on both 2004 and 2006 SHARE data (drop-off questionnaire).

Chapter 3

The Fixed-Effects Zero-Inflated Poisson Model**

3.1 Introduction

Count data models have become increasingly popular in many fields of empirical economics and other social sciences; see, for example, Cameron and Trivedi (1998), Wooldridge (2002, Chapter 19), Winkelmann (2003), or Cameron and Trivedi (2005, Chapter 20). Applications include, for example, studies in transportation (on the number of accidents or trips), demography (on the number of births), health economics (on the number of doctor visits or hospital stays), industrial organization (on the number of patents), marketing (on the number of products purchased) and labor economics (on the number of job market transitions, for example). Models for cross-section data range from the standard Poisson model to models allowing for overdispersion such as the negative binomial model, and hurdle models or zero inflated models that account for unusually large numbers of zero outcomes (see, e.g., Lambert 1992). Our focus here is on the latter type of models.

Count outcomes are particularly common in many medical and public health studies, with data that often present a large number of zeros. In order to adjust for extra zero counts, and to avoid biased parameter estimates and misleading inferences, various modifications of the Poisson regression model have been proposed. In the recent literature, various examples of empirical modeling of count measures of health care can be found. There are mainly two streams of literature, one that considers utilization of health care as a two-part decision making process (hurdle models)[1], given the latent class, and distinguishes between users and non-users; the other approach that considers individuals belonging to latent classes and distinguishes between low and high users (finite mixture negative binomial models)[2]. Deb and Trivedi (1997, 2002) argue that the distinction between low and

---

**This chapter is based on "The Fixed-Effects Zero-Inflated Poisson Model with an Application to Health Care Utilization", M.C. Majo, and A. van Soest. *Acknowledgments:* I am particularly grateful to Arthur van Soest, Jochem de Bresser and Vasilios Andrikopoulos for valuable assistance.

[1]Mullahy 1986, Pohlmeier and Ulrich 1995.

[2]Deb and Trivedi 1997.

high users of health care is a better approach, which has been supported by the subsequent literature (see, for example, Deb and Holmes 2000). In some applications, and given different distributional assumptions on the traditional hurdle model[3], it has been found that the hurdle model performs better than the finite mixture models. However, Winkelmann (2004) shows that the finite mixture approach outperforms the traditional hurdle model, unless in the latter different distributional assumptions are made.

Since the seminal article of Hausman et al. (1984), many studies have also used panel data models for count data, such as the (static or dynamic) fixed-effects Poisson and negative binomial models and a random effects version of the (static) zero-inflated Poisson model (Crepon and Duguet 1997; Wang et al. 2002). Fixed-effects models are more flexible than random effects models and are often found to outperform the corresponding random effects models in empirical studies. To the best of our knowledge, there are no existing studies that use a fixed-effects version of the (static) zero-inflated Poisson model. This study fills this gap. We show that the zero inflated Poisson model with fixed-effects can be estimated in a similar way as the fixed-effects logit and the fixed-effects Poisson or negative binomial models. We then apply this model and show that it outperforms several existing panel data models for count data. We analyze data having excess zero counts, namely three types of health care service utilization, using micro level data from the first two waves (2004 and 2006) of the Survey of Health, Ageing and Retirement in Europe (SHARE).[4] We compare our zero inflated Poisson model with fixed-effects (ZIP_FE) with the Poisson (P) and the negative binomial (NB) model, in order to determine which model better fits the data. We conclude that ZIP_FE represents an interesting alternative to other panel data models for count data with excess zeros.

The remainder of the paper is organized as follows: Section 3.2 describes frequently applied count data models and introduces the zero inflated model (ZIP) and its extension with fixed-effects for panel data (ZIP_FE). Section 3.3 presents the data that we use for the application. Section 3.4 presents the results and Section 3.5 concludes.

3.2 Panel Data Models for Count Data

3.2.1 Poisson and Negative Binomial Models

A frequently applied model for the distribution of the count observations Y_{it} in panel data is the Poisson (P) regression model. It assumes that the conditional distribution of Y_{it} for individual (or cross-section unit) i in time period t, given (strictly exogenous) regressors X_{it} and an individual effect α_i, is a Poisson distribution with parameter μ_{it}:

$$Po(y;\mu_{it}) = exp(-\mu_{it})\mu_{it}^{y}/y!, \qquad \text{for} \quad y = 0, 1, 2, \ldots, \tag{3.1}$$

where

$$\mu_{it} = exp(x_{it}'\beta + \alpha_i) \tag{3.2}$$

[3]See for example Jimenez-Martin et al. 2002, and Bago D'Uva 2006.

[4]Börsch-Supan and Jürges 2005.

Here β is a vector of unknown parameters to be estimated. In the fixed-effects version of the model, no assumptions are made on α_i and they are treated as unknown nuisance parameters. In the random effects version, it is assumed that the α_i are independent of all X_{it} and follow a specific distribution, usually a Gamma distribution (with a mean normalized to one). Finally, the pooled version of the model treats the panel data set as a cross-section, assuming $\alpha_i = 0$ for all i.

The Poisson model has the properties

$$E(Y_{it}|X_{it}, \alpha_i) = Var(Y_{it}|X_{it}, \alpha_i) = \mu_{it} \tag{3.3}$$

It therefore assumes that data are "equidispersed": the conditional variance is equal to the conditional mean. In practice, it is often found that this assumption is too restrictive, and the data are better described by a model allowing for "overdispersion", that is a variance that is larger than the mean.

The most common model allowing for overdispersion is the negative binomial model (NB). The NB model accounts for overdispersion through an additional parameter θ:

$$Pr(Y_{it} = y \,|\, \mu_{it}, \theta) = \frac{\Gamma(y + \theta^{-1})}{y!\,\Gamma(\theta^{-1})} \left(\frac{\theta^{-1}}{\theta^{-1} + \mu_{it}}\right)^{\theta^{-1}} \left(\frac{\mu_{it}}{\theta^{-1} + \mu_{it}}\right)^{y}, \tag{3.4}$$

for $\quad y = 0, 1, 2, \ldots.$

Here μ_{it} is defined in the same way as for the Poisson model. In the NB model, we have:

$$E(Y_{it}|X_{it}, \alpha_i) = \mu_{it} \qquad \text{and} \qquad Var(Y_{it}|X_{it}, \alpha_i) = \mu_{it} + \theta\, \mu_{it}^2 \tag{3.5}$$

The parameter θ therefore reflects overdispersion. The NB model can be derived as a mixture distribution of a Poisson model in which the Poisson parameter follows a Gamma distribution with coefficient of variation (standard error divided by the mean) equal to $\sqrt{\theta}$ (Cameron and Trivedi 2005, p. 675); the Poisson model is the special case of the NB model with $\theta = 0$.

We use the type of parametrization of the NB as defined by Hausman et al. (1984)[5] and follow their derivation of the model. Assuming that, for a given individual i, the y_{it} are independent over time, then $\sum_t y_{it}$ has a NB distribution with parameters θ_i and $\sum_t \mu_{it}$. Conditional on these total counts, the conditional fixed-effects NB log-likelihood contribution of individual i is given by:

$$\begin{aligned} L_i \;=\; & ln\Gamma(\sum_t \mu_{it}) + ln\Gamma(\sum_t y_{it} + 1) - \sum_t ln\Gamma(y_{it} + 1) \\ & - ln\Gamma(\sum_t y_{it} + \sum_t \mu_{it}) + \sum_t (ln\Gamma(y_{it} + \mu_{it})) - \sum_t (ln\Gamma(\mu_{it})) \end{aligned} \tag{3.6}$$

Note that the heterogeneity parameter θ does not appear in the log-likelihood, so that θ is not estimated. Standard numerical maximization routines can be applied to maximize

[5]See also Cameron and Trivedi (1998).

the log-likelihood function of the conditional fixed-effects estimator, and estimation procedures are already implemented in several econometric packages (e.g. Stata). Allison and Waterman (2002) demonstrate that this model is not a true fixed-effects method, because the individual effects and the covariates do not enter in exactly the same way; in particular, they influence the conditional variance in different manners; see Rabe-Hesketh and Skrondal (2008). As a consequence, it is possible in this model to estimate the coefficients of time invariant regressors.

3.2.2 Zero-inflated Poisson Model

It often happens that the data are characterized by a larger frequency of extra zeros than a P model or an NB model predicts, and that whether or not the outcome is zero is driven by different factors than the mean of the positive outcomes. A popular approach to account for these features of the data is the zero inflated Poisson regression model (ZIP; Lambert 1992). The ZIP distribution can be seen as a mixture of the Poisson distribution (with probability p) and a degenerate distribution with point mass one at zero (with probability $(1-p)$; see Johnson et al. 1992, or Lambert 1992). For a Poisson distribution with parameter μ, this gives the following probability function:

$$f(y;\tilde{p},\mu) = \begin{cases} (1-\tilde{p}) + \tilde{p}\,Po(0;\mu) & \text{if } y=0, \\ \tilde{p}\,Po(y;\mu) & \text{if } y>0 \end{cases} \tag{3.7}$$

Here $0 < \tilde{p} \leq 1$. The Poisson distribution is the special case with $\tilde{p} = 1$. If $\tilde{p} < 1$, the distribution has a larger probability of zero outcomes than the corresponding Poisson distribution. It is easy to show that the mean and variance of this distribution are given by:

$$E(Y) = \tilde{p}\,\mu \qquad \text{and} \qquad Var(Y) = \tilde{p}\,\mu + \tilde{p}(1-\tilde{p})\,\mu^2 \tag{3.8}$$

Thus the ZIP model also incorporates (a special form of) overdispersion: for all $\tilde{p} < 1$, the variance is larger than the mean.

A problem with the ZIP distribution written in this way is that there are two types of zeros: the extra zeros, and the zeros from the Poisson model. This makes it hard to say something about $\tilde{p}$ without also estimating μ. This problem can be avoided by writing the ZIP distribution in an alternative way – as a mixture of a *truncated* Poisson distribution (μ) and a degenerate distribution with all its mass at zero, with weights $p = \tilde{p}[1 - Po(0;\mu)]$ and $1-p$ (see, e.g., Lee et al. 2002):

$$f(y;p,\mu) = \begin{cases} (1-p) & \text{if } y=0, \\ p\,Po(y;\mu)/[1-Po(0;\mu)] & \text{if } y>0 \end{cases} \tag{3.9}$$

The probability mass function of this distribution can also be written as:

$$Pr(y|\mu) = \begin{cases} (1-p) & \text{if } y = 0, \\ p\,\dfrac{exp(-\mu)\mu^y}{y!\,[1-exp(-\mu)]} & \text{if } y > 0 \end{cases} \tag{3.10}$$

This parametrization has the advantage that $1-p$ is simply the probability of outcome zero, while μ is now the parameter of the truncated Poisson distribution describing the non-zero outcomes. As a consequence, and as will be demonstrated below, it is more convenient to take this parametrization as the starting point of the econometric model than to take the parametrization with $\tilde{p}$.

To obtain the (static) zero inflated panel data model, we specify p and μ for each observation (i,t) as follows:

$$p_{it} = \frac{exp(X'_{it}\beta^p + \alpha_i^p)}{1 + exp(X'_{it}\beta^p + \alpha_i^p)} \tag{3.11}$$

$$\mu_{it} = exp(X'_{it}\beta^\mu + \alpha_i^\mu) \tag{3.12}$$

We consider the fixed-effects version of the model – we make no assumptions on the individual effects α_i^p and α_i^μ and treat them as nuisance parameters. The parameters of interest are β^p and β^μ. The parameters β^p determine which factors determine whether Y_{it} is zero or not; the parameters β^μ determine the conditional distribution of Y_{it} (and its mean and variance) given that Y_{it} is positive.

Estimation of β^p is straightforward, since our model specification implies that whether Y_{it} is positive or not is now explained by a fixed-effects logit model. We can therefore estimate β^p using the standard conditional maximum likelihood estimator of Chamberlain (1980). For the case of two time periods (as in our empirical example), this boils down to estimating a binary logit model explaining whether i changes from $Y_{i1} = 0$ to $Y_{i2} > 0$ in the subsample of observations with $Y_{i1} = 0$ and $Y_{i2} > 0$ or $Y_{i1} > 0$ and $Y_{i2} = 0$ (discarding all the other observations), with regressors $X_{i2} - X_{i1}$. The estimates of the slope coefficients in this logit model are consistent estimates for β^p.[6]

Estimation of β^μ is less standard. We focus on the case of two time periods ($t = 1, 2$), which is also what we have in our empirical example. First, we discard all observations with $Y_{i1} = 0$ or Y_{i2}=0. Second, we apply conditional maximum likelihood on the remaining observations, conditioning on $Y_{i1} + Y_{i2}$. This is similar to the usual conditional maximum likelihood for the FE Poisson model, but using the truncated Poisson distribution instead of the Poisson distribution. Starting from the truncated Poisson distribution with probabilities

$$Pr(y_{it} = k|X_{it}, \alpha_i^\mu, y_{it} > 0) = \frac{\mu_{it}^k exp(-\mu_{it})}{k!\,(1 - exp(-\mu_{it}))}, \tag{3.13}$$

with

$$k = 1, 2, \ldots; \qquad t = 1, 2; \qquad \mu_{it} = exp(x'_{it}\,\beta^\mu + \alpha_i^\mu),$$

[6] As always in fixed-effects models, only time varying regressors can be included.

and using that outcomes in the two time periods are conditionally independent given X_{it} (and α_i^{μ}), it can be easily shown that the conditional likelihood contribution for an observation i with $y_{i1} = k > 0$ and $y_{i2} = w-k > 0$, conditional on $X_{i1}, X_{i2}, \alpha_i^{\mu}$, $y_{i1}+y_{i2} = w$, $y_{i1} > 0$ and $y_{i2} > 0$, is given by:

$$\begin{aligned} LC_i &= P(y_{it} = k \,|\, y_{i1} + y_{i2} = w, y_{i1} > 0, y_{i2} > 0, X_{i1}, X_{i2}, \alpha_i^{\mu}) = \\ &= \frac{w!\, \mu_{i1}^{k} \mu_{i2}^{(w-k)}}{k!\,(n-k)!\,[(\mu_{i1} + \mu_{i2})^{w} - \mu_{i1}^{w} - \mu_{i2}^{w}]} \end{aligned} \tag{3.14}$$

With $\lambda_{it} = exp(X_{it}'\beta^{\mu}) = \mu_{it} exp(-\alpha_i^{\mu})$, this can also be written as

$$\begin{aligned} P(y_{it} &= k \,|\, y_{i1} + y_{i2} = w, X_{i1}, X_{i2}) = \\ &= \frac{w!\, \lambda_{i1}^{k} \lambda_{i2}^{(w-k)}}{k!\,(n-k)!\,[(\lambda_{i1} + \lambda_{i2})^{w} - \lambda_{i1}^{w} - \lambda_{i2}^{w}]} \end{aligned} \tag{3.15}$$

The important thing is that this expression no longer depends on α_i^{μ}: as in the FE-Poisson model (see Hausman et al. 1984, for example), in this FE-truncated Poisson model, the sum of the outcomes $y_{i1} + y_{i2}$ is a sufficient statistic for the individual effect α_i^{μ}. As a consequence, this conditional maximum likelihood estimator maximizing $\sum LC_i$ (where the summation is over the subsample of observations with $Y_{i1} > 0$ and $Y_{i2} > 0$) only involves maximization over β^{μ} and will give a consistent estimator of β^{μ}.

The actual estimation can be done using maximum likelihood routines in Stata (see Gould et al. 2006). The syntax for the conditional likelihood to estimate β^{μ} is given in Appendix B (in Stata 9.2).

The ZIP_FE model combines two attractive features of count data models. First, it makes it possible to account for fully flexible fixed individual effects in panel data, whereas previous applications of the ZIP model have either used cross-sectional data, or, in a few cases, ZIP models with random effects, which are more restrictive than the fixed-effects specification. For example, Wang et al. (2002) used a random effects ZIP model to account for inter-hospital variation in hospital stays within diagnosis related groups, and Crepon and Duguet (1997) used a random effects ZIP model to analyze innovation in firms on the basis of the number of patents. To our knowledge our current study is the first time that fixed-effects are introduced in a ZIP setting. Second, the ZIP_FE model has the same flexibility of the ZIP model for cross-section in dealing with zero observations. While our derivations (and the Stata code in Appendix B) are for the case of two panel waves only, generalizing the estimator to the case of more than two waves is in principle straightforward. It requires much more notation and programming, however, and is therefore left for future work.

3.3 Data

This paper uses data from the Survey of Health, Ageing and Retirement in Europe (SHARE). It is a multidisciplinary and cross-national panel database of micro data on health, socio-economic status, and social and family networks of individuals aged 50 or over. The project started in 2004 (baseline study) in 11 European countries. In 2006 the second wave has been carried out, extending the study to four additional countries: the Czech Republic (CZ), Ireland (IE), Israel (IL), and Poland (PL). Since we are interested in the longitudinal dimension of the data, we consider only those countries that present data in both waves, namely: Austria (AT), Belgium (BE), Denmark (DK), France (FR), Germany (DE), Greece (GR), Italy (IT), the Netherlands (NL), Spain (ES), Sweden (SE), and Switzerland (CH). The survey provides extensive information at both the household and the individual level (see Börsch-Supan and Jürges 2005), and ensures a good level of comparability across countries and over time.

The final sample consists of 34,350 observations – a balanced panel of 17,175 individuals observed in two years. The outcome variables representing health care utilization that are used here are the number of doctor visits during the past twelve months (DOCT), the number of visits to a general practitioner during the past twelve months (GP), and the number of visits to a specialist and outpatient treatments in a clinic or an emergency room (SPOUTER) during the past twelve months. The SPOUTER variable, defined as a count, has been obtained as the difference between the reported variables counting DOCT and GP visits. To be precise, respondents were asked to answer the following questions:

- **HC002_**: "Since last year, about how many times in total have you seen or talked to a medical doctor about your health? Please exclude dentist visits and hospital stays, but include emergency room or outpatient clinic visits." (0,...,98).
- **HC003_**: "How many of these contacts were with a general practitioner or with a doctor at your health care center?" (0,...,98).

Table 3.1 shows how the dependent and independent variables used in our analysis are defined. The covariates are in line with the factors usually considered to explain the demand for health care. The socio-economic characteristics include the logarithm of family income adjusted for household size[7] (LOGINCOME), and occupational status categorized as employed (EMP), retired (RETIRED), and not employed (NOTEMP; the base category).[8] Gender (FEM), age (AGE) and educational qualification (categorized as low, EDUQUAL1, the base category; medium, EDUQUAL2; and high, EDUQUAL3) are added to those models where time invariant regressors can be included.[9] Household composition is controlled for using a dummy for living with a partner or having a spouse (MSTAT2, with living as a

[7]Total household income has been divided by the square root of the household size; the imputations provided by the SHARE team were used to replace missing values.

[8]We also controlled for household wealth, but it was never significant; we therefore excluded it from the final model. Results are available upon request.

[9]Thus to the pooled and random effects models, see Section 3.4.

single being the base category). Health status variables considered are: a dummy whether the individual considers his health to be less than good (SPHS), and dummy variables for the prevalence of at least two chronic conditions (CHRONIC), one ore more limitations with activity of daily living (ADL), and one ore more physical limitations (MOBILIT).

Table 3.2 shows summary statistics of our estimation sample for each of the two waves. The changes in the means from wave 1 to wave 2 are all in line with the notion that respondents are older and less healthy in wave 2 than in wave 1. In the second wave, they are more often retired and less often employed, have lower income, are more likely to have lost their spouse, more often have health problems, and more often visit a doctor than in the first wave. In all cases, the three outcome variables DOCT, GP, and SPOUTER, present evidence of strong overdispersion, with the unconditional variance being much larger than the mean. Table 3.3 shows summary statistics of our dependent variables DOCT, GP, and SPOUTER, only for positive counts in both waves. This provides a first evidence that the process underlying the contact decision is different from the second stage process, that is, once the contact has been made. We can see that, once we remove the zero counts, the distribution changes in all cases, presenting a smaller overdispersion than in the full sample.

Table 3.4 and Figure 3.1 show the distribution of the three outcome variables. The maximum number of consultations is 98 for each of the three services.[10] It can be seen that, especially for SPOUTER visits, there is a large number of zeros, with more than 50% of the respondents reporting zero visits in both waves. For DOCT and GP visits the distribution is less skewed than for the SPOUTER distribution, but still, a large number of zeros is found in both cases (almost 15% and 20% of zero counts, respectively).

In this situation of highly overdispersed data and a large frequency of extra zeros in the distribution, the traditional count data models, such as the P and the NB, may not be appropriate to fit the health care utilization data, whereas their inflated variants may be more appropriate. On the other hand, overdispersion and zeros can also be explained by individual effects, and the extent to which they do is not something that can be derived from the raw data. The next section will address this by comparing the estimates of various panel data models, focusing on the ZIP_FE model introduced in Section 3.2.

3.4 Application to Health Care Utilization Data: Results

This section presents the estimation results for several cross-section and panel data versions (pooled, random effects, and fixed-effects) of the P and the NB model, and for the ZIP_FE model introduced in Section 3.2. All models use the same estimation sample of 34,350 observations (balanced panel of 17,175 individuals observed twice), P and NB (both random and fixed-effects), and ZIP_FE, which is defined by the expression given in Section 3.2.

[10]This is the maximum number that can be reported; respondents with more than 98 visits are also coded as 98.

3.4.1 Poisson and Negative Binomial Models

Tables 3.5, 3.6, and 3.7 show the estimation results for the three types of health care service use that we consider. The models used in these tables have all been presented in Section 3.2.

It is interesting to compare the results for the panel data models to the results for the P and NB with pooled data both for the estimates obtained and also the precision of the estimates. The parameter estimates generally seem more precise in the panel data models, which have smaller standard errors. This may be because both point estimates and standard errors in the pooled model are estimated inconsistently if individual effects matter. Most of the estimated coefficients have the same sign in the three models, but there are a few notable exceptions. In particular, it is interesting to note that in all three cases (DOCT, GP and SPOUTER), the fixed-effects model specifications differ from the pooled and random effects panel models where it comes to the coefficients of LOGINCOME: income has a positive and significant effect according to the fixed-effects specifications, whereas in both the pooled and the random effects panel models these coefficients are negative and significant for all three cases (DOCT, GP, and SPOUTER), and insignificant only for SPOUTER visits in the random effects P. This suggests that individual effects are negatively correlated with income, leading to a negative bias in the estimates of the logincome coefficients in the pooled and random effects models. The effect of RETIRED is negative and significant in the pooled model (with the exception of the NB where it is not significant for GP, and SPOUTER visits). In the NB it is always insignificant in both the random and the fixed-effects specifications in all three cases of consultations, differing from the P where it is always negative and significant in all three cases of DOCT, GP and SPOUTER visits. Marital status (MSTAT2) also changes sign, though the pattern is less consistent across models and dependent variables and it is often insignificant in the non-fixed-effects models. All the estimated coefficients of the other variables have the same sign and significance in the three models.

We can see also that the overdispersion parameter θ is particularly large in the SPOUTER visits case, where the difference between the variance and the mean was the largest (see Table 3.2).

Tables 3.8 presents the model selection tests. To assess which model between P and NB (random effects) performs better, the significance of the θ parameter can be tested by a likelihood ratio test (since the two models are nested), with $H_0 : \theta = 0$ versus $H_1 : \theta \neq 0$. For all three health care services analyzed, θ is significantly different from zero, implying that NB is preferred over P. We use a Hausman test to choose between random and fixed-effects models (for both P and NB and for all three health care services). The Hausman test tests the null hypothesis that the random effects assumptions on the individual effects are valid, against the fixed-effects alternative without assumptions on the individual effects. The small p-values in the table indicate that random effects models are rejected against the corresponding fixed-effects models in all cases, implying that fixed-effects models are preferred in all cases.

3.4.2 ZIP_FE

Table 3.9 shows the estimates of the parameters of the model ZIP_FE. As explained in Section 2, the ZIP_FE generates two separate models. First, a count data model predicts counts of the truncated Poisson model for respondents with at least one visit. Second, a fixed-effects logit model is used to explain whether an outcome is zero or not. This model uses only the transitions from zero to a positive outcome or the reverse. If we look at Table 3.9, the first part 'COUNT' is the response variable (DOCT, or GP, or SPOUTER) predicted by the truncated model estimated by conditional maximum likelihood, and the second part 'LOGIT' refers to the logistic model predicting whether a respondent is likely to have at least one visit in a given year. We first look at the 'COUNT' portion of the output, which refers to the respondents who have at least one consultation per year. The effect of income on the number of DOCT visits in a year is significantly positive (holding all other variables in the model constant) and the effect is similar in size to the effect in the fixed-effect models in Table 3.5. The same is for GP visits, with an increase of GP consultations in a year by a factor of about $exp(0.029) = 1.029$ for every unit increase in the logincome. The income effect is not significant for SPOUTER visits. If we compare it with the models in the previous section, we see that the sign is the same that we had in the fixed-effects models, with the exception of SPOUTER, where coefficients were positive and significant. If we look at the 'LOGIT' portion of the output, which predicts whether outcomes are positive or zero, we find a significant effect for SPOUTER only: the higher a respondent's logincome, the more likely the respondent will have a visit. The estimated marginal effect of a 10 percent income increase for an average respondent (with probability 0.48 that SPOUTER is positive) is about 0.10*0.043*0.48*(1-0.48)*100% = 0.11 percentage points.

The estimated coefficients for MSTAT2 are positive and significant for the number of visits, in line with the fixed-effects models in the previous section, whereas in both the pooled and the random effects panel models these coefficients were negative and significant. Respondents who are married or living with a partner tend to visit a doctor more often than single respondents, once they have decided to go at least once (keeping all other variables in the model constant). In the 'LOGIT' part of the model, however, we find the opposite effect: a non-single-respondent has less probability to have a DOCT or a GP consultation than a single respondent with identical scores for the other predictors. This is an example where the effect in the two equations is quite different, supporting the use of the ZIP model which has the flexibility to capture this.

All the other variables are consistent with the models presented in the previous section for the 'COUNT' part of the model. Occupational status is not significant in the 'LOGIT' portion of the model. If we look at the 'COUNT' portion of the model, an employed (retired) respondent decreases her SPOUTER visits by $exp(0.067) = 1.07$ ($exp(0.061) = 1.06$) compared to a respondent who is nor employed neither retired, everything else being the same. Health status is positive and significant for all estimated coefficients in both the 'COUNT' and the 'LOGIT' model portions (where a higher score in the health status variable, means a worse health status for the respondent), with the exception of ADL that

is not significant for the zero/positive decision.
All in all we find a strong income-health care visit gradient for the number of visits given that this is positive for DOCT and GP, while the income effect is absent in the 'LOGIT' portion of the model. In SPOUTER visits we find the opposite, the income-health care visit gradient is in the decision to have at least one visit or not.

Table 3.10 shows the log-likelihood, AIC and BIC (respectively, Akaike and Schwarz information criteria) for the estimated models. The information criteria AIC and BIC are used in comparison of non-nested models, where a log-likelihood test cannot be performed. The ZIP_FE model outperforms all the alternative models for GP and SPOUTER, whereas the fixed-effects NB should be preferred over the other models for DOCT visits. This results are also in line with Table 3.4, where we showed excess zeros for both GP and SPOUTER.

3.5 Conclusions

In this paper we defined and estimated a zero-inflated Poisson model with fixed-effects to identify respondent- and health-related characteristics associated with health care demand using a two-wave panel. This is a new model that is proposed to model count measures of health care utilization and account for the panel structure of the data. The estimation method and syntax developed in this paper can accommodate ZIP models with fixed-effects in both the logistic (already available in Stata) and the truncated Poisson part (for which we have developed the syntax). The computer program for the maximum likelihood estimation in Stata provides a flexible tool for analyzing the health care service count variables. We find that controlling for the portion of respondents that are certain zeros in one of the two years of the two waves does make a difference for counts with a larger number of zeros, where traditional count data models are not able to disentangle the effects. All in all we find a strong income-health care visit gradient for the "non certain zeros" group for DOCT and GP, while the income effect is absent in the "certain zeros" group. In SPOUTER visits we find the opposite, the income-health care visit gradient is in the "certain zeros" group. In general, the previous applications of the ZIP model have used cross-sectional data, with a few exceptions to random effects. To our knowledge this is the first time that fixed-effects are introduced in a ZIP setting. The ZIP_FE model has some attractive features. It makes it possible to account for individual effect in panel data: fixed-effects can explain overdispersion, where P model can not. It allows the correction for extra zeros defining two latent classes of low users in the probability of visiting a doctor, and high users in the conditional positive number of visits. Extending the estimator and the estimation algorithm to the case of more than two time periods and developing model selection tests will be further steps in future research developments.

Table 3.1: Variables Definition

Variable Name	
DOCT	number of visits to a medical doctor (GP, specialist, outpatient, ER) last year
GP	number of visits to a general practitioner (GP) last year
SPOUTER	number of doctor visits excluding GP (specialist, outpatient, ER) last year
logincome	ln of annual household income (€), adjusted for household size
emp	occupational status; 1 if employed
retired	occupational status; 1 if retired
fem	gender; 1 if female
eduqual2	1 if medium educational qualification
eduqual3	1 if high educational qualification
age	respondent's age at the time of the interview
mstat2	partnership status; 0 if single, 1 if married or living with a partner
sphs	1 if the respondent considers her health status to be less than good
chronic	1 if the respondent has 2 or more chronic conditions
mobilit	1 if the respondent has 1 or more mobility limitations
adl	1 if the respondent has 1 or more limitations with activity of daily living
wave	1 if year 2006 (wave 2)

Table 3.2: Summary Statistics by Wave – Full Sample

	Variable	Mean	Std. Dev.	Min.	Max.
WAVE 1					
	DOCT	6.200	9.032	0	98
	GP	4.418	7.098	0	98
	SPOUTER	1.783	4.712	0	98
	logincome	10.196	1.597	0	15.43
	emp	0.288	0.453	0	1
	retired	0.492	0.500	0	1
	eduqual2	0.268	0.443	0	1
	eduqual3	0.221	0.415	0	1
	fem	0.540	0.498	0	1
	age	64.00	9.539	50	99
	mstat2	0.745	0.436	0	1
	sphs	0.675	0.468	0	1
	chronic	0.411	0.492	0	1
	mobilit	0.470	0.499	0	1
	adl	0.082	0.274	0	1
	N	17175			
WAVE 2					
	DOCT	6.753	9.292	0	98
	GP	4.581	6.733	0	98
	SPOUTER	2.172	5.406	0	98
	logincome	9.952	1.768	0	15.43
	emp	0.242	0.428	0	1
	retired	0.543	0.498	0	1
	fem	0.540	0.498	0	1
	eduqual2	0.268	0.443	0	1
	eduqual3	0.221	0.415	0	1
	age	66.00	9.539	52	101
	mstat2	0.729	0.444	0	1
	sphs	0.724	0.447	0	1
	chronic	0.431	0.495	0	1
	mobilit	0.486	0.500	0	1
	adl	0.096	0.295	0	1
	N	17175			

Table 3.3: Summary Statistics by Wave – Positive Counts

	Variable	N	Mean	Std. Dev.	Min.	Max.
WAVE 1						
	DOCT+	13599	7.491	9.504	1	98
	GP+	12045	5.884	7.775	1	98
	SPOUTER+	4921	4.510	6.748	1	98
WAVE 2						
	DOCT+	13599	8.061	9.651	1	98
	GP+	12045	5.945	7.183	1	98
	SPOUTER+	4921	4.891	7.108	1	98

Table 3.4: Fraction of Respondents with Zero and Non-Zero Visits

	Wave 1			Wave 2		
Contacts (0,...,98)	DOCT	GP	SPOUTER	DOCT	GP	SPOUTER
0	**0.14**	**0.21**	**0.57**	**0.13**	**0.18**	**0.52**
1	0.12	0.15	0.13	0.11	0.15	0.12
2	0.12	0.14	0.10	0.11	0.15	0.11
3	0.09	0.09	0.06	0.09	0.10	0.06
4	0.10	0.11	0.04	0.10	0.12	0.05
5	0.07	0.05	0.03	0.07	0.05	0.03
6	0.07	0.06	0.02	0.07	0.06	0.02
7	0.02	0.01	0.01	0.03	0.01	0.01
8	0.03	0.02	0.01	0.04	0.02	0.01
9	0.01	0.01	0.00	0.01	0.01	0.00
$\geq$10	0.23	0.15	0.03	0.24	0.15	0.07
	N 17175			N 17175		

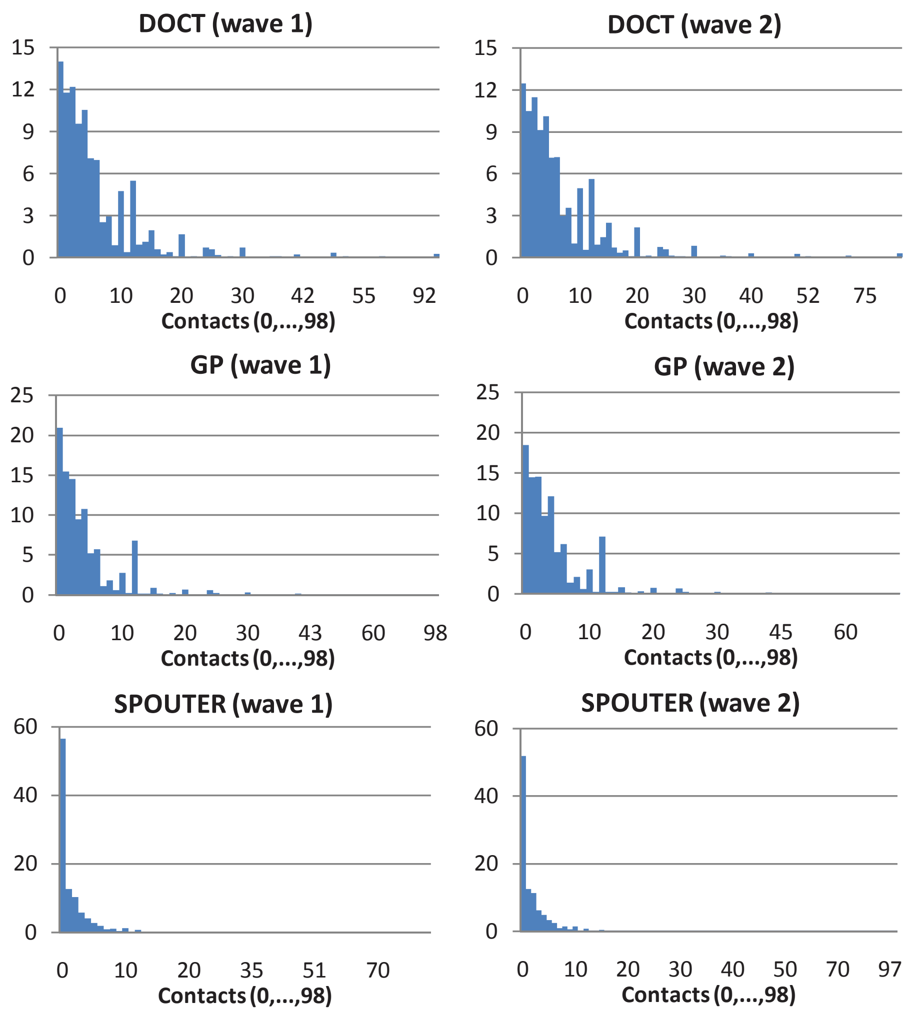


Figure 3.1: Fraction of Respondents with Zero and Non-Zero Visits by Wave

Table 3.5: Doctor Visits

DOCT	POOLED DATA		RANDOM EFFECTS		FIXED EFFECTS	
	P	NB	P	NB	P	NB
logincome	-0.043***	-0.057***	-0.004*	-0.033***	0.024***	0.020***
	(0.004)	(0.004)	(0.002)	(0.003)	(0.003)	(0.005)
emp	-0.310***	-0.275***	-0.251***	-0.199***	-0.154***	-0.056**
	(0.026)	(0.025)	(0.014)	(0.018)	(0.018)	(0.027)
retired	-0.070***	-0.036*	-0.047***	-0.018	-0.038***	0.005
	(0.020)	(0.020)	(0.011)	(0.014)	(0.012)	(0.020)
eduqual2	0.018	0.048**	-0.009	0.067***		
	(0.020)	(0.019)	(0.017)	(0.015)		
eduqual3	0.004	0.052**	-0.060***	0.067***		
	(0.022)	(0.021)	(0.018)	(0.016)		
fem	0.018	0.054***	0.080***	0.088***		
	(0.018)	(0.017)	(0.014)	(0.013)		
age	-0.001	0.000	0.005***	0.003***		
	(0.001)	(0.001)	(0.001)	(0.001)		
mstat2	-0.020	-0.026	-0.011	0.016	0.048**	-0.022
	(0.020)	(0.020)	(0.014)	(0.014)	(0.024)	(0.030)
sphs	0.431***	0.440***	0.336***	0.382***	0.223***	0.185***
	(0.018)	(0.018)	(0.010)	(0.013)	(0.011)	(0.017)
adl	0.341***	0.340***	0.232***	0.200***	0.185***	0.068***
	(0.026)	(0.027)	(0.010)	(0.016)	(0.011)	(0.021)
mobilit	0.275***	0.274***	0.218***	0.220***	0.156***	0.111***
	(0.016)	(0.016)	(0.008)	(0.012)	(0.009)	(0.015)
chronic	0.438***	0.444***	0.291***	0.432***	0.194***	0.205***
	(0.016)	(0.016)	(0.008)	(0.011)	(0.009)	(0.014)
wave	0.031***	0.022*	0.038***	0.040***	0.066***	0.080***
	(0.012)	(0.012)	(0.005)	(0.008)	(0.004)	(0.008)
Constant	1.645***	1.642***	0.994***	0.470***		0.584***
	(0.091)	(0.088)	(0.062)	(0.064)		(0.059)
θ			0.692			
Observations	34350	34350	34350	34350	32418	32418
No. id	17175	17175	17175	17175	16209	16209
Log-likelihood			-104563	-94006	-44973	-34949

Base categories: single, not employed, eduqual1, male, wave 1.
In P and NB fixed-effects only time varying regressors can be included.
Standard errors in parentheses, adjusted for clustering on 17175 id in the pooled model.
*** p<0.01, ** p<0.05, * p<0.1.

Table 3.6: GP Visits

GP	POOLED DATA		RANDOM EFFECTS		FIXED EFFECTS	
	P	NB	P	NB	P	NB
logincome	-0.052***	-0.070***	-0.013***	-0.038***	0.024***	0.023***
	(0.005)	(0.004)	(0.003)	(0.003)	(0.003)	(0.005)
emp	-0.290***	-0.252***	-0.244***	-0.182***	-0.124***	-0.006
	(0.026)	(0.025)	(0.017)	(0.020)	(0.022)	(0.030)
retired	-0.063***	-0.034	-0.041***	-0.009	-0.026*	0.025
	(0.023)	(0.021)	(0.012)	(0.016)	(0.014)	(0.022)
eduqual2	-0.069***	-0.045**	-0.097***	0.015		
	(0.022)	(0.020)	(0.018)	(0.016)		
eduqual3	-0.172***	-0.128***	-0.228***	-0.065***		
	(0.025)	(0.022)	(0.020)	(0.018)		
fem	0.013	0.035*	0.060***	0.058***		
	(0.019)	(0.018)	(0.015)	(0.014)		
age	0.005***	0.006***	0.010***	0.007***		
	(0.001)	(0.001)	(0.001)	(0.001)		
mstat2	-0.038*	-0.048**	-0.022	-0.016	0.092***	-0.008
	(0.021)	(0.020)	(0.015)	(0.015)	(0.028)	(0.033)
sphs	0.397***	0.403***	0.322***	0.347***	0.189***	0.140***
	(0.018)	(0.018)	(0.011)	(0.014)	(0.013)	(0.018)
adl	0.308***	0.307***	0.213***	0.188***	0.160***	0.061***
	(0.028)	(0.028)	(0.012)	(0.018)	(0.013)	(0.023)
mobilit	0.251***	0.251***	0.211***	0.190***	0.139***	0.078***
	(0.017)	(0.017)	(0.009)	(0.013)	(0.011)	(0.016)
chronic	0.402***	0.413***	0.271***	0.392***	0.151***	0.162***
	(0.017)	(0.016)	(0.009)	(0.012)	(0.010)	(0.016)
wave	-0.029**	-0.031***	-0.021***	0.004	0.024***	0.057***
	(0.013)	(0.012)	(0.005)	(0.008)	(0.005)	(0.008)
Constant	1.184***	1.240***	0.563***	0.510***		0.659***
	(0.094)	(0.085)	(0.068)	(0.072)		(0.065)
θ			0.765			
Observations	34350	34350	34350	34350	31074	31074
No. id	17175	17175	17175	17175	15537	15537
Log-likelihood			-89218	-83006	-35515	-29554

Base categories: single, not employed, eduqual1, male, wave 1.
In P and NB fixed-effects only time varying regressors can be included.
Standard errors in parentheses, adjusted for clustering on 17175 id in the pooled model.
*** p<0.01, ** p<0.05, * p<0.1.

Table 3.7: Specialist, Outpatient, and Emergency Room Visits

SPOUTER	POOLED DATA		RANDOM EFFECTS		FIXED EFFECTS	
	P	NB	P	NB	P	NB
logincome	-0.015*	-0.015**	0.004	-0.019***	0.022***	0.020**
	(0.008)	(0.007)	(0.005)	(0.005)	(0.005)	(0.009)
emp	-0.363***	-0.343***	-0.282***	-0.225***	-0.191***	-0.039
	(0.049)	(0.047)	(0.024)	(0.029)	(0.030)	(0.045)
retired	-0.081**	-0.053	-0.069***	0.018	-0.075***	-0.046
	(0.040)	(0.040)	(0.019)	(0.024)	(0.022)	(0.035)
eduqual2	0.218***	0.255***	0.211***	0.233***		
	(0.036)	(0.035)	(0.028)	(0.023)		
eduqual3	0.358***	0.403***	0.319***	0.387***		
	(0.040)	(0.041)	(0.031)	(0.025)		
fem	0.029	0.061*	0.084***	0.182***		
	(0.035)	(0.033)	(0.024)	(0.020)		
age	-0.016***	-0.015***	-0.009***	-0.011***		
	(0.002)	(0.002)	(0.001)	(0.001)		
mstat2	0.029	0.037	0.006	0.160***	-0.085*	0.113**
	(0.040)	(0.038)	(0.024)	(0.022)	(0.044)	(0.046)
sphs	0.507***	0.507***	0.402***	0.432***	0.294***	0.217***
	(0.039)	(0.039)	(0.017)	(0.022)	(0.020)	(0.030)
adl	0.417***	0.421***	0.298***	0.169***	0.240***	0.013
	(0.049)	(0.053)	(0.019)	(0.028)	(0.021)	(0.039)
mobilit	0.335***	0.341***	0.264***	0.260***	0.197***	0.126***
	(0.032)	(0.031)	(0.015)	(0.020)	(0.016)	(0.027)
chronic	0.522***	0.528***	0.401***	0.524***	0.308***	0.279***
	(0.031)	(0.030)	(0.014)	(0.019)	(0.016)	(0.026)
wave	0.172***	0.159***	0.169***	0.160***	0.165***	0.179***
	(0.024)	(0.025)	(0.008)	(0.015)	(0.008)	(0.015)
Constant	0.637***	0.522***	0.175	-0.744***		-1.046***
	(0.185)	(0.185)	(0.108)	(0.105)		(0.104)
θ			2.020			
Observations	34350	34350	34350	34350	21606	21606
No. id	17175	17175	17175	17175	10803	10803
Log-likelihood			-65319	-56963	-25551	-17382

Base categories: single, not employed, eduqual1, male, wave 1.
In P and NB fixed-effects only time varying regressors can be included.
Standard errors in parentheses, adjusted for clustering on 17175 id in the pooled model.
*** p<0.01, ** p<0.05, * p<0.1.

Table 3.8: Model Selection

	DOCT	GP	SPOUTER
Likelihood ratio test of $\theta = 0$ (P vs NB – random effects)			
θ	0.692	0.765	2.020
Chibar2(01)	1.0E+05	7.7E+04	7.5e+04
Pr$\geq$Chi2	0.000	0.000	0.000
Hausman test (P model) – fixed vs random effects			
Chi2(9)	2116.10	2077.41	296.73
Pr>Chi2	0.0000	0.0000	0.0000
Hausman test (NB model) – fixed vs random effects			
Chi2(9)	1618.90	1519.11	453.13
Pr>Chi2	0.0000	0.0000	0.0000

Table 3.9: ZIP_FE

	DOCT	GP	SPOUTER
COUNT			
logincome	0.023***	0.029***	-0.001
	(0.003)	(0.003)	(0.007)
emp	-0.162***	-0.128***	-0.067*
	(0.019)	(0.025)	(0.037)
retired	-0.054***	-0.028*	-0.061**
	(0.012)	(0.015)	(0.027)
mstat2	0.091***	0.128***	0.126**
	(0.025)	(0.030)	(0.058)
sphs	0.197***	0.205***	0.217***
	(0.012)	(0.015)	(0.028)
adl	0.185***	0.172***	0.213***
	(0.011)	(0.014)	(0.025)
mobilit	0.141***	0.117***	0.068***
	(0.009)	(0.012)	(0.021)
chronic	0.158***	0.111***	0.084***
	(0.009)	(0.011)	(0.021)
wave	0.060***	0.002	0.078***
	(0.005)	(0.006)	(0.010)
LOGIT			
logincome	0.013	-0.037	0.043**
	(0.026)	(0.023)	(0.018)
emp	0.058	0.028	-0.170*
	(0.142)	(0.123)	(0.100)
retired	0.127	0.036	-0.052
	(0.130)	(0.108)	(0.081)
mstat2	-0.666***	-0.488**	0.001
	(0.231)	(0.192)	(0.151)
sphs	0.343***	0.252***	0.297***
	(0.076)	(0.067)	(0.057)
adl	0.037	0.113	0.021
	(0.175)	(0.140)	(0.086)
mobilit	0.332***	0.272***	0.256***
	(0.082)	(0.070)	(0.053)
chronic	0.745***	0.514***	0.443***
	(0.095)	(0.075)	(0.053)
wave	0.149***	0.191***	0.253***
	(0.042)	(0.036)	(0.028)
Nonzero observations	27198	24090	9842
Log-likelihood (COUNT)	-36947	-26234	-10708
Zero observations	5220	6984	11764
Log-likelihood (LOGIT)	-1724	-2341	-3938
Pseudo R2	0.047	0.033	0.034

Base categories: single, not employed, wave 1.
Standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1.

Table 3.10: Log Likelihood and Information Criteria for Estimated Models

Variable	Model	N	Log(L)	K	AIC	BIC
DOCT						
	RE P	34350	-104563	15	209155	209282
	RE NB	34350	-94006	16	188045	188180
	FE P	32418	-44973	9	89964	90039
	FE NB	32418	-34949[a]	10	69917[b]	70001[c]
	ZIP_FE('count')	27198	-36947	9	73912	73986
GP						
	RE P	34350	-89218	15	178466	178593
	RE NB	34350	-83006	16	166043	166178
	FE P	31074	-35515	9	71048	71123
	FE NB	31074	-29554	10	59127	59211
	ZIP_FE('count')	24090	-26234[a]	9	52486[b]	52559[c]
SPOUTER						
	RE P	34350	-65318	15	130667	130794
	RE NB	34350	-56963	16	113957	114092
	FE P	21606	-25551	9	51120	51192
	FE NB	21606	-17382	10	34785	34864
	ZIP_FE('count')	9842	-10708[a]	9	21434[b]	21499[c]

Notes: RE, random effects; FE, fixed-effects; AIC, Akaike information criterion: $AIC = -2log(L) + 2K$; BIC, Schwarz information criterion: $BIC = -2log(L) + K\,log(N)$; where L is the maximized log likelihood of the model, K is the number of parameters; and N is the number of observations. [a] Model with the bigger log likelihood value; [b] Model preferred by AIC; [c] Model preferred by BIC.

Conclusions

Ensuring socio-economic equity and reactivity of health care systems is an explicit policy target in many OECD countries, where "equity" is defined as equal access to health care services for people with equal health problems. Policy makers should have insight in the inequality changing effects of various health care systems, as lack of access and quality may cause or at least reinforce the positive association between socio-economic status (SES) and health, the so-called SES gradient in health. There is ample evidence that mortality and morbidity, the relative incidence of a disease or condition that alters health and the quality of life, are inversely related to SES correlates such as income, education, or wealth (Deaton 2002). This requires policies which are aimed at containing the prevalence of chronic diseases associated with population aging, and at dedicating more resources to preventive measures.

In the first chapter we analyzed the relationship between income as a measure of SES and the use of several health care services for the 50+ population in the US and a number of European countries. Using a health production framework, we analyzed the potential income effects and how they vary with prices and other institutional features. We found clear evidence of a positive income gradient for several health care services, particularly for specialist visits, outpatient services, and dental care. These are also the services for which we found the clearest positive association between the income gradient and public expenditure on health care at the aggregate (country) level. Our results suggest that countries with higher public health expenditures do not automatically get closer to the policy goal of health care equity, i.e. equal access for those with the same needs. On the contrary, our results suggest that the extra services that the extra money can buy disproportionally benefit the richer part of the (older) population. Validating the theoretical predictions requires more detailed insight in the prices and characteristics of various types of health care services than is currently available. More work on, for example, what is covered by which insurance is needed.

In the second chapter we focused on preventive clinical service utilization by the population aged 50 and over in a number of European countries with varying health care systems. The analysis of education first, and then of three indicators of SES – education, income, and employment status – suggests that economic and social resources are associated with whether respondents use preventive services. With more education and skills, respondents are more likely to have better jobs with employer-sponsored health benefits. More informed respondents are also likely to better navigate their way through increasingly complex health care systems, with more educated respondents being usually more patient. Income is similarly related to the use of preventive services: with higher economic status comes greater purchasing power and greater likelihood of medical access. The effectiveness and efficiency of polices aimed at controlling the growing costs of health care also depend on how much individuals understand the health risk involved and decide to act accordingly. Better and more accessible information on the medical effectiveness of the intervention and the associated risk of death is important. Therefore it seems that informative materials focused on specific socio-demographic segments need to be developed to encourage screening.

Chapter 3 was dedicated to defining and estimating a zero-inflated Poisson model with fixed-effects to identify respondent- and health-related characteristics associated with health care demand using a two-wave panel. This is a new model that is proposed to model count measures of health care utilization and account for the panel structure of the data. Our findings suggest that controlling for the portion of respondents that are certain zeros in one of the two years of the two waves does make a difference for counts with a larger number of zeros, where traditional count data models are not able to disentangle the effects. All in all we find a strong income-health care visit gradient for the "non certain zeros" group for doctor and GP visits, while the income effect is absent in the "certain zeros" group. In specialist, outpatient and emergency room consultations we find the opposite: the income-health care visit gradient is in the "certain zeros" group. The model that has been developed in this chapter has some attractive features. It makes it possible to account for individual effect in panel data: fixed-effects can explain overdispersion, where Poisson model can not. It allows the correction for extra zeros defining two latent classes of low users in the probability of visiting a doctor, and high users in the conditional positive number of visits. Extending the estimator and the estimation algorithm to the case of more than two time periods and developing model selection tests need to be explored in future research.

Appendix A

Data Sources in Chapter 1

The data used in Chapter 1 are drawn from the following surveys:

- Survey of Health, Ageing and Retirement in Europe (SHARE), wave 1 – release 2. The first wave covers eleven European countries (Austria, Germany, Belgium, Denmark, Spain, France, Greece, Italy, the Netherlands, Sweden and Switzerland); it is the first European survey that provides cross-nationally comparable micro-data on the economic, social and health situation of the 50+ population. It was launched in 2004 at present it has two waves available. However the variable income is not directly comparable across the two waves since it is gross in the first wave and net in the second wave. We selected the first wave to make it directly comparable with the Health and Retirement Survey (HRS);

- Health and Retirement Survey (HRS) which surveys Americans over the age of 50 every two years. The study paints an emerging portrait of an aging America's physical and mental health, insurance coverage, financial status, family support systems, labor market status, and retirement planning.

SHARE has been designed after HRS, which facilitates the use of these surveys for comparative purposes. Preliminary results from SHARE (Börsch-Supan et al. 2005, pp. 89–94) clearly indicate that health status is positively associated with SES. Lower socio-economic groups experience poorer health status and have higher health care needs: people with lower socio-economic background and with poor health used comparatively more family physician and hospital services. In contrast, specialist services were comparatively less used by people with lower socio-economic background and with poor health.

Appendix B

Stata Syntax for ZIP_FE Model

The syntax below shows how to estimate a ZIP fixed-effects model (ZIP_FE) via conditional maximum likelihood with Stata. You need to know how to use the optimization tool in Stata, see Gould et al. (2006).

```
set more off
capture program drop ZIP_FE_model

program define ZIP_FE_model
version 9.1

args todo b lnf
tempvar theta1 lambda last nonz w sln0 sln r0 r nb0 nb1 nb00 nb2 L2

local by "$MY_panel"
local byby "by `by'"
sort `by' wave
local y "$ML_y1"

mleval `theta1' = `b'

quietly {
gen double `lambda' = exp(`theta1')

`byby': gen  double `last'= (_n==_N)
`byby': egen double `nonz'= min(`y')
`byby': egen double `w'   = sum(`y')
`byby': gen  double `sln0'= lngamma(`y'+1)
`byby': egen double `sln' = sum(`sln0')
`byby': gen  double `r0'  = `y'*ln(`lambda')
`byby': egen double `r'   = sum(`r0')
`byby': egen double `nb0' = sum(`lambda')
`byby': gen  double `nb1' = `nb0'^`w'
`byby': gen  double `nb00'= `lambda'^`w'
```

```
‘byby’: egen double ‘nb2’ = sum(‘nb00’)

‘byby’: gen  double ‘L2’  = lngamma(‘w’+1) - ‘sln’ + ‘r’ - ln( ‘nb1’ - ‘nb2’ ) /*
*/ if (‘last’ == 1 & ‘nonz’>0)
mlsum ‘lnf’ = ‘L2’ if (‘last’ == 1 & ‘nonz’>0)
}
end

sort id wave
global MY_panel id

ml model d0 ZIP_FE_model (y = x1 x2, nocons) if nonz>0
ml check
ml search
ml maximize, difficult
```

Bibliography

Adams P., Hurd M.D., McFadden D., Merrill A., and Ribeiro T. (2003), "Healthy, wealthy and wise? Tests for direct causal paths between health and socioeconomic status", *Journal of Econometrics*, 112, 3–56.

Allison P.D., and Waterman R.P. (2002), "Fixed-effects negative binomial regression models", *Sociological Methodology*, 32, 247–265.

Atella V., Brindisi F., Deb P., and Rosati F.C. (2004), "Determinants of access to physician services in Italy: a latent class seemingly unrelated probit approach", *Health Economics*, 13(7), 657–668.

Avitabile C., Jappelli T., and Padula M. (2008), "Screening tests, information, and the health-education gradient", *CSEF Working Paper*, 187.

Bago d'Uva T. (2006), "Latent class models for utilisation of health care", *Health Economics*, 15, 329–343.

Bago d'Uva T., Jones A.M., and van Doorslaer E. (2009), "Measurements of horizontal inequity in health care utilisation using European panel data", *Journal of Health Economics*, 28, 280–289.

Banks J., Marmot M., Oldfield Z., and Smith J.P. (2007), "The SES health gradient on both sides of the Atlantic", *IZA Discussion Paper*, 2539.

Böhning D. (1998), "Zero-inflated Poisson models and C.A.MAN: a tutorial collection of evidence", *Biometrical Journal*, 40, 833–843.

Börsch-Supan A., Brugiavini A., Jürges H., Mackenbach J., Siegrist J., and Weber G. (2005), "Health, Ageing and Retirement in Europe – First results from the Survey of Health, Ageing and Retirement in Europe", Mannheim Research Institute for Economics of Ageing (MEA), Mannheim.

Börsch-Supan A., et al. (2008), "Health, Ageing and Retirement in Europe (2004–2007) – Starting the Longitudinal Dimension", Mannheim Research Institute for the Economics of Ageing (MEA), Mannheim.

Börsch-Supan A., Jürges H. (Eds.) (2005), "The Survey of Health, Ageing and Retirement in Europe – Methodology", Mannheim Research Institute for the Economics of Ageing (MEA), Mannheim.

Bullough B. (1972), “Poverty, ethnic identity and preventive health care”, *Journal of Health and Social Behavior*, 13.

Busse R., Schreyögg J., and Smith P.C. (2008), “Variability in healthcare treatment costs amongst nine EU countries – Results from the HealthBASKET project”, *Health Economics*, 17(S1), S1–S8.

Cameron C., and Trivedi P.K. (1998), “Regression analysis of count data”, *Econometric Society Monographs*, Cambridge University Press, Cambridge, UK.

Cameron C., and Trivedi P.K. (2005), “Microeconometrics: Methods and applications”, *Econometric Society Monographs*, Cambridge University Press, New York, NY.

Carman K.G., and Kooreman P. (2010), “Flu shots, mammograms, and the perception of probabilities”, *Netspar Discussion Paper*, No. 03/2010–014.

Chamberlain G. (1980), “Analysis of covariance with qualitative data”, *Review of Economic Studies*, 47, 225–238.

Chamberlain G. (1984), “Panel data”, Handbook of Econometrics, Griliches A. and Intriligator M.D. (Eds.), North-Holland, Amsterdam, 2, 1247–1318.

Chao A., Connell C.J., Cokkinides V., Jacobs E.J., Calle E.E., and Thun M. (2004), “Underuse of screening sigmoidoscopy and colonoscopy in a large cohort of US adults”, *American Journal of Public Health*, 94(10).

Cherrington A., Corbie-Smith G., and Pathman D.E. (2007), “Preventive medicine”, *Elsevier*, 45, 282–289.

Coffield A.B., Maciosek M.V., McGinnis J.M., Harris J.R., Caldwell M.B., Teutsch S.M., Atkins D., Richland J.H., and Haddix A. (2001), “Priorities among recommended clinical preventive services”, *American Journal of Preventive Medicine*, 21, 1–9.

Courbage C., and Coulon A. de (2004), “Prevention and private health insurance in the U.K.”, *The Geneva paper series on risk and insurance*, 29(4), 719–727.

Crepon B., and Duguet E. (1997), “Research and development, competition and innovation. Pseudo maximum likelihood and simulated maximum likelihood methods applied to count data models with heterogeneity”, *Journal of Econometrics*, 79, 355–378.

Culyer A.J. (2006), “The bogus conflict between efficiency and vertical equity”, *Health Economics*, 15(11), 1155–1158.

Dave D., and Kaestner R. (2006), “Health insurance and the ex-ante moral hazard: evidence from Medicare”, *NBER Working Paper*, 12764.

Deaton A. (2002), “Policy implications of the gradient of health and wealth”, *Health Affairs*, Millwood, 21(2), 13–30.

Deb P., and Holmes A.M. (2000), “Estimates of use and costs of behavioural health care: A comparison of standard and finite mixture models”, *Health Economics*, 9, 475–489.

Deb P., and Trivedi P.K. (1997), "Demand for medical care by the elderly: A finite mixture approach", *Journal of Applied Econometrics*, 12, 313–336.

Deb P., and Trivedi P.K. (2002), "The structure for demand for health care: Latent class versus two-part models", *Journal of Health Economics*, 21, 601–625.

Docteur E., and Oxley H. (2003), "Health-care systems: lessons from the reform experience", *OECD Health Working Paper Series*, 9, OECD, Paris.

European Commission (2001), "The future of health care and care for the elderly: guaranteeing accessibility, quality and financial viability", Commission of the European Communities, Brussels.

European Parliament (2006), "European guidelines for quality assurance in breast cancer screening and diagnosis", P6_TA(2006)0449.

Fabbri D., and Monfardini C. (2002), "Public vs. private health care services demand in Italy", *Giornale degli Economisti e Annali di Economia*, 62(1), 93–123.

Fiore A.E., Wortley P.M., and Bridges C.B. (2007), "Missed opportunities for the prevention of influenza", *Preventive Medicine*, 45(1), 88–89.

Geiger T.M., Miedema B.W., Geana M.V., Thaler K., Rangnekar N.J., and Cameron G.T. (2008), "Improving rates for screening colonoscopy: Analysis of the health information national trends survey (HINTS I) data", *Surgical Endoscopy*, 22(2).

Gøtzsche P.C., and Nielsen M. (2006), "Screening from breast cancer with mammography", *Cochrane database of Systematic Reviews*, 4.

Gould W., Pitblado J., and Sribney W. (2006), "Maximum likelihood estimation with Stata", College Station, TX: Stata Press.

Greene W. (2008), "Functional forms for the negative binomial model for count data", *Economic Letters*, 99, 585–590.

Grossman M. (1972), "On the concept of health capital and the demand for health", *Journal of Political Economy*, 80(2), 223–255.

Grossman M. (2000), "The human capital model", Handbook of Health Economics, Culyer A.J. and Newhouse J.P. (Eds.), Elsevier, Amsterdam, 1A, 347–408.

Grossman M., and Kaestner R. (1997), "Effects of education on health", The Social Benefits of Education, Behrman J.R. and Stacey N. (Eds.), Ann. Arbor, University of Michigan Press.

Grossman M., and Rand E. (1974), "Consumer incentive for health services in chronic illnesses", Consumer Incentives for Healthcare, Mushkin S.J. (Eds.), Milbank Memorial Fund, New York, 114–151.

Guimaraes P. (2008), "The fixed-effects negative binomial model revisited", *Economic Letters*, 99, 63–66.

Hausman J.A. (1978), “Specification tests in econometrics”, *Econometrica*, 46(6), 1251–1271.

Hausman J.A., Hall B.H., and Griliches Z. (1984), “Econometric models for count data with an application to the patents-R&D relationship”, *Econometrica*, 52(4), 909–938.

Holland W.W., Stewart S., and Masseria C. (2006), “Policy brief: Screening in Europe”, *Policy Brief*, European Observatory on Health Systems and Policies, Copenhagen.

Hunt S.M., McKenna S.P., McEwen J., Williams J., and Papp E. (1981), “The Nottingham health profile: Subjective health status and medical consultations”, *Social Science & Medicine*, 15(3), Part 1, 221–229.

Hurd M.D., and Kapteyn A. (2003), “Health, wealth and the role of institutions”, *Journal of Human Resources*, 38(2), 386–415.

Jimenez-Martin S., Labeaga J.M., and Martinez-Granado M. (2002), “Latent class versus two-part models in the demand for physicians services across the European Union”, *Health Economics*, 11, 301–321.

Jimenez-Martin S., Labeaga J.M., and Martnez-Granado M. (2004), “An empirical analysis of the demand for physician services across the European Union”, *European Journal of Health Economics*, 5(2), 150–165.

Johnson N., Kotz S., and Kemp A.W. (1992), “Univariate discrete distributions”, Wiley, 2nd edition, New York.

Jonas D.E., Russel L.B., Chou J., and Pignone M. (2009), “Willingness-to-pay to avoid the time spent and discomfort associated with screening colonoscopy”, *Health Economics*, 19(10), 1193–1211.

Jones A.M., Koolman X., and van Doorslaer E. (2006), “The impact of supplementary private health insurance on the use of specialists in selected European countries”, *Annales d'Economie et Statistique*, 83/84, 251–275.

Katz S.J., and Hofer T.P. (1994), “Socioeconomic disparities in preventive care persist despite universal coverage”, *JAMA*, 272(7).

Kenkel D.S. (1991), “Health behavior, health knowledge, and schooling”, *Journal of Political Economy*, 99(2), 287–305.

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

Kenkel D.S. (2000), “Prevention”, Handbook of Health Economics, Culyer A.J. and Newhouse J.P. (Eds.), Elsevier, New York, 1676–1720.

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

Lee M., and Kobayashi S. (2001), “Proportional treatment effects for count response panel data: Effects of binary exercise on health care demand”, *Health Economics*, 10, 411–428.

Lee A.H., Stevenson M.R., Wang K., and Yau K.K.W. (2002), “Modeling young driver motor vehicle crashes: data with extra zeros”, *Accid. Anal. Prev.*, 34, 515–521.

Long J.S. (1997), “Regression models for categorical and limited dependent variables”, *Advanced Quantitative Techniques in the Social Sciences*, Sage Publications.

Manning W.G., Morris C.N., Newhouse J.P., et al. (1981), “A two-part model of the demand for medical care: preliminary results from the Health Insurance Study”, Health, Economics, and Health Economics, Sheffler R.M. and Rossiter L.F. (Eds.), Amsterdam, North-Holland, 103–123.

Marmot M. (1999), “Acting on the evidence to reduce inequalities”, *Health Affairs*, Millwood, 18(3), 42–44.

Maurer J. (2007), “Modelling socioeconomic and health determinants of health care use: A semiparametric approach”, *Health Economics*, 16, 967–979.

Meer J., Miller D.L., and Rosen H.S. (2003), “Exploring the health-wealth nexus”, *Journal of Health Economics*, 22, 713–730.

Michaud P.C., and van Soest A. (2008), “Health and wealth of elderly couples: causality tests using dynamic panel data models”, *Journal of Health Economics*, 27, 1312–1325.

Mullahy J. (1986), “Specification and testing of some modified count data models”, *Journal of Econometrics*, 33, 341–365.

Mullahy J. (1999), “It’ll only hurt a second? Microeconomic determinants of who gets flu shots”, *Health Economics*, John Wiley & Sons (Ltd.), 8(1), 9–24.

OECD (2004), “OECD Health Data 2004: A comparative analysis of 30 countries”, Organisation for Economic Co-operation and Development (OECD), Paris.

OECD (2007), “OECD Health Data 2007”, Organisation for Economic Co-operation and Development (OECD), Paris.

OECD (2007), “Health at a glance”, Organisation for Economic Co-operation and Development (OECD), Paris.

Oliver A., and Mossialos E. (2008), “Equity of access to health care: outlining the foundations for action”, *Journal of Epidemiology and Community Health*, 58, 655–658.

Parente S.T., Salkever D.S., and DaVanzo J. (2005), “The role of consumer knowledge of insurance benefits in the demand for preventive health care among the elderly”, *Health Economics*, 14(1), 25–38.

Picone G., Sloan F., and Taylor D.Jr. (2004), “Effects of risk and time preference and expected longevity on demand for medical tests”, *Journal of Risk and Uncertainty*, 28(1).

Pohlmeier W., and Ulrich V. (1995), “An econometric model of the two-part decision making process in the demand for health care”, *Journal of Human Resources*, 30, 339–361.

Rabe-Hesketh S., and Skrondal A. (2008), "Multilevel and longitudinal modeling using Stata", College Station, TX: Stata Press.

Sen A. (2002), "Why health equity?", *Health Economics*, John Wiley & Sons Ltd., 11(8), 659–666.

Smith J.P. (1999), "Healthy bodies and thick wallets: The dual relationship between health and economic status", *The Journal of Economic Perspectives*, 13(2), 145–166.

Stargardt T. (2008), "Health service costs in Europe: Cost and reimbursement of primary hip replacement in nine countries", *Health Economics*, 17, S9–S20.

Syme S.L., and Balfour J.L. (1998), "Social determinants of disease", Maxcy-Rosenau-Last Public Health & Preventive Medicine, Wallace R.B. (Eds.), Appleton & Lange 14th edition, Stamford, 795–810.

Tan S.S., Redekop W.K., and Rutten F.H. (2008), "Costs and prices of single dental fillings in Europe: a micro-costing study", *Health Economics*, 17, S83–S93.

U.S. Census Bureau (2005), "Income, poverty, and health insurance coverage in the United States: 2005", *Current Population Reports*, Washington, DC.

U.S. Department of Health and Human Services (2000), "Healthy people 2010: National health promotion and disease prevention objectives", U.S. Department of Health and Human Services, Public Health Service, Washington, DC, 91–50212.

Van Doorslaer E., Masseria C., and Koolman X. (2006), "Inequalities to access in medical care by income in developed countries", *Canadian Medical Association Journal*, 174(2), 177–183.

Van Doorslaer E., Wagstaff A., van der Burg H., Christiansen T., De Graeve D., Duchesne I., Gerdtham U., Gerfin M., Geurts J., Gross L., Häkkinen U., John J., Klavus J., Leu R.E., Nolan B., O'Donnell O., Propper C., Puffer F., Schellhorn M., Sundberg G., and Winkelhake O. (2000), "Equity in the delivery of health care in Europe and the US", *Journal of Health Economics*, 19, 553–583.

Vuong Q.H. (1989), "Likelihood ratio tests for model selection and nonnested hypotheses", *Econometrica*, 57(2), 307–333.

Wang K., Yau K.W.K., and Lee A.H. (2002), "A zero-inflated Poisson mixed model to analyze diagnosis related groups with majority of same-day hospital stays", *Computer Methods and Programs in Biomedicine*, 68, 195–203.

Wang P.D., and Lin R.S. (1996), "Sociodemographic factors of Pap smear screening in Taiwan", *Public Health*, 110, 123–127.

Wee C.C., McCarthy E.P., Davis R.B., and Phillips R.S. (2000), "Screening for cervical and breast cancer: in obesity an unrecognized barrier to preventive care?", *Annals of Internal Medicine*, 132, 697–704.

Winkelmann R. (2003), *Econometric analysis of count data*, 4th edition, Springer-Verlag, Berlin.

Winkelmann R. (2004), "Health care reform and the number of doctor visits – An econometric analysis", *Journal of Applied Econometrics*, 19, 455–472.

Witt J. (2008), "The effect of information in the utilization of preventive health-care strategies: An application to breast cancer", *Health Economics*, 17, 721–731.

Wooldridge J. (2002), *Econometric analysis of cross-section and panel data*, MIT-Press, Cambridge, MA.

WHO (2002), "Breast cancer screening", IARC Handbooks of Cancer Prevention, 7, IARC Press, Paris.

WHO (2004), "World report on knowledge for better health: strengthening health systems", World Health Organization (WHO), Geneva.

WHO (2005), "Preventing chronic diseases: A vital investment", World Health Organization (WHO), Geneva.

Sommario (Abstracts in Italian)

Capitolo 1: Questo capitolo analizza l'incidenza del reddito sull'utilizzo delle prestazioni mediche in adulti oltre i 50 anni di età negli Stati Uniti ed in alcuni Paesi europei caratterizzati da differenze istituzionali nei sistemi sanitari. Lo studio si basa su dati individuali ottenuti da diverse indagini che forniscono informazioni sulla situazione socio economica, sullo stato di salute e sull'utilizzo delle prestazioni medico-sanitarie di adulti oltre i 50 anni di età. Con il modello probit vengono stimate le probabilità che un individuo riceva specifiche prestazioni mediche (visite dal medico di base, da un medico specialista, dal dentista, ricovero in ospedale con degenza, operazioni chirurgiche senza degenza). Oltre ad includere variabili quali reddito e caratteristiche demografiche, si considera anche lo stato di salute dell'individuo (utilizzando misure dello stato di salute sia soggettive che oggettive). Si analizza come vari la relazione tra reddito e utilizzo delle prestazioni sanitarie tra paese e paese. Successivamente tali differenze tra i paesi vengono messe in relazione con le caratteristiche istituzionali dei diversi paesi, come ad esempio la spesa sanitaria totale e pubblica pro capite, se valga o meno il sistema dell'impegnativa da parte del medico di base per ottenere visite specialistiche, e copayment per le prestazioni mediche.

Capitolo 2: In questo capitolo viene analizzata la relazione tra l'utilizzo di esami di prevenzione e lo status socio economico di individui oltre i 50 anni di età in vari Paesi europei che presentano sistemi sanitari diversi. Lo studio utilizza dati della indagine su Salute, Invecchiamento e Pensioni in Europa (in inglese "Survey of Health, Ageing and Retirement in Europe", SHARE), una banca dati multidisciplinare e multipaese che fornisce dati individuali su status socio-economico, salute e utilizzo di esami di prevenzione. Con il modello probit vengono ottenute le probabilità che un individuo riceva specifici esami di prevenzione (vaccinazione contro l'influenza, analisi del sangue, colonscopia, ricerca del sangue occulto nelle feci, esame dell'occhio e, per le donne, mammografia). Oltre a controllare per il livello d'istruzione e caratteristiche demografiche, si considerano anche ulteriori variabili economiche e lo stato di salute dell'individuo (utilizzando misure dello stato di salute sia soggettive che oggettive). L'analisi empirica mostra come, iniziando dal livello di educazione e successivamente analizzando anche gli altri due indicatori di status socio-economico – reddito e status occupazionale –, sia le risorse economiche che sociali incidano sull'utilizzo di esami di prevenzione. In particolare, il livello di educazione emerge come un fattore determinante nell'utilizzo di tali esami.

Capitolo 3: Il campo delle analisi quantitative che riguarda i sistemi sanitari è spesso caratterizzato da dati di conteggio con eccesso di zeri. In questo capitolo viene definito un modello zero-inflazionato di Poisson ad effetti fissi per identificare le caratteristiche individuali e relative allo stato di salute, associate alla domanda per prestazioni mediche. Questo nuovo modello viene utilizzato per analizzare l'utilizzo delle prestazioni mediche, definite come variabili di conteggio, tenendo conto anche della dimensione longitudinale dei dati. La stima dei parametri è ottenuta via massima verosimiglianza condizionata. Viene presentata una applicazione ai dati individuali della indagine su Salute, Invecchiamento e Pensioni in Europa (in inglese "Survey of Health, Ageing and Retirement in Europe", SHARE) relativi al biennio 2004–2006. Le stime così ottenute vengono paragonate ai risultati stimati con i modelli standard per dati panel per la modellazione di dati di conteggio. I risultati mostrano che quando si distingue tra outcome che presenti "zero" o "non zero" in uno dei due anni, esiste una differenza per quelle variabili di conteggio che presentano numerosi zero nella distribuzione.