Modelling Unobserved Heterogeneity in Health and Health Care: an Extended Latent Class Approach

Paolo Li Donni

Ph.D.

University of York

Economics

July 2010

Abstract

Unobserved heterogeneity is one of the main concerns for applied economists, this is particularly so when modelling health and health related behaviours. This thesis illustrates four studies on modelling unobserved heterogeneity using some recent developments in latent class analysis. Chapter 2 examines two sources of individual unobserved heterogeneity when subjective indicators are used to measure health status: variations in unobservable true health and differences in self-reporting behaviour for a given level of "true health". These two sources are separately identified using both objective (biomarkers) and subjective health indicators.

Chapter 3 examines the so called positive correlation test. This test rejects the null of absence of private information in a given insurance market when individuals with greater coverage experience more of the insured risk. This test is shown to lead to puzzling results where there exists multiple sources of private information (multidimensional heterogeneity). An alternative strategy proposed uses a finite number of heterogeneous types and extends the standard adverse and favourable selection definitions into *local* and *global* ones. We implement a finite mixture model to identify the unobserved types and test the multidimensionality of private information. We apply these approaches to the US long-term care and Medigap insurance markets.

Chapter 4 further considers the issue of the asymmetric information in the insurance markets, by investigating how to disentangle the incentive effect,

due to the structure of insurance contracts, from the selection effect. Alternative econometric strategies are evaluated to empirically disentangle these two effects when multiple dimensions of unobserved heterogeneity affect the reliability of the standard positive correlation. The proposed application focuses on the effect of Medigap insurance on having an inpatient hospital stay and compares the probit and the recursive bivariate probit model with a discrete multiresponse finite mixture model.

Finally Chapter 5 examines the relationship between unobserved risk preferences and four insurance purchase decisions in the US: Medigap, long-term insurance, life insurance and annuity. Standard economic theory assumes that individuals take decisions over a set of risky domains according to their own risk preferences which are stable across decision contexts. This assumption of context-invariant risk preference has caused debate in the literature concerning its validity. This chapter proposes an empirical strategy to test whether risk preferences are multidimensional and whether they differ across insurance choices. This empirical appraisal provides a simple way to model non-preference factors - such as context specificity - which can play an important role in determining multiple demand for insurance conditional on individual general risk preferences. Our findings are largely consistent with the hypothesis of domain-general components of risk preferences, although context specificity is important especially between choices over insurance contracts that are “closer” in context (e.g. long-term care and Medigap insurance).

Contents

Acknowledgments

I express my gratitude to my supervisors Andrew Jones and Nigel Rice for their valuable guidance and advice during my training as PhD student. I would also like to thank all the people at the Centre for Health Economics for the friendly environment, help and support during these years.

My greatest thank goes to my wife and family for their important support and sustaining me along the way.

Declarations

The present work is the result of research I have conducted as a PhD candidate in the Department of Economics and Related Studies at the University of York. My research was partially funded by the Economic & Social Research Council.

Chapters 2, 3 and 4 are written in co-authorship with Valentino Dardanoni. My major contributions are: data assembly, construction of the variables and estimation of empirical models. I have contributed, together with my co-author, to the concepts and definition of the methods of analysis and model specification. In addition I also wrote a substantial part of the papers following my co-author's useful advice and comments during the development of the work. Chapters 2, 3 and 4 have been submitted for publication in Health Economics, Quantitative Economics and the Journal of Health Economics and are currently under review.

Chapter 5 is sole-authored.

Chapter 1

Introduction

Many concepts that are of interest to applied social scientists cannot be observed directly. Examples of these constructs involve individual preferences, health status, attitudes towards risky behaviours or utilization of health care resources. Since these dimensions are often not directly observed or only partially observed, they introduce important sources of unobserved heterogeneity in empirical analysis of many issues relevant in economics (Heckman [69]).

Unobserved heterogeneity is one of the main concerns for applied economists and particularly when health related issues are investigated since health status is intrinsically multidimensional and unobservable. There are several different strategies that applied economists exploit to overcome this issue and to model unobserved heterogeneity. The adopted solution is usually related specifically to the research context and depends more generally on the availability of data (experimental or observational data) or on the econometric strategy employed to deal with unobservables.

In the policy evaluation setting, for example, researchers focus on the relationship between the outcome and a variable describing a policy or a programme. Causal interpretation of this relationship can be achieved as long as unobserved factors do not affect the programme participation decision, otherwise model parameters are not identifiable. The endogeneity of the policy variable has motivated the use of social experiments as offering a potential solution to identify causal effects. Experimental data are well suited to control for unobserved factors since they make it possible to vary the particular variable of interest, holding other covariates at controlled levels. This is possible because the experimental environment data generating process can be directly controlled. In contrast observational data, which usually come from population surveys, are generated in an uncontrolled environment, leaving open the possibility that unobserved factors affect the relationships of interest. Since experimental data are expensive to produce (on a large scale) and their implementation may pose some ethical concerns, most applied economic studies use survey data (sometimes with some specific characteristics) and develop econometric techniques to deal with unobserved factors. A common strategy is to use proxy or observable indicators derived from questionnaire items designed to elicit responses related to these unobservable dimensions. An important technique is based on latent class modelling.

A latent variable model is generally constructed using a set of observable individual characteristics, often termed "indicators" or "responses", that are considered a manifestation of the underlying latent construct of interest. This type of model is often non-linear and, in addition to the manifest variables, the model includes one or more unobserved or latent variables representing the constructs of interest. Two main assumptions define the causal mechanisms underlying the responses. First, it is assumed that the responses on the indicators are the result of an individual's position on the latent variable(s). The second assumption, known as the local independence axiom, states that manifest variables have nothing in common after controlling for the latent variable(s). Depending on the distribution of the latent and manifest variables one obtains different types of latent variable models. When both indicators and latent variables are categorical, and therefore assumed to come from a multinomial distribution, the latent variable model takes the specific name of latent class analysis (Bartholomew and Knott [11]).

This thesis uses survey data and some recent developments in latent class analysis to model individual unobserved heterogeneity in three different contexts: health status, health care utilization and insurance coverage, and multiple demand for insurance and individual risk preference stability. The first study uses data from the Health Survey for England to disentangle the effect of individual characteristics on health production on the one hand, and its self-reporting effect on the other. Our focus is to study unobserved heterogeneity in the self-assessment of health after conditioning on unobserved heterogeneity in health production. Our econometric approach relies on latent class analysis to identify unobserved health and employs a wide set of health indicators including both objective and self-reported measures. In addition we use some recent developments in latent class analysis which allow us to model explicitly the residual association between subjective indicators and self-assessed health by setting an explicit recursive structure. This is the first contribution of the paper, which makes our approach substantially different from those relying on the traditional MIMIC model. The second contribution of this study is related to the indicators used as manifestations of individual health status. In fact "true" health is constructed using not only self-assessed health and other self-reported health measures, but also objective indicators observed through biomarkers. Our results show the existence of two unobserved classes of individuals representing those in good health and those in ill-health. Moreover, our findings provide further evidence of the existence of unobserved heterogeneity in self reporting behaviour since individual characteristics, such as socio-economic status and education, cannot be ignored as predictors of self-assessed health even after conditioning on unobserved individual 'true' health-types.

The third and the fourth chapters are intrinsically related since they both focus on the role of unobserved heterogeneity in the demand for health insurance and health care utilization. In particular the third chapter examines the standard prediction of the well-known Rothschild-Stiglitz adverse selection model. This model predicts that when individuals have private information about their actual risk, the insurance contract will be adversely selected, with high risk individuals choosing higher insurance coverage. Ex post, this will cause a positive correlation between risk and coverage. This observation has been empirically implemented using the so called positive correlation (PC) test which rejects the null of absence of asymmetric information in a given insurance market when, conditional on consumers' characteristics used by companies to price contracts, individuals with more coverage experience more of the insured risk. Contrary to the theoretical prediction of positive rick-coverage correlation, a number of empirical studies of insurance markets have found a negative risk-coverage correlation, a phenomenon which has been named favourable selection. This has been justified by claiming that there are two conflicting sources of private information, namely individual's actual risk and risk attitudes, so that two types of people buy insurance: high risk individuals and high risk averse individuals. Ex post, the former are higher risk than predicted and the latter are lower risk; in aggregate, those who buy more insurance do not have higher claims. The main contribution of the chapter is to evaluate how the standard PC test performs in the presence of multidimensional private information. Our claim is that under multidimensional private information there is the possibility that the insurance contract is *both* adversely *and* favourably selected by different individuals. Since the PC test relies on a single statistic to appraise the risk-coverage correlation, multidimensional private information may cause serious problems in detecting selection effects using observable data, since the PC test averages out selection effects which may pull in different directions.

We provide a substantially new methodology, relying on latent class analysis, to investigate the existence of private information which has three appealing characteristics: 1) it is possible to identify unobserved types by cross-classifying the relevant private information variables, 2) it extends the standard adverse and favourable selection definitions into *local* and *global* ones, 3) it is possible to test directly for the absence of selection effects into insurance contracts and for the multidimensionality of private information by imposing restrictions on the behaviour of these unobservable types. To show the validity of our approach we study separately the Medigap and the long-term insurance markets in the US. These two markets differ not just by the type of insurance, but also for in the level of regulation imposed in the Medigap market. Results show that the PC test shows an absence of significant residual risk-coverage correlation in both samples, while our finite mixture model reveals the existence of significant residual heterogeneity with large selection effects, which is stronger for the regulated Medigap market.

The fourth chapter is also focused on asymmetric information, but provides substantially different information compared to the previous paper. In particular it studies how Medigap affects the utilization of health care services and then focuses on disentangling the incentive effect, induced by insurance contracts on health care utilization, from selection effects due to the unobserved private information in insurance purchase decisions. Our main contribution is to provide evidence on how the standard econometric strategies developed to deal with this issue perform compared with a discrete multiresponse finite mixture model which controls for selection. In particular we compare our results to those obtained by the probit and the bivariate probit model, and find the residual effect of insurance on health care is small after controlling for individual unobserved types.

The final chapter studies whether individual risk preferences are stable across multiple insurance choices. In general there are two different points of view regarding risk preference generality. Classical economic theory assumes that individuals have the same attitude to bear risk in different contexts, while a recent and important literature, mostly related to behavioral economics, finds that context is the main factor and poses serious concerns for the internal validity of the risk-preference invariance principle. This issue has motivated a growing body of research. A subset of these studies focus on this principle by looking at multiple demand for insurance. In particular they study whether individuals who bear risk in one insurance domain are also willing to bear the same risk in another insurance domain. Clearly if risk preferences are general, the risk preferences pattern across insurance domains should be stable. To test this principle they consider the residual correlation between insurance choices after conditioning on predicted and realized risk. We propose an alternative framework to examine this issue using latent class analysis to identify unobserved types which differ in their level of risk aversion. Our approach has two appealing characteristics: 1) it allows us to disentangle the effect of risk preference on insurance choice from the residual correlation introduced by non-preference factors; this can be done by modelling correlations between insurance choices conditioned on unobserved risk preferences; 2) if risk preferences are stable across the relevant choices then there exists a unique latent variable which affects each of these choices; this can be tested directly by imposing restrictions on the relevant insurance behaviours. To show the applicability of our methodology, we use data from the Health and Retirement Study on four insurance purchase decisions: life insurance, Medicare supplemental insurance (Medigap), log-term care insurance and annuity. Our results show the existence of a stable pattern of individual risk preferences over different insurance domains, which supports the idea of a domain-general component of preferences. In addition we also provide further evidence that context plays an important role in determining insurance choices particularly when insurance coverage decisions are "closer" in coverage type.

Chapter 2

Reporting Heterogeneity in Subjective Health Measures: an Extended Latent Class Approach

2.1 Introduction

Self-assessed health status (SAH) is a widely employed subjective health indicator in empirical research. It is based on the simple question "How is your health in general?" with a response framed in ordered categories ranging from "very good" or "excellent" to "poor" or "very poor". It is often assumed that these responses are generated by a corresponding continuous latent variable representing self-perceived health. Several studies found that SAH is a good predictor of mortality, morbidity and subsequent use of health care (Idler and Benyamini [77]). Furthermore Gerdtham *et al.* [62] showed that a continuous health measure obtained from the ordinal responses of SAH is highly correlated with other individual health measures.

As subjective indicator SAH has caused some concern among researchers related to the idea that individuals may link differently the same level of true health with the SAH's categories. The existence of these differences in self-reporting behavior is convincingly supported by empirical findings (Crossley and Kennedy [31], Groot [65]). For example, Crossley and Kennedy [31] exploit a particular feature of the Australian National Survey in which SAH question was asked to respondents and again to a random subsample. Results show that the distribution of SAH for this subgroup of respondents changes significantly between the two questions and that this variation depends on age, income and occupation. Bago d'Uva *et al.* [44] find that reporting heterogeneity may depend on the individuals' concept of what health means, on their expectations of their own health, their use of health care, and on their comprehension of the the health questions asked in the survey. Etilé and Milcent [50] find evidence that reporting heterogeneity is associated with socioeconomics status, while Lindeboom and van Doorslaer [90] find that age and gender, but not income or education, affect reporting behaviour. Johnston *et al.* [79] studied reporting heterogeneity in hypertension and found that the probability of false negative reporting is significantly income graded and then self-reported health measures might underestimate true income-related inequalities in health.

This source of measurement error in the mapping of true health into SAH that we call - following Shmueli [107] - unobserved reporting heterogeneity, has also been termed 'state-dependent reporting bias' (Kerkhofs and Lindeboom [85]), 'scale of reference bias' (Groot [65]), 'response category cut-point shift' (Sadana *et al.* [105], 2000; Murray *et al.* [96]).

To account for self-reporting behavior a possible approach is to use an ordered probit with cut point shift. This model allows the cut-points defining the mapping of latent health into the SAH's categories to depend on observable variables (Terza [110]). Although this approach allows for both index and cut off points shifts, it requires strong a priori restrictions on parameters to solve identifiability problems especially when the set of covariates on latent health and the cut-off point overlap (Lindeboom and van Doorslaer [90]).

There are many other papers that have analysed SAH. Van Doorslaer and Jones [112] use the McMaster 'Health Utility Index Mark' (HUI) to scale the intervals of SAH. They assume a stable mapping of HUI on the latent health determining SAH. Therefore the position of an individual ranked according to HUI should correspond to her rank according to SAH. They exploit this relationship between HUI and SAH to estimate an interval regression model where the upper (lower) bound of these intervals corresponds to the upper (lower) value on HUI's empirical distribution corresponding to the empirical cumulative frequency of SAH. A second approach was proposed by Kerkhofs and Lindeboom [85] and Lindeboom and van Doorslaer [90]. They stratify the population in several groups according to some individual characteristics and then estimate an ordered response model of SAH on HUI as proxy of true health. This estimation approach allows differences both with respect to cut-points and index-shift. On the same fashion Etilé and Milcent [50] use latent class analysis to construct a synthetic measure of clinical health and estimate a generalized ordered logit to investigate the effect of socio-economic status (e.g income level) on self reporting behavior. To assess the magnitude of reporting heterogeneity related to income they follow Kerkhofs and Lindeboom [85] and assume that all the information on true health are captured by the synthetic measure of clinical health, that is, they argue that individual characteristics should be *ignorable* to predict SAH. Therefore reporting heterogeneity is tested considering whether these characteristics have a significant effect after conditioning on clinical health.

Etilé and Milcent's [50] approach follows a two steps procedure. First they build a synthetic index of "true" health based on a set of self-reported health conditions including SAH and then regress the SAH on this index and other socioeconomic characteristic to test self reporting heterogeneity. However this approach does not model endogenously the self reporting behaviour and the heterogeneity in the health production. Thus, another possible approach which allow to model jointly self reporting and "true" health status heterogeneity relies on the use of multiple indicators. Shmueli [107] estimates a structural equation model exploiting some features of multiple indicators-multiple causes (MIMIC) modelling to shape the relationships between true health and a set of indicators (Joreskog and Goldberger [81]). The latent class approach offer the same advantage of the MIMIC model to undertake in a single step the estimation of the effect of covariates on both SAH and 'true' health.

In this paper we use some recent developments on finite mixture models to provide an empirical assessment of reporting heterogeneity using a set of "manifest" (objective and subjective) health indicators in a recursive model with unobserved latent classes. In particular our aim is to investigate how to disentangle the effect of individual characteristics on health production on the one hand, and its self-reporting effect on the other hand. Further we evaluate the magnitude of some individual characteristics on self-reporting heterogeneity considering the residual association between self-reported indicators conditioning on true health. Our econometric approach allows us to model explicitly the residual association between indicators allowing an explicit recursive structure, which make our approach substantially different from those relying on MIMIC (Shmueli [107]).

Our approach also differs from the previous literature regarding the type of indicators exploited to measure "true" health. In fact "true" health, intended as clinical and physical health status, is constructed using both subjective (SAH and other self reported health conditions) and objective indicators (biomarkers).

On the one hand, this avoids the arbitrariness of excluding SAH itself from the set of measures indicating clinical health. On the other hand, there is a great deal of interest among researchers on biological measures for several reasons. Biomarkers can be used not only to validate respondents' self-reported health measures but also to identify true health status and compare different groups of individuals (Banks *et al.* [7]); using biological measures give also the possibility to take into account the preclinical levels of disease even when the respondents may not have been aware. Using individual biomarkers, Johnston *et al.* [79] provide an important piece of evidence and show the existence of reporting heterogeneity comparing self-reported and objective hypertension.

We identify two unobserved classes representing people in good health and those in ill-health respectively. Our main finding provides further evidence of heterogeneity in self reporting behavior. In fact after conditioning on unobserved individual 'true' health-types personal characteristics cannot be ignored to predict SAH. In particular for a given level of "true" health people with higher income, better education and living in less deprived areas tend to report systematically better health. Moreover there is evidence that individual characteristics affect differently the reporting behaviour in each category of SAH.

The paper is organized as follows. In the next section we describe the data we use on our analysis. The following section explains the methods and the empirical strategy we exploit. Empirical findings are found in section 2.4. The last section discusses the results and concludes.

2.2 Data and variable definition

We use cross-sectional data from two waves (2003 and 2004) of the Health Survey for England (HSE). This is a large survey covering a wide range of fields related to socioeconomic status, health and life-style. In 2003 the major focus of the survey was cardiovascular disease (including heart attacks and strokes), which is one of the largest causes of death in England. Even when this type of disease is not fatal, it brings ill-health and disability which might deeply affect individuals' life. Therefore it is extremely important to obtain objective measures of health risk in order to measure individual "true" health, intended as absence of physical disability or illness. For this reason HSE is very suitable for our aims because it contains some biological measures (biomarkers) obtained from a blood sample and used in general as an indicator of a biological state. The same biological measures are available for all the sample aged 16 years or over in 2003, while only for a subgroup which represents minority ethnic groups in 2004.

The Health Survey for England is composed by two parts, an interviewer-administered interview (Stage 1), and a visit by a nurse to carry out measurements and take a blood sample (Stage 2). At each stage participants are asked to decide whether to proceed with the following stage or not. Therefore someone may agree to take part at Stage 1 but decide not to continue to Stage 2.

In the first stage an individual questionnaire is administered in order to collect information on general health, eating habits, physical activity, smoking, drinking, family cardiovascular disease history and socioeconomic status (e.g. income, employment status, educational background). At the end of this stage respondents are asked to proceed with stage 2 by fixing an appointment with a qualified nurse. In this second stage nurses ask more information on health and health care utilization. For those to older than 35 in 2003 and 16 in 2004 the nurses also ask to provide a fasting blood sample and a blood pressure measurement.

For our analysis we consider a homogeneous sample of individuals aged 30 or over excluding cases with incomplete or inconsistent information on the relevant socioeconomic, demographic, health and life-style variables. The original sample size of individuals eligible to have a nurse visit was about 10,000 observations. After cleaning for missed (or inconsistent) observations, the remaining sample size consists of 3,381 observations. This reduction in the sample size is mainly due to the biomarker variables which are of main interest in the analysis. First notice that these variables are available only for an eligible subsample. Second for each measures used to define the objective health measures, we consider only those individuals with a valid lab test result. This derived measure is directly available in the HSE data set and excludes observation with a potentially unreliable lab test result - for example individual has eaten or smoked before the blood test, etc.

There are three important sets of variables relevant for our analysis (see tables A.1 and A.2). The first set includes three binary objective health indicators obtained considering only valid measures of the lab tests and excluding all the cases in which lab test results have been affected by individual behavior (e.g. people that have smoked or eaten before the nurse visit, etc.). The first indicator

(BPN) takes a value 1 if individual has normal blood pressure measured by a qualified nurse with Dinamap and Omron measures. Our second objective indicator (CHL) is a binary measure derived from the total cholesterol/high density lipoprotein ratio in the blood. This ratio is more indicative of cardiovascular disease than total cholesterol since it consider both high and low density lipoprotein cholesterol. Then the variable CHL takes a value 1 if individual has the cholesterol ratio below a sex-adjusted threshold indicating a low risk of cardiovascular disease. In particular for men an acceptable ratio of total cholesterol/high density lipoprotein is 4.5 or below, and for women is 4.0 or below. Finally our last objective indicator is based on the c-reactive protein (CRP) blood test. CRP may be used to screen apparently healthy people for cardiovascular disease (CVD). If the CRP level in the blood drops, it means that individual are getting better and CVD risk factor is being reduced. The CRP indicator takes 1 if individual have a lab test score lower 3.0 mg/L, associated to low chance of having a sudden heart problem.

Since available biomarkers capture mainly cardiovascular risk and thus have a limited role in capturing overall true health, we also include two additional health indicators: SAH and self-reported limiting longstanding illness (LLI). This latter variable is available directly from the survey and it was derived considering whether individual has longstanding illness and whether daily activities are limited due to this illness. LLI takes 1 if individual has no chronic limitations on daily activities.

As part of the health questionnaire of the HSE, respondents were asked: "how is your health in general?" The response categories were excellent, very good, good, fair, bad or very bad. We combine three SAH's categories (very bad, bad and fair) in one category representing poor health, because only a relatively small fraction (5.5%) of the sample reports very bad health. The SAH variable thus consists of three ordered categories: poor, good and very good health. Notice however that our definition of "true" health is different from the broader definition which include individual mental and psychosocial status. In fact these indicators refer mainly to "true" health as absence of physical pain, physical disability, or an objective measured condition that is likely to cause illness (e.g. high blood pressure can be interpreted as symptoms of a cardiovascular disease, etc.).[1]

[1]Notice that this definition of "true" health can be extend to include mental health using a mental score included in the HSE dataset. However its inclusion would render slower the estimation of the model throughout the EM algorithm, since this indicator is primarily coded using a score defined in 12 point scale and a binary discretization would be arbitrary.

Finally we follow Johnston *et al.* [79] to define the last set of variables which provides information on socio-economic status, demographic characteristics and life-style. Socioeconomic status is measured using the equivalised income and an overall index of multiple deprivation (IMD2004). This is a composite index of relative deprivation at small area level, based on seven domains of deprivation involving for example income, employment, health deprivation and disability, education, crime and living environment.[2] This survey is also rich in information on individual life-style. These variables offer a good opportunity to better identify individual health. In particular there are detailed information on past and present smoking behavior as well as physical activities, sport intensity and the daily number of portion of fruits and vegetables. HSE provides also a three levels fat score ranged from "low fat" to "high fat" eating habits derived considering the consumption of cheese, fish, fried food, meat, etc. However including life-style variables across regressors may introduce an important source of endogeneity since individuals who are in good health may also have healthier behaviours.

2.3 The Model

Our aim is to study the association between SAH and "true" health, using recent developments on latent class analysis, which allow covariates to affect latent class membership, and possibly residual association among indicators after conditioning on latent health-types (Huang and Bandeen-Roche [75], Bartolucci and Forcina [13] and Dardanoni, Forcina and Modica [36]).

Since we are mainly interested into disentangle the effect of individual on self-reporting behaviour, let U be a latent discrete variable with two categories representing individuals in good ($U = 0$) and bad ($U = 1$) health. The main problem when one wants to study the relationship between true health and self-reported health is to disentangle the effect that some personal characteristics have on "true" health's variations from the effect that the same variables have on self-reporting.

Following Kerkhofs and Lindeboom [85], Etilé and Milcent [50] suggest the following strategy to distinguish between these two effects. They assume

[2] The Equivalised income variable is provided by the HSE. It is computed using the McClement score for each household (dependent on number, age and relationships of adults and children in the household), and then dividing the total household income by this score to get an equivalised household income.

that "true" health is entirely captured by a synthetic measure of clinical health (which they denote H^0) for which the following ignorability condition holds (see Wooldridge [121], p. 63):

$$Pr(Y_{sah} = i \mid H^0, \boldsymbol{z}) = Pr(Y_{sah} = i \mid H^0) \qquad i = 1, 2, 3 \tag{2.1}$$

where H^0 was obtained using a latent class model with self-reported health measures as indicators. The assumption above relies on the fact that the effect of covariates $\boldsymbol{z}$ on "true" health is entirely captured by H^0. Thus, if SAH is a reliable indicator of individual health, then any differences on personal characteristics should not affect the distribution of SAH after conditioning on H^0, which means that $\boldsymbol{z}$ is ignorable to predict SAH. Therefore if the assumption above holds, a test of self-reporting behavior can be easily performed regressing Y_{sah} on H^0 and $\boldsymbol{z}$ and testing whether parameters of personal characteristics are still significant conditioning on the synthetic measure of clinical health.

Our approach can be considered an extension of Etilé and Milcent's [50] test from two points of view. First, using a LC approach the estimation of the effect of covariates on 'true' health and SAH is undertaken in a single step. This means that we estimate endogenously individual "true" health U by taking information both from subjective (SAH, LLI) and objective indicators (BPN, CRP, CHL), so that all available information from health indicators, including SAH, are used to estimate individual "true" health. This also makes clear the effect of covariates on both unobserved U and SAH. Second we allow for residual association between health indicators in order to capture any *adaptation effect* of individuals to their own health condition. This means that the effect of subjective health measure, such as Y_{lli}, on SAH status is not only driven by U but it could also affect *indirectly* self-reporting behavior of SAH itself - for example people may adapt to a chronical limitation status measured by Y_{lli} and report systematically better health (see e.g. Groot [65]).

Let $\boldsymbol{Y} = (Y_{sah}, Y_{lli}, Y_{bpn}, Y_{chl}, Y_{crp})$ be the vector of observable response variables. As a simple starting point, consider the traditional latent class analysis (see e.g. Goodman [63]), which implies the existence of a discrete U such that these observables $\boldsymbol{Y}$ are independent conditionally on U. This is also named "local independence" and is expressed as:

$$Pr(Y_{sah}, Y_{lli}, Y_{bpn}, Y_{chl}, Y_{crp}) = \sum_{u=0}^{1} Pr(Y_{sah} \mid u) \cdots Pr(Y_{crp} \mid u) Pr(u)$$

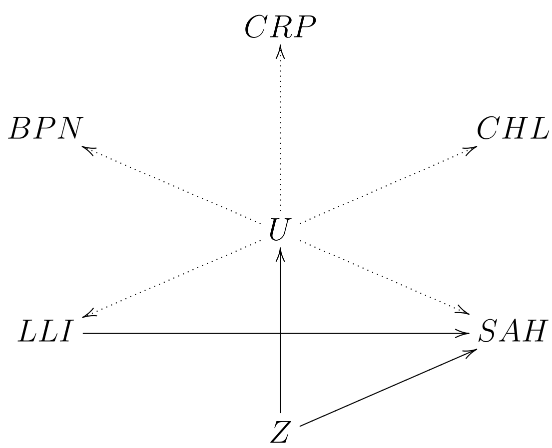
Clearly a U that makes these responses conditionally independent captures elements of individuals' "true" health. However, the local independence assumption is too restrictive for our purposes since it does not allow responses and latent health-types to depend on covariates, and it does not allow residual association between any response after conditioning on U.

Our model assumes that the joint distribution of responses $(U, \boldsymbol{Y})$ conditional on the set of observable covariates $\boldsymbol{z}$ (describing demographic, socioeconomic and life-style individual characteristics) is fully determined by the following set of conditional distributions of observables, and by the marginal distribution of U:

$$\begin{aligned} & Pr(U = 1 \mid \boldsymbol{z}), \\ & Pr(Y_{sah} = i \mid Y_{lli}, \boldsymbol{z}, U) \\ & Pr(Y_{lli} = 1 \mid U) \\ & Pr(Y_{bpn} = 1 \mid U) \\ & Pr(Y_{chl} = 1 \mid U) \\ & Pr(Y_{crp} = 1 \mid U) \end{aligned} \tag{2.2}$$

which can be equivalently formulated in terms of the *directed acyclic graph* (see Pearl [98]) reported in figure 2.3.

Figure 2.1: *Directed Acyclic Graph of model* (2.2)



We then model these conditional probabilities as linear functions of the covariates $\boldsymbol{z}$ using a logit link to form a multivariate regression system of logit equations:

$$\begin{aligned}
&\Pr\left(U=1\,|\boldsymbol{z}\right)=\Lambda\left(\alpha_u\left(u\right)+\boldsymbol{z}'\boldsymbol{\gamma}\right)\\
&\Pr\left(Y_{sah}>1\,|u,y_{lli},\boldsymbol{z}\right)=\Lambda\left(\alpha_1\left(u,y_{lli}\right)+\boldsymbol{z}'\boldsymbol{\beta}_1\right)\\
&\Pr\left(Y_{sah}>2\,|u,y_{lli},\boldsymbol{z}\right)=\Lambda\left(\alpha_2\left(u,y_{lli}\right)+\boldsymbol{z}'\boldsymbol{\beta}_2\right)\\
&\Pr\left(Y_{lli}=1\,|u\right)=\Lambda\left(\alpha_3\left(u\right)\right)\\
&\Pr\left(Y_{bpn}=1\,|u\right)=\Lambda\left(\alpha_4\left(u\right)\right)\\
&\Pr\left(Y_{crp}=1\,|u\right)=\Lambda\left(\alpha_5\left(u\right)\right)\\
&\Pr\left(Y_{chl}=1\,|u\right)=\Lambda\left(\alpha_6\left(u\right)\right)
\end{aligned}\tag{2.3}$$

where Λ is a logit link function $\Lambda=e^t/(1+e^t)$. Note that $\alpha_j\left(u,y_{lli}\right)$, $j=1,2$, represents all the possible combinations between U and Y_{lli}; since U and Y_{lli} are binary this means we have 4 α parameters in the second and third equation that could be alternatively expressed as $a_{j,1}+a_{j,2}U+a_{j,3}Y_{lli}+a_{j,4}Y_{lli}U$. Notice that for sake of generality we do not assume any relationships between individual characteristics, "true" health and self-reported health status. Thus the same set of covariates affecting "true" unobserved health may also potentially affect reporting behavior. The system of equations (2.3) makes clear how the effect of individual characteristics on "true" health is separated from the effect on reported health, since individual characteristics affect separately the unobserved latent health status U and the SAH. In particular γ parameters capture the effect of individual characteristics on health status, while the β's the effect on reported health status.

Parameters in model (2.3) are estimated by the EM algorithm.[3] In the E step the posterior probability of latent class U given the observed configuration $\boldsymbol{y}$ is computed. The M-step maximizes a likelihood function that is further refined in each iteration by the E-step. Details on estimation and identification of model (2.3) can be derived by looking at the Appendix of Dardanoni, Forcina and Modica [36] and at Bartolucci and Forcina [13].

It is well known that the EM algorithm may converge even if the model is not identified, a crucial issue for finite mixture models. Local identification of model (2.3) can be obtained using the numerical test described by Forcina [59], which consists in checking that the Jacobian of the transformation between the parameters of the observable responses and the mixture model parameters is of full rank for a wide range of parameter values.

[3] We are grateful to Antonio Forcina for kindly providing the Matlab code for the estimation.

We propose two tests of reporting heterogeneity. The first test is nothing but the ignorability condition of $\boldsymbol{z}$ in the equation determining SAH (compare with Etilé and Milcent's [50] equation (1) above and Kerkhofs and Lindeboom [85]), that is:

$$Pr(Y_{sah} = i \mid U, Y_{lli}, \boldsymbol{z}) = Pr(Y_{sah} = i \mid U, Y_{lli}) \qquad i = 1, 2, 3 \tag{2.4}$$

which can be performed by testing whether $\boldsymbol{z}$ has a significant influence on SAH in model (2.3), that is, $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = \mathbf{0}$. Our second test is more specific and is focused on whether individual characteristics affect different parts of SAH distribution after conditioning on U and Y_{lli}, that is, testing whether $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2$. Both tests are performed by estimating a restricted model and computing a LR-test which has a chi-square asymptotic distribution.

2.4 Results

2.4.1 Generalized ordered logit results

As benchmark of our analysis we use the results obtained by a generalized ordered logit model of SAH on the set of indicators and individual characteristics:

$$\begin{aligned} \Pr\left(Y_{sah} > 1 \,|\boldsymbol{z}, \boldsymbol{w}\right) &= \Lambda\left(\alpha_1 + \boldsymbol{z}'\boldsymbol{\eta}_1 + \boldsymbol{w}'\boldsymbol{\theta}\right) \\ \Pr\left(Y_{sah} > 2 \,|\boldsymbol{z}, \boldsymbol{w}\right) &= \Lambda\left(\alpha_2 + \boldsymbol{z}'\boldsymbol{\eta}_2 + \boldsymbol{w}'\boldsymbol{\theta}\right) \end{aligned} \tag{2.5}$$

(compare with the second and third equations of system (2.3)), where $\boldsymbol{z}$ is the vector of socio-economic, demographic and life styles characteristics as above, and $\boldsymbol{w}$ is a vector of health variables. Notice that following Etilé and Milcent [50] (and to make the regression systems (2.3) and (2.5) directly comparable), we assume that only the coefficients of $\boldsymbol{z}$ are allowed to vary across the categories of SAH; while health variables are assumed to affect uniformly the SAH's distribution.

We first tested the null hypothesis of parallel lines - called the *proportional odds assumption* in the statistical literature (Agresti [4], p. 275) - by imposing the restriction that $\boldsymbol{\eta}_1 = \boldsymbol{\eta}_2$; the likelihood ratio test is equal to 63.25 with 37 df (p-value .0045). Thus, the hypothesis of parallel lines is rejected.

Table (A.4) shows the estimated coefficients from the generalized ordered logit. Results show that with this specification there are several individual characteristics that affect SAH. Under the assumption that biomarkers fully capture "true" health, differences in SAH should reflect heterogeneity in reporting behaviours, rather than genuine variation in the health status. Thus people with higher income who live in less deprived area tend to report better health. Therefore our results on income related differences in the probability to report better health are close to many other papers about the existence of income-related reporting heterogeneity (Hernandez-Quevedo *et al.* [71], Etilé and Milcent [50], Lindeboom and van Doorslaer [90]). Let consider now results on health related life variables. One observes that sport and physical activities, also increase the probability to report good health, while individuals who eat more portions of fruits and vegetables per day, are more likely to report very good health than just at least good health. Clearly it is hard in this framework to claim that the effect of individual characteristics on the probability to report subjective health is only due to reporting heterogeneity since it doesn't clearly distinguish the effect of personal characteristics on"true" health from the effect on self-reporting behavior - for example life styles may have a positive effect on "true" health, but they may also systemically induce individual to over(under) report individual subjective health. For this reason we then also estimate model (2.3) where observables are allowed to affect both individuals' health and reporting heterogeneity.

2.4.2 Results from model (2.3)

Intercepts

Since U is binary, table (A.5) shows $2^{*}4 + 2^{*}2^{2} + 1 = 17$ estimated intercepts α. In particular there is 1 parameter to describe the class membership probability, 2 parameters for each of the 4 health indicators, and 4 parameters to describe the effect of U and Y_{LLI} on SAH for people who report at least good or very good health.

A glance at the table reveals that people with good health ($U = 0$) are much more likely to have desirable lab test scores and no limiting longstanding illness compared to people with ill-health ($U = 1$). Furthermore, it is easily checked that people with $U = 0$ are also much more likely to report at least good or very good health conditional on $Y_{lli} = 0, 1$. Regarding the effect of Y_{lli} on SAH, it is also easily checked that the probability to report at least good or very good health is also increasing in Y_{lli} conditional on $U = 0, 1$.[4] This result provides significant evidence on the existence of heterogeneity in self-reporting behavior related to differences in self-perceived limiting illness.

Variations in the unobserved "true" health U.

The first column of table (A.6) reports the estimated parameters $\boldsymbol{\gamma}_u$ of both health related variables and socioeconomic characteristics:

- the effects of demographics characteristics on unobserved "true" health have the expected sign. In fact, ill-health is positive and statistically related to age, but negatively with sex and ethnicity. As expected women tend generally to have better health than men (Wingard [119]);
- socio-economic characteristics play an important role in health determination. In particular those with higher education have lower probability of being classified with poor health than those with no qualification. Health status is also strongly and positively correlated with income and social class as showed by estimated coefficients of equivalised income and social class. This results support findings obtained by Johnston *et al.* [79] who found in the same data no significant income gradient for self-reported chronic hypertension, but a clear negative gradient for objective 140/90 hypertension. This indicates that individuals living in most deprived household are significantly more likely to have hypertension. Another important role on determining individual health is also played by the index of multiple deprivation. Individuals who live in highly deprived area register a lower level of health than those living in less deprived area, although the effect seems to statistically vanishes as the deprivation decreases. Finally there is a small and negative statistical significant effect on health of the number of months individual lived in the same area;
- unobserved health is also related with individual life-style characteristics; individual who are no smokers who practice sport regularly with a moderate physical activity and follow a diet with a low fat content have a greater probability of having good unobserved health; the opposite holds for individuals who are obese with cardiovascular conditions in the family. Finally

[4] Just as an example, the probability that people with $U = 0$ report very good health is on average .52 if $Y_{lli} = 0$ while it is .97 if $Y_{lli} = 1$.

ill-health seems to be negatively correlated with the parents' age and with to the number of units drank in the heaviest day in the last seven days. This last result is clearly unexpected, although it could be related to a sort of measurement error in the drinking-unit variable.

Variations in reporting behaviour

The discussion above shows significant variation in unobservable health status by personal characteristics, representing a considerable source of unobserved heterogeneity in health production which should be taken into account. In the present section, we analyse the direct relationships between self-reported health and observable characteristics conditional on true unobserved health status and no limiting longstanding illness. Recall that for the sake of generality we have assumed (see (2.3)) that the set of covariates $\boldsymbol{z}$ affecting "true" unobserved health may also potentially affect reporting behavior.

We first tested the parallel line assumption by estimating a restricted model in which individual characteristics have the same effect on different categories of the SAH distribution. The value of log-likelihood for the restricted and unrestricted model is equal to -10979 and -10953. The value of likelihood ratio test is 53.72 which is clearly rejected with 37 d.f (p-value = .037). Results on unobserved heterogeneity in self-reported behavior are reported in the last four columns of table (A.6). They differ slightly with respect of SAH classes and can be summarized as:

- after conditioning on U and Y_{lli} a wide set of variables (such as ethnicity, education and individual life styles) are not anymore statistically significant in model (2.3) in both SAH's categories as compared with the results obtained in the generalized order logit model discussed above. In fact only some variables which appeared to significantly affect both SAH'categories in model (2.5) are also significant in both at least good and very good health in model (2.3). For example, after conditioning for unobserved health-types, having any qualification or being an individuals from ethnic minorities has no effect on SAH, contrary to the generalized logit model.
- individual with higher income tend to self report better health status. On the contrary people living in the most deprived area and with low educational attainment report systematically a worse level of health. The

magnitude of the effect does not differ significantly among SAH's classes. Therefore our results support the idea that reporting heterogeneity in SAH depends on socio-economic conditions, as suggested by [50];

- age affects self-reporting behavior. In particular elders tend to report better health than expected. This result seems plausible with previous findings (Lindeboom and van Doorslaer [90], Groot [65]) and confirms the apparently puzzle between self reported health and age, which is related to the existence of individual adaption to chronical ill conditions;

- individual's life-styles tend also to affect differently reporting behavior. Physical activities and sports increases the probability to report "very good" health but not at least "good" health. Thus, our findings are very similar to those obtained by Johnston *et al.* [79], indicating that individual with healthy life may over estimate their health status and then tend to over report subjective health. This effect is also increasing in the effort required by the activity itself, but is smaller compared with those obtained with the generalized ordered logit. However notice that these relationships between SAH, unobserved health and individual's health related behaviour may also reflect a potential source of endogeneity. A possible solution which can be addressed in the future is to include in this framework individual behaviour in the past which should affect subsequent health status.

2.5 Discussion and final remarks

The present study explores the relationships between socio-economic, demographic and life-style personal characteristics and health. In particular we test the existence of reporting heterogeneity on SAH implementing the approach proposed by Etilé and Milcent [50] and Kerkhofs and Lindeboom [85]. Our empirical strategy is innovative in two ways. First we use some recent developments on LCA to disentangle the effect of personal characteristics on self-reporting behavior from the effect on heterogeneity in health production, and we allow residual association between self-reported indicators in order to capture differences related to reporting heterogeneity after conditioning on latent "true" health. Second, to identify unobserved individual latent health we use not only subjective measures, but also objective indicators, such as biological measures, which help to validate respondents' self reports and to identify individual health by taking into account the pre-clinical levels of disease even when the respondents may not have been aware (Banks *et al.* [7]).

Our results confirm the existence of systematic self-reporting bias which has been found in many other empirical investigations. However, after conditioning on individual unobserved health-type, we find that several individual characteristics do not have a statistically significant effect on self-reporting behavior compared with the generalized order logit which does not distinguish explicitly between heterogeneity in health and reporting behavior.

Appendix A

Tables

Table A.1: Variable Definitions

Variable	Definition
chl	Total/high density lipoprotein cholesterol (1 = if lab test score is good, 0 otherwise)
crp	C-reactive protein (1 = if lab test score is lower than 3 mg/L, 0 otherwise)
bpn	Blood pressure (1 = normotensive with Dinamap and Omron readings), 0 otherwise)
lli	Limiting Longstanding Illness (1 = if no Limiting Longstanding Illness, 0 otherwise)
sah	Self-Assessed Health status (1 = "poor health", 2 = "good", 3 = "very good")
mar	1 = if individual is married, 0 otherwise
age	age of individuals
women	1 = female, 0 otherwise
black	1 = black, 0 otherwise
white	1 = white, 0 otherwise
noqual	1 = no qualification, 0 otherwise
eduh	1 = second level or higher, 0 otherwise
scl2	1 = social class for skilled non-manual and skilled manual
scl3	1 = social class professional and managerial technical
eqvinc	Equivalised income
imd3	1 = third quintile of Overall Index of Multiple Deprivation
imd4	1 = fourth quintile of Overall Index of Multiple Deprivation
imd5	1 = fifth quintile of Overall Index of Multiple Deprivation (most deprived)
hse04	1 = if individual belongs to HSE 2004, 0 otherwise
bmil	value of Body Mass Index if it is lower than 18.5, 0 otherwise
bmih	value of Body Mass Index if it is higher than 29.9, 0 otherwise

Table A.2: Variable Definitions

Variable	Definition
drinkun	# of drinking units in the heaviest day
agema	age of mother
agepa	age of father
demam	1 = whether the mother is dead
depa	1 = whether the father is dead
famcvd	1 = whether there are cardiovascular conditions in the family history
smacc	1 = if someone smokes in the accommodation
smkc	1 = if individual smokes currently
smkevr	1 = if individual has ever smoked
smkex	1 = if individual is an ex smoker
smkoc	1 = if individual smokes occasionally and is an ex-smoker
sportm	1 = moderate sport activity
sportr	1 = regular sport activity
hrsspt	# of hours of sport per week
phy2	1 = medium physical activity level
phy3	1 = high physical activity level
veg	# of portions of fruits and vegetables per day
fatt2	1 = if individual's diet has a medium fat score
fatt3	1 = if individual's diet has a high fat score
livehm	# of months individual has lived in this local year
urban	1 = if individual lives in an urban area

Table A.3: Descriptive Statistics

	Mean	S.D		Mean	S.D
chl	0.5927	0.4913	drinkun	3.1656	2.4252
crp	0.7403	0.4385	agema	70.1709	11.9926
bpn	0.6613	0.4733	agepa	68.9665	11.5362
lli	0.5862	0.4925	demam	0.4285	0.4949
sah	2.1572	0.7579	depa	0.5894	0.4920
mar	0.7071	0.4551	famcvd	0.1230	0.3285
age	49.3688	13.0599	smacc	0.8041	0.3968
women	0.5211	0.4996	smkc	0.1736	0.3788
black	0.0337	0.1805	smkevr	0.3921	0.4883
white	0.8550	0.3520	smkex	0.2729	0.4455
noqual	0.1878	0.3906	smkoc	0.1401	0.5106
eduh	0.7749	0.4176	sportm	0.2691	0.4435
scl2	0.3620	0.4806	sportr	0.1227	0.3281
scl3	0.5034	0.5000	hrsspt	1.2022	3.0489
eqvinc	3.1720	2.7412	phy2	0.4046	0.4908
imd3	0.2037	0.4028	phy3	0.3022	0.4593
imd4	0.1768	0.3816	veg	3.8218	2.4619
imd5	0.1336	0.3403	fatt2	0.1520	0.3590
hse04	0.1685	0.3744	fatt3	0.0301	0.1710
bmil	0.5025	3.0435	livehm	160.0535	208.1209
bmih	18.8296	14.3018	urban	0.1792	0.3836

Sample Size=3,381

Table A.4: Results for the Ordered Logit Model

	β_1	S.E.	β_2	S.E.
mar	0.198*	0.11	0.063	0.09
age	0.000	0.01	0.003	0.01
women	0.105	0.11	-0.035	0.08
black	0.313	0.28	0.405	0.25
white	0.777**	0.22	0.689**	0.17
noqual	-0.862**	0.31	-0.594**	0.22
eduh	-0.695	0.31	-0.281	0.21
scl2	0.066	0.14	-0.202	0.13
scl3	0.219	0.16	0.060	0.13
eqvinc	0.711**	0.26	0.397**	0.16
imd3	-0.076	0.14	-0.139	0.11
imd4	-0.254*	0.15	-0.227*	0.11
imd5	-0.798**	0.16	-0.429**	0.14
hse04	-0.058	0.18	-0.005	0.13
bmil	0.010	0.01	-0.011	0.01
bmih	-0.002	0.00	-0.008**	0.01
drinkun	0.022	0.02	0.017	0.02
agema	-0.004	0.00	-0.001	0.01
agepa	0.009*	0.00	0.002	0.01
demam	-0.231	0.14	0.099	0.11
depa	0.014	0.14	0.101	0.10
famcvd	-0.243	0.16	-0.105	0.13
smacc	0.276*	0.16	0.054	0.12
smkc	-0.271	0.25	-0.237	0.18
smkevr	-0.164	0.20	0.141	0.14
smkex	-0.045	0.21	-0.034	0.15
smkoc	0.007	0.14	0.008	0.09
sportm	0.405**	0.15	0.230**	0.06
sportr	0.522*	0.27	0.419**	0.16
hrsspt	-0.013	0.02	0.026**	0.01
phy2	0.622**	0.12	0.269**	0.10
phy3	0.715**	0.15	0.436**	0.11
veg	0.004	0.02	0.034	0.01
fatt2	0.209	0.15	0.016	0.10
fatt3	-0.032	0.28	0.084	0.23
livehm	-0.000	0.01	-0.000	0.00
urban	0.201	0.14	0.073	0.11
intercept	-0.426**	0.66	-2.761**	0.50

Parameters of vector $\boldsymbol{w}$

	θ	S.E.
lli	1.589**	.07
crp	0.162*	.08
chl	0.172	.09
bpn	0.157**	.07

** Significant at the 5% level;
* Significant at the 10% level.

Table A.5: Estimated intercepts of model 2.3

	α	S.E.	Prob.
$U = 1$	0.1615	0.1249	0.54
$chl \mid U = 0$	1.6278	0.0853	0.84
$chl \mid U = 1$	-0.5323	0.0596	0.37
$crp \mid U = 0$	1.8694	0.0852	0.86
$crp \mid U = 1$	0.5100	0.0539	0.62
$bpn \mid U = 0$	1.9212	0.0934	0.87
$bpn \mid U = 1$	-0.1277	0.0556	0.47
$lli \mid U = 0$	0.8031	0.0599	0.69
$lli \mid U = 1$	-0.0378	0.0521	0.49
$sah > 1 \mid U = 0, lli = 0$	-0.1717	0.1109	0.46
$sah > 1 \mid U = 0, lli = 1$	2.5915	0.1517	0.93
$sah > 1 \mid U = 1, lli = 0$	-1.6320	0.1219	0.16
$sah > 1 \mid U = 1, lli = 1$	0.6901	0.1063	0.67
$sah > 2 \mid U = 0, lli = 0$	0.0806	0.1097	0.52
$sah > 2 \mid U = 0, lli = 1$	3.4650	0.2195	0.97
$sah > 2 \mid U = 1, lli = 0$	-1.3091	0.1478	0.21
$sah > 2 \mid U = 1, lli = 1$	1.3354	0.1652	0.79

Table A.6: Estimated covariates' coefficients of model 2.3

	γ_u	S.E.	β_1	S.E.	β_2	S.E
mar	-0.1079	0.2216	-0.0159	0.1303	0.1689	0.1172
age	0.1493**	0.0183	-0.0070	0.0103	0.0130**	0.0067
women	-2.5172**	0.2734	-0.0290	0.1366	-0.0849	0.1199
black	-2.2479**	0.5418	0.6449	0.3430	0.1330	0.3403
white	-0.1400	0.3893	1.1140**	0.2352	0.3628	0.2324
noqual	-0.2231	0.6467	-0.5668	0.3942	-0.7306**	0.2591
eduh	-1.0248	0.6119	-0.3372	0.3477	-0.4912*	0.2532
scl2	-0.5605	0.3170	0.0747	0.2156	-0.2078	0.1485
scl3	-0.3788	0.3267	0.1987	0.2198	0.1065	0.1601
eqvinc	-0.8082*	0.2967	0.5116**	0.2334	0.5070**	0.2174
imd3	0.3715	0.2514	-0.1228	0.1463	-0.0855	0.1373
imd4	0.4941*	0.2775	-0.0789	0.1684	-0.3462**	0.1472
imd5	1.5698**	0.3452	-0.8177**	0.2251	-0.4448**	0.1659
hse04	0.2451	0.3162	0.1559	0.1990	-0.2741	0.1878
bmil	-0.0721	0.0468	0.0042	0.0134	-0.0510	0.0399
bmih	0.1355**	0.0119	-0.0083*	0.0046	0.0044	0.0057
drinkun	-0.1297**	0.0456	0.0364	0.0284	0.0057	0.0236
agema	-0.0137	0.0108	0.0083	0.0076	-0.0078	0.0048
agepa	-0.0183*	0.0106	-0.0027	0.0069	0.0082	0.0046
demam	0.0627	0.2542	0.0705	0.1646	-0.0412	0.1427
depa	-0.0534	0.2495	0.0554	0.1482	0.1146	0.1529
famcvd	0.5728*	0.3313	-0.2160	0.2311	-0.1146	0.1499
smacc	-0.6574**	0.3137	-0.1505	0.1971	0.2991**	0.1556
smkc	0.8732**	0.4396	-0.7655**	0.2664	0.2281	0.2349
smkevr	-0.0649	0.3346	0.1928	0.2082	-0.0301	0.1961
smkex	0.0471	0.3558	-0.0078	0.2282	-0.0309	0.1929
smkoc	0.0710	0.2354	-0.0269	0.1363	0.0546	0.1340
sportm	-1.1110**	0.2469	-0.1289	0.1425	0.4725**	0.1519
sportr	-1.1047**	0.3918	-0.1735	0.2262	0.8126**	0.2994
hrsspt	0.0487	0.0324	0.0706**	0.0281	-0.0432	0.0288
phy2	-0.5596**	0.2572	0.5506**	0.1708	0.3364**	0.1229
phy3	-1.3925**	0.2989	0.5664**	0.1809	0.6059**	0.1489
veg	0.0005	0.0400	0.0270	0.0245	0.0344	0.0219
fatt2	-0.0002	0.2652	-0.2242	0.1615	0.2951**	0.1393
fatt3	0.7564	0.5752	0.6370	0.4567	-0.2132	0.2617
livehm	-0.0011**	0.0005	-0.0005	0.0004	-0.0001	0.0002
urban	0.3106	0.2605	0.0596	0.1620	0.1475	0.1448

** Significant at the 5% level;
* Significant at the 10% level.

Chapter 3

Testing for Selection Effects in Insurance Markets with Unobservable Types

3.1 Introduction

The effects of private information on the efficient operation of insurance markets has been one of the most active research topics in economics, starting from the classic Rothschild and Stiglitz [104] (RS) paper. The standard RS adverse selection model predicts that when individuals have private information about their actual risk, the insurance contract will be adversely selected, with high risk individuals choosing higher insurance coverage; ex post, this will cause a positive correlation between risk and coverage. This observation has inspired the seminal contribution by Chiappori and Salanié [25], who considered the testable implications of asymmetric information in insurance markets, and proposed the so called Positive Correlation (PC) test. The PC test rejects the null of absence of private information in a given insurance market when, conditional on consumers' characteristics used by companies to price contracts, individuals with more coverage experience more of the insured risk. The coverage-risk correlation has been shown to be a robust implication of competitive insurance markets under private information in many different settings (see e.g. Chiappori *et al.* [24]).

The PC test has inspired a large and growing literature on empirical testing for asymmetric information in insurance markets. A recent paper by Cohen and Spiegelman [30] reviews almost a hundred empirical applications focusing on automobile, annuities, life, reverse mortgages, long-term care, crop and health insurance markets.[1]

Contrary to the theoretical prediction, a number of empirical analyses of insurance markets have found *negative* risk-coverage correlation, a phenomenon which has been named *favorable* selection.[2] For example, in their seminal paper on favourable selection in the US long-term care insurance market, Finkelstein and McGarry [58] find negative correlation between insurance purchase and nursing home use, both unconditionally and after conditioning on risk classification by insurers. As an explanation of this puzzle, Finkelstein and McGarry argue that there are two conflicting sources of private information, namely individual's actual risk and risk attitudes, so that two types of people buy insurance: high risk individuals and high risk aversion individuals. Ex post, the former are higher risk then predicted; the latter are lower risk; in aggregate, those who buy more insurance do not have higher claims. The same puzzling negative risk-coverage correlation has also been found in a recent paper by Fang *et al.* [53], who show that, conditional on controls for Medigap prices, individuals with Medigap insurance tend to spend *less* on medical care. This is explained again by the existence of multiple dimensions of private information, with cognitive ability being one of the key sources of favourable selection. The possibility that multidimensional private information may invalidate standard insurance model predictions has also been the subject of recent theoretical work, which has shown that the positive risk-coverage correlation may not necessarily follow from the existence of private information (see e.g. Chiappori *et al.* [24], de Meza and Webb [37], Smart [108], Villeneuve [114], Wambach [117]).

In practice, especially when insurance companies do not use all relevant information to price insurance contracts, either because it is unobservable or because it cannot be used for regulatory constraints or political economy considerations, it is quite likely that there are many individual characteristics which are not used by insurance companies to price contracts but may influence insurance coverage and actual risk occurrence. Thus, in many instances it is possible that there is multidimensional residual heterogeneity which affects the risk-coverage correlation, even after conditioning on variables used by insurers to price contracts. Notice however that multidimensionality of private information does not

[1] See also Einav *et al.* [47] for a recent review of testing for asymmetric information in insurance markets.

[2] Alternatively, some authors refer to this situation as *advantageous* or *propitious* selection. See Hemenway [70] for an early discussion of the relevance of favourable selection in insurance markets.

need to be necessarily related to regulation, which can have many additional other effects on market.

The first question addressed by this paper is then: How does the standard PC test perform in the presence of multidimensional private information? We argue that under multidimensional private information there is the possibility that the insurance contract is *both* adversely *and* favourably selected by different individuals. Since the PC test relies on a single statistic to appraise the risk-coverage correlation, multidimensional private information may cause serious problems in detecting selection effects using observable data, since the PC test averages out selection effects which may pull in different directions.

We then show how unobserved heterogeneity can be modeled by assuming a finite number of heterogeneous "types", which result by cross-classifying the relevant private information variables, and extend the standard adverse and favourable selection definitions into *local* and *global* ones. Using recent advances in finite mixture modelling, we show how risk and coverage probabilities can be estimated for the unobserved types in order to detect selection effects in the market. Tests for the absence of selection and the multidimensionality of private information can be performed by imposing restrictions on the behavior of these unobservable types.

To show the potential applicability of our approach, we look at the US long-term care and Medigap insurance markets, which differ substantially since the former is not heavily regulated as the latter. While the PC test shows absence of significant residual risk-coverage correlation in both samples, our finite mixture model shows the existence of significant residual heterogeneity with large selection effects. Not surprisingly, we find that selection effects are stronger in the regulated market.

3.2 The Positive Correlation test

In this section we review the main properties of the PC test in the standard Rothschild-Stiglitz model of adverse selection. As explained by Cohen and Spiegelman [30], when looking at the risk-coverage correlation, there is no unique way to define coverage and risk. When insurers offer different contracts, high risk individuals may choose contracts with more comprehensive coverage; when insurers offer a single product, high risk agents have a greater probability to buy insurance. Similarly, risk may be defined by higher expected claims, higher payouts in the event of a claim, or both.

To keep the analysis simple, we follow the original Chiappori and Salanié [25] analysis and assume we observe a binary variable $I \in \{0,1\}$ which takes value 1 if an individual has bought an insurance contract which protects from a fixed loss, and a binary variable $O \in \{0,1\}$ which takes value 1 if the individual incurs the loss. The terms of the contract depend on a set of individual observable characteristics $\boldsymbol{x}$ used by the insurance company. Thus, individuals with the same $\boldsymbol{x}$ are offered the same contract. In practice conditioning on $\boldsymbol{x}$ is key in these testing procedures. We consider first the analysis conditional on $\boldsymbol{x}$, that is, we consider the population of individuals who have the same characteristics $\boldsymbol{x}$.

Let $R \in \{0,1\}$ be a private information variable denoting risk type. Since R is not observed, the underlying question is *what can we learn about the distribution of O, I, R when we observe only O, I.* An important point to notice is that the positive risk-coverage correlation may arise, even in the absence of adverse selection, because of moral hazard. In particular, since moral hazard implies ex post positive correlation between risk and coverage even in the absence of adverse selection, finding positive risk-coverage correlation in the data does not provide concluding evidence of adverse selection, but negative or zero correlation in the data is not compatible with adverse selection. This is a well known issue, and an important task of current literature is disentangling the two effects. Advances are being made exploiting exogenous variations or the panel structure of data (see e.g. Abbring [2]-[1], Cardon and Hendel [21], Chiappori *et al.* [26], Dionne *et al.* [41], Einav *et al.* [48]); the issue is discussed with references in the reviews of Cohen and Spiegelman [30] and Einav *et al.* [47].

To focus on selection effects, let us assume that there is no moral hazard. This assumption is only partially restringing. Notice, in fact, that moral hazard effect of insurance differs between individuals. In particular some individuals whose behaviour is more responsive to insurance may be also more likely to buy (ex ante) insurance. Thus, we would still view this as selection, in the sense that individuals are selecting insurance on the basis of their anticipated behavioural response to it. Therefore the moral hazard effect we are not considered is then net net of the moral hazard induced by the selection of a contract by different types, who anticipate that their behavior will change after buying the contract. In particular it could be viewed as a variation affecting each individual in the loss occurrence and consumption subsequent to the contract purchasing.[3] Under this assumption we have:

$$P(O = 1 \mid R, I = 0) = P(O = 1 \mid R, I = 1). \tag{3.1}$$

The classic RS definition of adverse selection, that higher risk individuals (i.e. those with higher loss probability) are more likely to buy insurance, can be written as:[4]

Definition 1. (Adverse Selection 1)*: The insurance contract I is adversely selected if*

$$\big(P(O = 1 \mid R = 1) - P(O = 1 \mid R = 0)\big) \cdot \big(P(I = 1 \mid R = 1) - P(I = 1 \mid R = 0)\big) > 0.$$

An alternative definition of adverse selection in this context can be derived by noting that riskier types, when compared to the population, are more likely to experience the loss; the insurance contract I is adversely selected by them since they are also more likely to buy insurance compared to the population. On the other hand, less risky types, who have lower expected claims, also adversely select I since they are less likely to buy insurance compared to the population. If we denote by $\bar{P_O}$ and $\bar{P_I}$ the average loss and insurance probabilities,[5] the insurance contract I is adversely selected by type $R = i$ if

$$\big(P(O = 1 \mid R = i) - \bar{P_O}\big) \cdot \big(P(I = 1 \mid R = i) - \bar{P_I}\big) > 0.$$

Definition 2. (Adverse Selection 2)*: The insurance contract I is adversely selected if it is adversely selected by types $R = 0, 1$.*

Now, Definitions 1 and 2 are not directly testable since they involve the unobserved variable R. Consider then the following two testable conditions:

[3]This assumption is crucial to show simply the theoretical equivalence between the differen definitions of the positive correlation test. However, notice that since our paper is mainly concentrated on studying the existence of residual (multidimensional) private information in the insurance markets, without disentangling the effect of adverse selection from moral hazard, we follow several other studies on this issue which assume no moral hazard (see e.g. Chiappori and Salanié [25], Cohen and Einav [29], Finkelstein and McGarry [58], Fang *et al.* [53], Cutler *et al.* [32]).

[4]If we label $R = 1$ as the high risk individuals, Definition 1 could be written as $P(O = 1 \mid R = 1) > P(O = 1 \mid R = 0)$ and $P(I = 1 \mid R = 1) > P(I = 1 \mid R = 0)$. Definition 1 is slightly more general since it does not imply any label on R, which will be useful later when unobservable types could reflect multidimensional characteristics.

[5]That is, $\bar{P_O} = P(R = 0)P(O = 1 \mid R = 0) + P(R = 1)P(O = 1 \mid R = 1)$ and $\bar{P_I} = P(R = 0)P(I = 1 \mid R = 0) + P(R = 1)P(I = 1 \mid R = 1)$.

Definition 3. (Positive Correlation Test 1)*:*

$$P(O = 1 \mid I = 1) > P(O = 1 \mid I = 0),$$

and

Definition 4. (Positive Correlation Test 2)*:*

$$\frac{P(O = 1, I = 1) \cdot P(O = 0, I = 0)}{P(O = 1, I = 0) \cdot P(O = 0, I = 1)} > 1.$$

Definitions 3 and 4 are two ways to implement the PC test in this context. Definition 3 says that the expected loss for consumers who chose to insure is greater than for consumers who did not. Definition 4 says that the odds ratio between O and I should be greater than one, that is, O and I should be positively correlated, and goes back to the original implementation of Chiappori and Salanié [25]. Both alternative definitions of the PC test are discussed in the literature; for example, the review of Einav *et al.* [47] privileges the first, while the review of Cohen and Spiegelman [30] considers both. The following result clarifies the relationship between the four definitions, and is proved in the appendix.

Proposition 1. *Definitions 1, 2, 3, and 4 are equivalent under Assumption* (3.1)*.*

The proposition shows that under no moral hazard, adverse selection, which involves the unobservable risk type R, is equivalent to the positive correlation property which involves only observables.

In practice, since the insurance contract depends on the observable variables $\boldsymbol{x}$, the PC test is performed conditional on $\boldsymbol{x}$. As suggested by Chiappori and Selanié [25], when the variables in $\boldsymbol{x}$ are discrete with a very limited number of distinct configurations, one can define a finite number of mutually exclusive and exhaustive configurations (strata), and test the independence of I and O in each stratum. Since this is equivalent to imposing no restriction on how $P(O, I \mid \boldsymbol{x})$ depends on $\boldsymbol{x}$, the analysis is nonparametric. On the other hand, if covariates are continuous or take so many values that most strata contain too few subjects, the nonparametric approach is not viable, and Chiappori and Selanié [25] suggest to test the conditional correlation between O and I with a standard binary probit model

$$
\begin{aligned}
I &= 1(\boldsymbol{x}'\boldsymbol{\beta}_I + \epsilon_I > 0), && (3.2)\\
O &= 1(\boldsymbol{x}'\boldsymbol{\beta}_O + \epsilon_O > 0) && (3.3)
\end{aligned}
$$

with ϵ_I and ϵ_O standard normal errors. In particular, since risk type R is not observed, one can always rewrite, say, $\epsilon_I = \gamma_I R + \eta_I$ and $\epsilon_O = \gamma_O R + \eta_O$, with η_I and η_O being idiosyncratic errors and γ_I and γ_O positive constants, so that the null of the absence of private information amounts to testing that the correlation between ϵ_I and ϵ_O is zero. Notice that model (3.2,3.3) assumes no moral hazard since I has no direct effect on O. In their seminal application of the PC test to the automobile insurance market, Chiappori and Selanié [25] find that the null of the absence of private information cannot be rejected under both the nonparametric and parametric testing procedures.

3.2.1 The Positive Correlation test with multidimensional private information

Suppose now there are two unobserved binary variables, R and P, so that, after cross classification $R \times P$, there are four unobservable "types" (L, L), (L, H), (H, L) and (H, H). As argued in the Introduction, the existence and the effect of multidimensional private information in insurance markets has been analyzed both in the theoretical and the empirical literature. For example, Smart [108] studied a competitive insurance market in which individuals differ with respect to both accident probability and degree of risk aversion showing that multiple dimension of private information can change the nature of equilibrium introducing "noise" into the problem of inferring agents' type. However, it is important to notice that both the theoretical and empirical literature use different interpretations of "risk preference" P, which have included not only the canonical Arrow-Pratt risk aversion specification, but also context's specific risk perceptions which may affect both insurance choice and loss occurrence (see Einav *et al.* [49] for a recent contribution on context's specific risk perceptions).

Suppose that insurance and loss follow a standard bivariate probit model with

$$I = 1(\boldsymbol{x}'\boldsymbol{\beta}_I + \gamma_I R + \delta_I P + \eta_I > 0),$$

$$O = 1(\boldsymbol{x}'\boldsymbol{\beta}_O + \gamma_O R + \delta_O P + \eta_O > 0)$$

so that, letting $\epsilon_I = \gamma_I R + \delta_I P + \eta_I$ and $\epsilon_O = \gamma_O R + \delta_O P + \eta_O$, model (3.2,3.3) above is a special case with $\delta_I = \delta_O = 0$. Consider the following four examples, where for simplicity we assume that $\boldsymbol{x}$ has been centered and the analysis is conditional on $\boldsymbol{x} = \mathbf{0}$:

1. In the first example, suppose $\gamma_I = \gamma_O = \delta_O = 1$ and $\delta_I =$ -1; thus, as in the standard model, risk type has a positive effect on both the insurance and claim probabilities, but there is an additional unobservable variable (risk preference) which has a positive effect on claims and a negative effect on insurance. Let R and P be independent, respectively with support $\{-1, 2\}$ with probability (2/3,1/3), and support $\{-2, 1\}$ with probability (1/3,2/3). It can be checked that there is zero correlation between ϵ_I and ϵ_O, so that the PC test would conclude that there is no selection in this insurance contract. However, after calculating the population probabilities $\bar{P_O}$ and $\bar{P_I}$ and the insurance and outcome probabilities for the four types, it can be seen that the while the insurance contract is adversely selected by types (H, H), there is *favourable* selection by types (L, L) since, compared with the population, types (L, L) have lower claim probability but a greater probability of buying the insurance contract.

2. In the second example, let $\gamma_I = \gamma_O = 1$, $\delta_I =$ -1 and $\delta_O = 0$. This captures a simple extension of the standard model which assumes that risk preference has a negative effect on insurance purchase. Let R and P be independent, equiprobable, and with support respectively equal to $\{-1, 1\}$ and $\{-2, 2\}$. There is positive correlation between ϵ_I and ϵ_O, which is interpreted by the PC test as evidence of adverse selection. However, after calculating the insurance and outcome probabilities for the four types and for the population, it can be seen that the insurance contract is favourably selected by types (L, L) and (H, H), and adversely selected by types (L, H) and (H, L).

3. In the third example, let $\gamma_I = \delta_O = 1$ and $\delta_I = \gamma_O = 0$, so that insurance choice is affected only by risk preference, and claim probability only by risk type, and suppose R and P be independent. There is zero correlation between ϵ_I and ϵ_O, but the insurance contract is adversely selected by types (L, L) and (H, H), and favourably selected by types (L, H) and (H, L).

4. In the final example, let $\gamma_I = \gamma_O = 0$, $\delta_O = 1$ and $\delta_I =$ -1, so that actual risk has no effect on I and O, but risk preference has a negative effect on insurance choice and a positive one on claim probabilities. There is negative

correlation between ϵ_I and ϵ_O, and the insurance contract is favourably selected by both types $P = L$ and $P = H$; the interest of this example lies in the fact that there is favourable selection in this market, but private information is actually *unidimensional*.[6]

It is important to stress that lacking a sound and universally accepted equilibrium insurance model under multidimensional unobserved heterogeneity, the examples above are only suggestive and just sketch some empirically or theoretically relevant cases. However, these examples do point at possible problems with the positive correlation test under multidimensional heterogeneity, which stem from the fact that using a single statistic to calculate correlation is bound to average out selection effects which pull in different directions.

3.3 Testing for selection effects with multidimensional private information: The FMP test

In an important stream of papers dealing with selection effects in insurance markets with multidimensional private information, Finkelstein and Poterba [57], Finkelstein and McGarry [58] and Cutler *et al.* [32] suggest using variables which are observable by the econometricians but not used by the insurance companies as *proxies* of private information variables. In practice, this involves adding such variables as covariates in the equations for the loss and insurance probabilities; if any of these proxies is found significant, this is taken as evidence of the existence of residual private information. As argued by Cutler *et al.* [32], the test can be performed both conditionally on insurers' risk classification $\boldsymbol{x}$ as suggested by Chiappori and Selanié [25], or unconditionally. They argue that "the unconditional relationships may be of greater interest, since we are primarily interested

[6]Of course this begs the question about the relationship between this example and those papers (e.g. Finkelstein and McGarry [58] and Fang *et al.* [53]) that, having found evidence for favourable selection in some insurance market, conclude that this implies the presence of *multidimensional* private information. The two views can be reconciled only if one assumes the existence of a further unobservable, namely actual risk R, which has a positive effect on both I and O. In other words, if one assumes that γ_I and γ_O are strictly positive, a necessary condition for finding negative correlation between ϵ_I and ϵ_O is that there must be some other unobservable with contrasting effect on I and O. Thus, having *assumed* the presence of one dimension of private information (actual risk), one concludes that private information must be multidimensional.

in how preferences mediate the insurance - risk occurrence relationship" ([32], p. 160), but in practice they find that in their applications the two sets of results are very similar. We refer to this testing procedure as the Finkelstein-McGarry-Poterba (FMP) test.

In a seminal application of this testing procedure to the long-term care insurance market, Finkelstein and McGarry [58] find that, after conditioning for risk classification by insurance companies, and proxing risk preference by seat belt use, preventive care activity, and wealth, all of these variables have a positive effect on the probability of buying insurance but a negative effect on nursing home use, which is taken as evidence of favourable selection in this market. In another application of the FMP test, Cutler *et al.* [32] consider five insurance markets (namely life, annuity, long-term care, Medigap, acute health) and concentrate in finding whether, extending the standard RS model of adverse selection, risk preference is a significant determinant of insurance choices and loss occurrence. As observable proxies for risk preference they use five variables (namely smoking, drinking, job risk, preventive care, seat belt); if, after controlling for $\boldsymbol{x}$, these variables are found to have the same (opposite) effect on the insurance and loss probabilities, this is interpreted as evidence of adverse (favourable) selection. Cutler *et al.* [32] find that individuals who engage in risky behaviours are less likely to buy insurance in all five markets considered. In addition, more risk tolerant individuals tend to have higher claims in life and long-term care insurance, but lower claims for annuities, while no systematic relationship is found in the Medigap and acute health insurance. They suggest that while in the annuity market the standard adverse selection model cannot be rejected, there is evidence for favourable selection in the life and long-term care insurance markets, and there is no concluding evidence on the nature of the selection effects in the Medigap and acute health markets.

3.4 Selection effects with multidimensional unobservable types

Suppose there is a set $V_1, \ldots, V_K$ of residual heterogeneity variables which affect insurance choice and outcomes after conditioning on $\boldsymbol{x}$. We assume that $V_1, \ldots, V_K$ are *discrete*, with V_k taking say l_k levels, $k = 1, \ldots, K$. This is a fairly innocuous assumption since any continuous variable can be approximated arbitrarily well by a discrete one, and it implies that we can cross-classify $V_1, \ldots, V_K$ into a single discrete variable which takes $l_1 \times \cdots \times l_K$ values, which identifies the set of heterogeneous "types". We do not make any assumption either on K or on each l_k; in practice, many different types are very close to each other and thus are empirically undistinguishable and can be lumped together. Let then $\mathcal{T} = \{1, 2, \ldots, M\}$ be the set of different unobserved heterogeneous types, and let T be a random variable with support in $\mathcal{T}$. Clearly since the label of the types is arbitrary, no order is assumed on T: for example, with two binary private information variables capturing risk type and risk preference, as shown by Smart [108] there is no ordering of the four types since single crossing of the indifference curves fails to hold. In this setting, different types are simply meant to capture heterogeneous insurance and claim behaviours after conditioning on $\boldsymbol{x}$, without any assumption on the underlying structure; what matters here is that a sufficient number of types is used to capture residual heterogeneity. Notice also that even if we could estimate insurance and loss probabilities for each type, without further assumptions one cannot disentangle the individual effect of each single unobservable characteristic.

Now, how can we define selection effects and test for multidimensional private information in the presence of a finite set of unobservable types? For the time being, consider again the analysis conditional on $\boldsymbol{x}$, and, to isolate selection effects, assume again no moral hazard:

$$P(O \mid T = t, I = 1) = P(O \mid T = t, I = 0), \ t \in \mathcal{T}. \tag{3.4}$$

Now, when private information is multidimensional it is well possible that an insurance contract is *both* adversely *and* favourably selected by different types. Thus, it makes sense to have both a *local* and a *global* definition of selection effects which naturally extends the unidimensional case:

Definition 5. (Local Selection): *The insurance contract I is adversely (favourably) selected by type $t \in \mathcal{T}$ if*

$$\big(P(O = 1 \mid T = t) - \bar{P_O}\big) \cdot \big(P(I = 1 \mid T = t) - \bar{P_I}\big) > (<) 0; \tag{3.5}$$

if an insurance contract is adversely (favourably) selected by *all* types $t \in \mathcal{T}$, since the labeling of the types is arbitrary, we have

Definition 6. (Global Selection:) *There is* global *adverse (favourable) selection* *when there is an appropriate rearrangement of the types such that*

$$\begin{array}{l} Pr(O = 1 \mid \boldsymbol{x}, T = 1) \leq \cdots \leq Pr(O = 1 \mid \boldsymbol{x}, T = M) \\ Pr(I = 1 \mid \boldsymbol{x}, T = 1) \leq (\geq) \cdots \leq (\geq) Pr(I = 1 \mid \boldsymbol{x}, T = M) \end{array} \tag{3.6}$$

with some inequality holding strictly.

Now, it is worth noting that if each private information variable $V_1, \ldots, V_K$ is *monotonic* in O and I, when a contract is globally selected we *cannot* exclude that private information is actually *unidimensional* (i.e. a globally selected contract under multidimensional private information is observationally equivalent to a unidimensional private information one). In other words, under the assumptions of no moral hazard and that outcome and insurance probabilities are monotone in each private information variable, we can *reject* the unidimensionality of private information only when the global selection inequalities do not hold, since if each V_k has a monotone effect, then there must be *at least two* unobservable characteristics which pull I or O in different directions. This observation suggests an empirical test for the unidimensionality of private information, when there are more than two types.

3.4.1 Implementation

In practice, since individuals differ by the set of variables used by insurers to price contracts, the analysis has to be performed conditional on $\boldsymbol{x}$. A key point is then the choice of $\boldsymbol{x}$. In many insurance markets there are many observable characteristics which *cannot* be used either by regulatory laws or political economy concerns. Thus, a natural choice is to use as conditioning variables those which are effectively used by insurers to price contracts so as to look at how people behave conditional on the menu of contracts they actually face. In highly competitive and unregulated markets, presumably the insurance companies' choice of $\boldsymbol{x}$ reflects more accurately actual insurance and loss probabilities, compared with heavily regulated markets. Therefore, in the latter markets it is more likely to find heterogeneous behavior of the types compared to the former (see e.g. Polborn *et al.* [99] on 'regulatory selection effects'). In this sense the unobserved heterogeneity T is implicitly captured by whatever affects O and I but is not used by insurance companies to price contracts.

The empirical implementation of the test requires identification of the insurance and outcome probabilities $Pr(I = 1 \mid \boldsymbol{x}, t)$ and $Pr(O = 1 \mid \boldsymbol{x}, t)$, and the possibility to estimate probabilities in terms of the unobserved types *for each vector $\boldsymbol{x}$ of insurance used controls*. The practical implementation of the test thus suggests the assumption that $Pr(I = 1 \mid \boldsymbol{x}, t)$ and $Pr(O = 1 \mid \boldsymbol{x}, t)$ are linearly additively separable in $\boldsymbol{x}$ and T, for example by assuming that

$$\begin{aligned} Pr(I = 1 \mid \boldsymbol{x}, t) &= F(\alpha_I(t) + \boldsymbol{x}'\boldsymbol{\beta}_I), \\ Pr(O = 1 \mid \boldsymbol{x}, t) &= F(\alpha_O(t) + \boldsymbol{x}'\boldsymbol{\beta}_O) \end{aligned} \tag{3.7}$$

for some appropriate link function F. Under this assumption, to analyze selection effects we need estimating the individual types effects $\alpha_I(t)$ and $\alpha_O(t)$ for $t \in \mathcal{T}$; appropriate equality and inequality restrictions on $\alpha_I(T)$ and $\alpha_O(T)$ can then be imposed to test for the absence of selection effects and the unidimensionality of private information.

3.5 Estimation

To analyze selection effects we need to estimate the parameters α's and β's in the nonlinear system (3.7), jointly with the membership probabilities $P(t)$, $t \in \mathcal{T}$. This can be accomplished by use of a *semiparametric finite mixture model*; this kind of models, which have become popular in economics after the seminal paper by Heckman and Singer [68], decompose the observed conditional joint distribution of O and I into a finite number of components with mixing probabilities $P(t)$:

$$P(O, I \mid \boldsymbol{x}) = \sum_{t \in \mathcal{T}} P(t) F(\alpha_O(t) + \boldsymbol{x}'\boldsymbol{\beta}_O) F(\alpha_I(t) + \boldsymbol{x}'\boldsymbol{\beta}_I).$$

Estimation of finite mixture models involves first a choice of an appropriate *link function F*; we use logit. Notice however that while relying on the parametric choice of F for modeling each component probability, no parametric structure is imposed on the unobserved heterogeneity variable T. Well known applications of semiparametric finite mixture models are Deb and Trivedi ([38] and [39]) in health economics and Cameron and Heckman [20] in education.

3.5.1 A discrete multiresponse finite mixture model

The semiparametric finite mixture model is well established and achieves non-parametric estimation of the unobserved heterogeneity parameters α's. However, conditionally on $\boldsymbol{x}$ we observe only two binary variables (that is, 3 parameters), while even with two mixture types there are 5 parameters to estimate (four α's and a mixture probability). Thus, it must rely on covariates' variation to achieve parameters' identification. In practice, it typically achieves identification of very few unobserved types; in many applications only two types are identified.

To achieve sharper identification of the heterogeneous types we may use a set of observable *indicators*, that is, observable manifestations of the unobserved heterogeneity which affects insurance choice and loss occurrence after conditioning on $\boldsymbol{x}$. Following the logic of the FMP testing procedure, appropriate indicators can be chosen as variables which are observable by the econometricians but not used by the insurance companies. Examples of variables which can be used as indicators include: wealth, cognitive abilities, occupational risk, risk reducing or increasing behavior such as preventive care, seat belt use, smoking and drinking or, if panel data are available, past insurance choices and claims. For simplicity, we assume that chosen indicators are actually binary variables (this restriction aims to simplify the notation, but our econometric analysis can be performed as long as these indicators are discrete).

Having found a set $Z_1, \ldots, Z_H$ of suitable indicators, the main model of interest is then augmented by an auxiliary system of conditional probabilities for the indicators

$$Pr(Z_h = 1 \mid t) = F(\alpha_{Z_h}(t)),\ h = 1, \ldots, H \tag{3.8}$$

which is instrumental for identifying the mixture components probabilities in (3.7), which are of primary interest. Notice that, since T defines mutually exclusive and exhaustive types, the auxiliary equation system (3.8) is saturated and the choice of the link function is purely one of convenience since no parametric assumption is imposed.[7]

[7] In the auxiliary system (3.8) one could also condition the probability of each indicator to the controls $\boldsymbol{x}$, in which case the unobservable types are defined relatively to $\boldsymbol{x}$. For example, if Z_h is the choice of wearing seat belts which acts as an indicator of risk preference, conditioning say on age, helps identifying risk attitudes relatively to age, while if no conditioning is made, one tends to identify unadjusted risk preferences. In our experience, for the purpose of estimating the parameters of the main system of interest, in practice there is little difference in the results

The discrete multiresponse finite mixture model is completed by the types membership probabilities $P(t)$. To force the types probabilities to lie between zero and one and sum to one, it is convenient use a multinomial logit parameterization:

$$Pr(T = t) = \frac{\exp(\alpha_T(t))}{\sum_{t=1}^{M} \exp(\alpha_T(t))}, \;\; \alpha_T(M) = 0 \tag{3.9}$$

so that the $M - 1$ logit parameters α_T are simply reparametrization of the membership probabilities, and do not impose any parametric restriction on the distribution of T.

The discrete multiresponse finite mixture model is defined by equations (3.7)-(3.8)-(3.9), with α's and β's being the model parameters. Model (3.7)-(3.8)-(3.9) can be seen as an instance of a discrete multivariate MIMIC (Joreskog and Goldberger[81]) model (see Goodman [63] for the seminal paper on finite mixture models with multivariate binary responses, and Huang and Bandeen-Roche [75] for a recent general treatment), which uses information on $\boldsymbol{x}$ and the observable joint distribution of the response variables $[I, O, Z_1, \ldots, Z_H] \equiv \boldsymbol{Y}$ to learn some relevant features of the unobservable conditional distribution $P(\boldsymbol{Y} \mid \boldsymbol{x}, T)$. Contrary to the MIMIC model, the unobserved heterogeneity T is not a continuous univariate variable on the real line, but an unstructured nonparametric variable.[8]

When F is the standard binary logit link, we can rewrite the model (3.7)-(3.8)-(3.9) in terms of M logits for I, O and each of the indicators Z_h, and $M - 1$ multinomial logits for the type membership probabilities. Thus, under the logit link, model (3.7)-(3.8)-(3.9) can be written more compactly as

$$\boldsymbol{\lambda}(\boldsymbol{x}) = \boldsymbol{B}(\boldsymbol{x})\boldsymbol{\psi} \tag{3.10}$$

where $\boldsymbol{\lambda}$ is the vector which collects the $(S + 1) \times M - 1$ logits, $\boldsymbol{B}(\boldsymbol{x})$ is a design matrix whose dependence on $\boldsymbol{x}$ reflects the dependence of the conditional distri-

obtained under the two approaches.

[8] For an intuitive explanation on how this can be done, let $S = H + 2$ denote the size of $\boldsymbol{Y}$, and consider the analysis conditional on a given value of $\boldsymbol{x}$. Conditionally on $\boldsymbol{x}$, we observe the joint distribution of $\boldsymbol{Y}$, which has $2^S - 1$ free parameters, but (3.7)-(3.8)-(3.9) contains $(S+1) \times M - 1$ parameters. Therefore, if a sufficient number of indicators are available, a large number of types can in principle be identified. For example, if $H = 4$, there are 63 observable parameters and $7 \times M - 1$ parameters to estimate, so that in principle up to 9 different types can be identified. However, this is only a necessary condition for the identification of the unobservable parameters; there are well known pathological examples in the literature on finite mixture models which show that this counting condition is not sufficient.

bution of I and O on $\boldsymbol{x}$, and $\boldsymbol{\psi}$ is the vector which collects the model parameters α's and β's. Now, it can be easily seen that α's are the only model parameters after conditioning for a given value of $\boldsymbol{x}$. However, since the α's are one-to-one and differentiable functions (i.e. reparametrizations) of the probabilities of interest, model (3.10) is actually nonparametric conditional on a given value of $\boldsymbol{x}$. It follows that when the variables in $\boldsymbol{x}$ are discrete, take a limited number of distinct configurations (strata), and sufficient observations are available for each stratum, in principle one can analyze selection effects in the discrete multiresponses finite mixture model in a nonparametric fashion, analogously to the PC test discussed in section 3.2 above. In particular, notice that the unconditional model is nonparametric, and that model (3.10) imposes the same parametric restriction to the data as the PC test -namely, the choice of the link function F in the two equations (3.7)- although nonparametric estimation of our model requires more observations in each strata.

Estimation of the model parameters in (3.10) can be obtained by the EM algorithm, which is the standard approach for maximum likelihood estimation of finite mixture models, and has been shown (Dempster *et al.* [40]) to converge to the maximum of the true likelihood. Given the binary nature of the response variables, the E-step is equivalent to compute, for each subject, the posterior probability of belonging to each unobservable type. The M-step requires maximization of a multinomial likelihood with individual covariates, with a suitable modification to allow for linear inequality constraints.[9] It is well known that the EM algorithm may converge even if the model is not identified, a crucial issue for finite mixture models. Conditions for parametric identification of model (3.10) are discussed in Theorem 1 of Huang and Bandeen Roche [75]; local identification in the nonparametric model can be obtained using the the numerical test described by Forcina [59], which consists in checking that the Jacobian of the transformation between the parameters of the observable responses and the mixture model parameters α's is of full rank for a wide range of parameter values.

Within model (3.10), the absence of selection effects and the multidimensionality of private information can be formally tested by imposing appropriate restrictions on the α parameters. For example, one rejects the null of the absence of selection effects when each of the equality constraints $\alpha_I(1) = \cdots = \alpha_I(M)$ and $\alpha_O(1) = \cdots = \alpha_O(M)$ is rejected. On the other hand, since finite mixture

[9]We are grateful to Antonio Forcina for kindly providing the Matlab code for the EM estimation.

models are invariant to types' permutation, the standard RS insurance model (i.e. the null of global adverse selection) for example can be tested by setting the inequalities $\alpha_I(1) \leq \cdots \leq \alpha_I(M)$ and $\alpha_O(1) \leq \cdots \leq \alpha_O(M)$; techniques of order restricted inference can be used to show that the LR test statistic is asymptotically distributed as a mixture of chi-squared distributions.[10]

Despite the usefulness of finite mixture models to detect underlying residual heterogeneity, one unresolved issue in their application is how to determine the number of unobserved types M. The currently preferred approach suggests the use of Schwartz's Bayesian Information Criterion (BIC) to guide this choice, which in certain conditions is known to be consistent and generally helps preventing overparametrization (see McLachlan and Peel [93] for a thorough introduction to finite mixture models and a review of existing criteria for the choice of the number of types). BIC is calculated from the maximized log-likelihood $L(\boldsymbol{\psi})$ by penalizing parameters' proliferation, $BIC(\boldsymbol{\psi}) = -2L(\boldsymbol{\psi}) + \upsilon log(n)$, where n denotes sample size and υ the number of parameters; the model with the lowest BIC is preferred.

Finally, it may be worth noticing that, since the number of types M is not predetermined, formal hypotheses tests performed in finite mixture models are in fact conditional on M, and pre-testing for the number of types may invalidate distributional results of the test statistics employed. Of course, while pre-testing is a common problem in most applied research whenever final estimates are obtained after searching for appropriate specification, this is an issue which should be kept in mind whenever test results differ significantly when performed under different values of M.

3.5.2 Relationship with FMP test

The multivariate discrete finite mixture model uses the same data as the FMP test, namely the response variables of main interest I and O, the set of variables used by insurers to price contracts $\boldsymbol{x}$, and the set of auxiliary observable variables Z_h. The main difference is that while the variables Z_h are used as *proxies* of residual private information by FMP, they act as *indicators* of the unobserved types in our model. While FMP test can appropriately detect the existence of private information in a simple and robust procedure, it may run into difficulties when

[10] See Gourieroux and Monfort [64] for a general exposition, and Dardanoni and Forcina [34] for a discussion on how the mixing weights can be calculated by simulations.

one tries to interpret the results in terms of selection effects and multidimensional private information. Our procedure, on the other hand, by clearly identifying a finite number of unobservable types, allows a precise interpretation of the nature of selection mechanisms in the insurance contract. Notice that the finite mixture model, while capturing relevant residual heterogeneity in a parsimonious and direct way, does not come at a cost of stronger parametric assumptions imposed on the data.[11]

3.5.3 Multiple outcomes

In many circumstances the insurance contract protects against multiple losses. For example, Medigap protects against high out of pocket expenses for several health care services, such as inpatient, outpatient and specialist visits. The framework above can then be extended by simultaneously considering say J binary outcomes O_j, $j = 1, \ldots, J$, which take value one if the individual experiences the loss of type j for which he is insured. Assuming again linear additive separability (which is key for implementing the proposed approach) the conditional probabilities of interest are

$$\begin{aligned} &Pr(I = 1 \mid \boldsymbol{x}, t) = F(\alpha_I(t) + \boldsymbol{x}'\boldsymbol{\beta}_I) \\ &Pr(O_j = 1 \mid I, \boldsymbol{x}, t) = F(\alpha_{O_j}(t) + \boldsymbol{x}'\boldsymbol{\beta}_{O_j}),\ j = 1, \ldots, J \end{aligned} \tag{3.11}$$

for a suitable link function F. For sharper identification of the mixture components, this system of equations of main interest is integrated by the auxiliary system, and the complete model is (3.8)-(3.9)-(3.11).

3.6 Application to the US long-term care insurance market

In a recent seminal paper Finkelstein and McGarry [58] (henceforth FM) study the long-term care insurance market in the USA. Long-term care expenditure

[11] In particular, the main parametric restriction imposed to the data by both approaches is the choice of the link function F in the I and O equations. However, if the test is performed unconditionally as suggested by Cutler *et al.* [32], the finite mixture model does not impose parametric restrictions, contrary to the FMP test where parameters' estimates of $P(O \mid Z_1, \ldots, Z_H)$ and $P(I \mid Z_1, \ldots, Z_H)$ still depend on the choice of a link function, if a linear separability assumptions is imposed to help interpretation of results.

risk is one the greatest financial risks faced by the elderly in the US; to get a quantitative feeling of its importance, the amount of expenditure in nursing home care in 2004 was about 1.2% of the US GDP. Furthermore, as argued by FM, long-term care insurance is a good market to study since it is not heavily regulated. Their data comes from the Health and Retirement Study (HRS); the average age of respondent is 77.

FM notice that in the sample there is *negative* correlation between insurance purchase and nursing home use; they also perform the PC test conditioning on risk classification by calculating, by means of a standard actuarial model,[12] the probability of nursing home use as estimated by insurance companies. They find no conditional correlation when they apply the test to the whole sample, and slightly significant negative conditional correlation when the test is applied to a more homogenous subsample of individuals who are likely to face the same menu of options.

Their overall interpretation of these results is that individuals may have private information not only on their risk type, but also on their preferences for insurance coverage, which operate in offsetting directions. They show that individuals who exhibit more cautious behavior - as measured either by their investment in preventive health care or by seat belt use are both more likely to have long-term care insurance coverage and less likely to use long-term care, so they conclude that the market is favorably selected. Since empirical evidence suggests that demand for nursing home use is relatively price inelastic, FM suspect that their results most likely reflect ex ante more than ex post private information.

3.6.1 Data and variables definition

We implement our testing procedure by using FM dataset as reported in table 4 of their paper (FM [58] pg. 948). This is a subsample of individuals in the top quartile of the wealth and income distribution without any health characteristics that might make them ineligible for long-term care insurance.[13] We use as insurance purchase and risk occurrence two binary variables, namely *Long-Term Care Insurance* which takes value one if the individual has long-term care insurance,

[12] They alternatively use as controls a rich set of covariates typically used by insurance companies to price contracts, but their results do not change significantly.

[13] FM's dataset is available in the AER website. We thank FM, and the AEA for their policy of providing data for published articles.

and *Nursing Home* which takes value one if the individual enters a nursing home in the following 5 years. In this sample about 17% of individuals have long-term care insurance in 1995 and 10% enter a nursing home in the following 5 years period. As observed characteristics used by insurance company ($\boldsymbol{x}$) we use the probability of entering a nursing home, which is calculated by FM from a standard actuarial model. We create 10 risk categories by considering deciles, so $\boldsymbol{x}$ is actually a vector of 9 dummies since we exclude the 5th decile.

As indicators for the residual unobserved heterogeneity we use the following binary variables: *Seat Belt* which takes value 1 if the subject always wears seat belts; *No Smoking and Drinking* which takes the value 1 if the subject either currently does not smoke or has less than three drinks per day; *Subjective Riskiness* which takes value 1 if the individual self-reported probability of nursing home utilization is higher than the insurance company estimated probability; *Preventive Care* which takes value 1 if the subject has taken more gender appropriate preventive care procedures in the past year than the median value.[14]

3.6.2 Results

We first attempted estimation of a standard semiparametric finite mixture model. However, in this sample the standard model does not robustly identify residual heterogeneity, since even with only two types the information matrix is badly conditioned and mixture parameters have very high standard errors. We thus estimate the extended model using the 4 indicators to set the auxiliary system (3.8). To choose the number of mixture types we use Schwartz's BIC, which achieves the minimum value with two mixture types. The model we estimate and report has thus 31 parameters: 18 regression coefficients β, 12 α's for the six responses and the two types, and 1 parameter for the marginal probability of T. For completeness, estimates of α's and β's and their standard errors are reported in Appendix B, but for economy of space estimated coefficients are not discussed in the main text. About 66% of individuals are of type 1, and 34% of type 2.

Table 3.1 below reports the conditional probabilities by types for the six responses.[15] Conditionally on $\boldsymbol{x}$ there seems to be a substantial difference in in-

[14] Gender appropriate preventive care is a discrete variable derived from a list of possible preventive activities. The median individual undertakes about 80% of these activities.

[15] The conditional probability for nursing home use and long-term care insurance are averaged across insurance risk classification.

surance purchase and nursing home use between the two types; type 1 individuals are 4 times more likely to buy a long-term care insurance, but almost 3 times less likely to use a nursing home than types 2. Thus the table hints at the presence of residual heterogeneity, and in particular at favourable selection.[16]

To understand the differences between the two types identified in this sample, it is instructive to look again at table 3.1: there seems to be a natural ordering of the types in terms of their cautiousness, such that, going from types 2 to types 1, there is a significant increase in the probability of using seat belts and preventive care, of refraining from smoking and drinking, and believing that one may need a nursing home in the near future with a higher probability than that predicted by insurance companies.

Table 3.1: Estimated conditional probabilities

	T=1	T=2
Seat Belt	0.9379	0.6361
Subjective Riskiness	0.5002	0.3523
Preventive Care	0.4567	0.2785
No Smoking and Drinking	0.9528	0.8034
Long-Term Care Insurance	0.2335	0.0568
Nursing Home	0.0667	0.1657

From the estimated joint distribution of the indicators and the unobserved types we can also get an estimate of the so called *posterior type probabilities* for some focal observable individual behaviour. Let $\tilde{\boldsymbol{Y}}$ be a vector of observable indicators of focal interest, then from the estimated joint distribution $P(T, \tilde{\boldsymbol{Y}})$, posterior type probabilities are obtained using

$$P(T = t \mid \tilde{\boldsymbol{Y}} = \tilde{\boldsymbol{y}}) = \frac{P(T = t, \tilde{\boldsymbol{Y}} = \tilde{\boldsymbol{y}})}{P(\tilde{\boldsymbol{Y}} = \tilde{\boldsymbol{y}})}.$$

We consider two focal behaviours: a 'cautious' individual who always wears a seat belt, does not smoke and drink, has a cautious estimate of his probability of

[16]The question naturally arises whether these differences in insurance choice and nursing home entry are simply due to sampling variation. The LR test statistic for the null of no residual heterogeneity in insurance choice (that is, $\alpha_{LTCI}(T = 1) = \alpha_{LTCI}(T = 2)$) and loss occurrence ($\alpha_{NH}(T = 1) = \alpha_{NH}(T = 2)$) are equal to 17.56 and 7.99, which are asymptotically distributed as a χ^2 with one d.f., overwhelmingly rejecting the null respectively with p-values lower than 10^{-5} and .005. For comparison, in the standard PC test (see Table 4, column 2 of Finkelstein and McGarry [58]), the hypothesis of the absence of private information has a p-value of .10.

needing a nursing home in the future and engages in preventive care, and a 'reckless' individual with the opposite attitudes.[17] Table 3.2 reports the estimated posterior type probabilities for these two individuals. A glance at the table confirms the presumption that types 1 are predominantly the 'cautious' types, while types 2 are the 'reckless' ones.

Table 3.2: Estimated posterior type probabilities

	T=1	T=2
Cautious	0.8872	0.1128
Reckless	0.0439	0.9561

Overall, our analysis seems to confirm Finkelstein and McGarry's [58] message that the standard unidimensional RS adverse selection model does not hold in this insurance market, with a little twist: at least in this sample, we cannot exclude the possibility that there is a *single* unidimensional private information variable (say 'cautiousness') which drives the favourable selection mechanism. In other words, the data are compatible with the possibility that there is underlying negative correlation between cautiousness and actual risk, but individuals do not act upon (or are ignorant about) such correlation. In the end, more cautious individuals would be taking insurance and at the same time present less risky outcomes.

3.6.3 The unconditional model

Table 3.3 below reports the estimated probabilities by types without conditioning on $\boldsymbol{x}$, and shows a broad agreement with the estimated probabilities in Table 3.1. Unconditional probabilities may be of independent interest, and are obtained in a saturated nonparametric model, so that they help checking the robustness of our results.[18]

[17] Thus $\tilde{\boldsymbol{Y}}$ = [Seat Belt, No Smoking and Drinking, Subjective Riskiness, Preventive Care].

[18] Table B.3 in the Appendix reports estimated α coefficients and standard errors; recall that in this saturated model the α's are simply reparametrizations of the conditional probabilities, so no parametric restriction is imposed.

Table 3.3: Estimated probabilities in the unconditional model

	T=1	T=2
Seat Belt	0.9369	0.6589
Subjective Riskiness	0.5138	0.3392
Preventive Care	0.4653	0.2762
No Smoking and Drinking	0.9486	0.8211
Long-Term Care Insurance	0.2462	0.0526
Nursing Home	0.0587	0.1664

3.7 Application to the US medigap insurance market

Our second application focuses on selection effects in the "Medigap" insurance market for elder individuals. A Medigap insurance plan is a health insurance contract sold by a private company to fill "gaps" in coverage of the basic Medicare plan. Medigap plans offer additional services and help beneficiaries pay health care costs (deductibles and co-payment) that the original Medicare plan does not cover, so that health care costs of Medigap enrollees is covered by both plans. As a relevant example for our application, Medicare's coinsurance or copayments for hospital stays, physician visits or outpatient care are covered by the Medigap plan.

The Medigap insurance market is quite interesting to study as a further application of our methods since, contrary to the long-term care market, it is highly regulated. In fact, Federal Law affects the Medigap market at least in three ways. First, Medigap plans are standardized into ten plans,"A" through "J", and the basic plan "A" must be offered if any other more generous plan is also offered. Second, there is a free enrolment period which lasts for six months from the first month in which people are both 65 years old and enrolled in Medicare. During this period Medigap cannot refuse any person even if there are pre-existing conditions. Third, pricing criteria are mainly based on individual's age and sex. Therefore, insurers are not free to offer any insurance contract at any price they choose (see e.g. Cutler *et al.* [32], Fang *et al.* [53] and Finkelstein [54], for theoretical and applied analysis of selection in the Medigap market).

In a recent influential paper, Fang *et al.* [53] consider private information in the Medigap market using data obtained by imputing HRS observations for year 2002 in the Medicare Current Beneficiary Survey. They find a negative correlation between Medigap supplemental coverage and ex post medical expenditure, and argue that individual cognitive ability is the main source driving favourable selection in the Medigap market.

3.7.1 Data and variables definition

We use data from waves 4-6 of the HRS, which covers respectively the following years: 1998, 2000 and 2002. This data set contains detailed information on individual insurance status and sources of supplemental coverage. Our main focus is to study selection effects in the last period (2002), but we exploit the panel nature of the dataset using past insurance and utilization decisions to help identify residual heterogeneity. Following Fang *et al.* [53] we define Medigap status (*Insurance 98*, *Insurance 00*, *Insurance 02*) to be equal to one if an individual is covered by Medicare and has deliberately purchased a supplemental plan additional to Medicare. Therefore we excluded from the dataset individuals who were younger than 65 years at 1998 and are also enrolled in any other public program different from Medicare or receive Medigap insurance coverage by his/her or spouse's former employer. Risk occurrence is measured by the following binary variables which take 1 if an individual: i) had any hospital stay (*Hospital 98*, *Hospital 00*, *Hospital 02*); ii) had more than five doctor visits (*Doctor 98*, *Doctor 00*, *Doctor 02*); iii) used any outpatient service such as surgery or home care facilities in the twelve months prior to the interview (*Outpatient 98*, *Outpatient 00*, *Outpatient 02*). As additional indicators to identify unobserved types we also use *Subjective Health*, which equals 1 if the individual reports good or very good health; *Wealth* which equals 1 if the individual is in the top wealth quartile; *Cognitive Skills* which takes value 1 if individual's performance is equal to or greater than the median score in the Telephone Interview for Cognitive Status (TICS); *Risk Tolerance* which takes value 1 if individual is less risk adverse than the median individual.[19] As observed characteristics used by insurance companies ($\boldsymbol{x}$) we use gender and 11 age dummies. Our sample is composed of individuals aged 65 or more in 1998. There are 1,231 observations in the sample; the median individual is 71 years old in 2002, and there are about 45% of females.

[19] This indicator of financial risk tolerance is based on a set of risk aversion measures provided by Kimball *et al.* [87], who estimate risk aversion for each respondent using a set of hypothetical income gambles, and has been also used by Fang *et al.* [53] in their analysis of Medigap insurance market.

3.7.2 Results

We first attempted estimation of the main model of interest (Medigap and the three medical care utilization variables in 2002) by means of a standard semi-parametric finite mixture model. In this sample the standard model does not robustly identify residual heterogeneity, since even with only two types the information matrix is badly conditioned and mixture parameters have very high standard errors. We then estimated the extended model using the 12 indicators discussed above to set the auxiliary system (3.8).

Table 3.4 reports the maximized log-likelihood and the BIC for different numbers of unobserved types. The BIC seems to indicate that four types are adequate to represent any residual unobserved heterogeneity. The model we estimate and report has four heterogeneous types,[20] and has 115 parameters: 48 regression coefficients β's, 64 α's for the 16 responses and the four types, and 3 parameters for the marginal probability of T. Again, for completeness we report estimates of α's and β's and their standard errors in Appendix B, but for economy of space estimated coefficients are not discussed in the main text.

Calculating the types membership probabilities, about 35% of individuals are of type 1, 38% of type 2, 17% of type 3 and 10% of type 4.

Table 3.4: Log-likelihood and BIC

	2LC	3LC	4LC	5LC
Log-Likelihood	-11563.60	-11255.03	-11144.18	-11104.86
BIC	23703.56	23207.40	23106.65	23148.99
# of parameters	81	98	115	132

Estimated coefficients are reported in Appendix B. Table 3.5 below reports the estimated conditional probabilities by types for the four variables of interest and the twelve auxiliary indicators. Conditionally on age and gender, there seems to be a striking heterogeneity in Medigap purchase decisions and loss occurrence across different types: types 3 and 4 are roughly ten times more likely to buy supplemental insurance than types 1 and 2, while types 1 and 4 are on average

[20] Estimated parameters for the other cases are available from the authors. The general picture emerging under three and five types is similar to the results discussed below. For robustness we also estimated the model controlling for age and sex also in the auxiliary equation system, without any significant difference in the conclusions.

two or three times more likely to use medical resources than types 2 and 3.[21]

Looking at Panel B in table 3.5, which reports conditional probabilities for the auxiliary indicators, we see that types 3 and 4 have much stronger preference for supplementary insurance and are more likely to be wealthier and with higher cognitive abilities; on the other hand, types 1 and 4 are much heavier users of medical care and report lower subjective health status. Financial risk tolerance does not seem to vary across types.[22] From the estimated joint distribution of the observed responses and T we can again calculate the posterior type probabilities for some focal individuals. Table 3.6 below reports the conditional probability of being a type $T = t$ for individuals who are: *i*) high risk (have used all the three types of medical care in previous periods, and reported low subjective health); *ii*) low risk (have not used any of the three types of medical care in previous periods, and reported high subjective health); *iii*) high risk preference (have not bought supplementary insurance in any previous periods, and have low wealth and cognitive skills); *iv*) low risk preference (have bought supplementary insurance in both previous periods, and have high wealth and cognitive skills).

[21] Since the differences in behaviour across types are so dramatic, it is very unlikely they are due to sample variations; in fact a LR tests for the equality of insurance purchase and loss occurrence across types rejects the null in each case with p-value lower than 10^{-6}. Regarding the possible effect of moral hazard on our estimates, while there is some evidence that the demand for hospital care is price inelastic (see e.g. Manning *et al.* [91]), there is no concluding evidence on the effect of supplementary insurance on the other medical care uses (see e.g. Ettner [51]). Given the magnitude of the effects we find in the data, and the similar pattern found in all the three types of use, it is unlikely that accounting for moral hazard might change significantly our conclusions.

[22] Fang *et al.* [53] report that cognitive ability is one of elements driving the selection into supplemental insurance, but financial risk tolerance does not have a significant effect.

Table 3.5: Estimated conditional probabilities

	T=1	T=2	T=3	T=4
Panel A: Main Equations				
Hospital 02	0.4200	0.1544	0.1975	0.5875
Doctor 02	0.7931	0.2791	0.2862	0.8729
Out. Services 02	0.3987	0.1689	0.2216	0.4318
Insurance 02	0.0634	0.0917	0.7113	0.8112
Panel B: Auxiliary Indicators				
Hospital 98	0.3550	0.1103	0.1310	0.4354
Doctor 98	0.7540	0.2145	0.2957	0.8451
Out. Services 98	0.2868	0.0831	0.1514	0.3546
Insurance 98	0.1038	0.1214	0.7916	0.8271
Hospital 00	0.3757	0.0867	0.1378	0.4326
Doctor 00	0.8627	0.2333	0.2727	0.8684
Out. Services 00	0.3337	0.1276	0.2089	0.3377
Insurance 00	0.0436	0.0538	0.8526	0.8691
Sub. Health	0.2771	0.5428	0.6732	0.1677
Cog. Skills	0.5947	0.6446	0.7228	0.7177
Wealth	0.2889	0.2681	0.4291	0.4365
Risk Tolerance	0.3721	0.3388	0.3794	0.3238

Estimated probabilities in the main system are averaged out for $\boldsymbol{x}$.

Notice that by risk preference we do not refer simply to financial risk aversion in the usual Arrow-Pratt sense, but also to the perception of the risk related to the specificity of the health insurance market. In a recent paper, Einav *et al.* [49] study five insurance coverage decisions (health, prescription drugs, dental, and short-term and long-term disability) to investigate how well an individual's willingness to bear risk in one context predicts his willingness to bear risk in other contexts. They find that among the five contexts considered, the magnitude of the domain-general component of preferences appears substantial; however, they also find higher correlation in choices that are closer in context (e.g health insurance and disability insurance), indicating the existence of non-trivial context specificity. Einav *et al.* [49] results support the choice of past insurance choices in modelling individual risk perception.

Table 3.6 suggests the existence of heterogenous types, with low and high actual risk and low and high preference for risk, which can be cross-classified as *high,high* (types $t = 1$), *low,high* (types $t = 2$), *low,low* (types $t = 3$) and *high,low* (types $t = 4$), with striking differences in estimated insurance and medical care choices.

Table 3.6: Estimated posterior probabilities

	T=1	T=2	T=3	T=4
High risk	0.6054	0.0003	0.0007	0.3937
Low risk	0.0057	0.7305	0.2634	0.0004
High risk preference	0.4996	0.4908	0.0070	0.0026
Low risk preference	0.0055	0.0085	0.5797	0.4063

The overall picture which emerges from our estimates is strongly suggestive of multidimensional residual private information, with strong *local* selection in the Medigap insurance market. Comparing with the average insurance and loss probabilities, the Medigap contract seems to be favourably selected by types 1 and 3, and adversely selected by types 2 and 4;[23] in other words the contract is favourably selected by individuals who have the same (high or low) propensity for insurance purchase and health care use, and adversely selected by individuals who have opposite attitudes. Thus, our results show the existence of heavy cross subsidization of some types at the expense of others. Notice that the selection effects we find should not necessarily be interpreted solely as evidence of asymmetric information.

On the other hand, after estimating a multivariate probit model for the four response variables of interest, the PC test in our sample shows a correlation coefficient equal to -0.051 (p-value .29) between insurance and doctor visits; 0.092 (p-value .07) between insurance and hospital visits; and 0.027 (p-value .59) between insurance and outpatient visits.[24] Thus the PC test in this sample suggests that there is no significant residual heterogeneity in the Medigap insurance market, in stark contrast with our results. This suggests that the PC test may run into serious problems in detecting selection effects when private information is multidimensional, since it appears to simply average out local selection effects. However, a refined version of the PC test of Fang *al.* [53] which compares claims for individuals with and without Medigap insurance under different conditioning sets, using proxies for propensity to use health care and buy insurance, paints a

[23] A glance at the estimated conditional probabilities in Table 3.5 suffices to realize that the standard RS model of adverse selection clearly does not hold, since it requires that insurance choice and medical care use probabilities have the same ordering across types. Formally, the LR test statistic for the null of global adverse selection is equal to 124.08; the conservative 1% critical value (Kodde and Palm ([88], page 1246) is equal to 32.196.

[24] The LR test statistic for conditional independence of insurance purchase and the three medical care use variables is equal to 4.98, and is asymptotically distributed as a chi-squared with three d.f. with p-value .17.

broad picture which does not contradict our results, since they find both adverse and favourable selection depending on the conditioning set used to control for unobservables. Our finite mixture model clarifies precisely who are the types who adversely and favourably select Medigap insurance, by simultaneously identifying the behaviours of unobservable types in both insurance purchase and claims.[25]

Table 3.7: FMP's Testing Procedure

Variables	Hospital 02		Doctor 02		Out.Serv.02		Ins. 02	
fem	0.0186	(0.0888)	0.0136	(0.0845)	-0.0289	(0.0870)	0.189	(0.0993)
age02==69	0.225	(0.394)	0.152	(0.319)	0.245	(0.389)	0.174	(0.303)
age02==70	0.256	(0.391)	0.143	(0.316)	0.258	(0.386)	0.104	(0.301)
age02==71	0.411	(0.396)	0.184	(0.322)	0.485	(0.389)	-0.105	(0.311)
age02==72	0.518	(0.414)	0.209	(0.338)	0.0903	(0.408)	0.193	(0.342)
age02==73	0.283	(0.425)	0.135	(0.354)	0.241	(0.416)	0.237	(0.339)
age02==74	0.231	(0.425)	0.110	(0.351)	0.0846	(0.419)	0.204	(0.355)
age02==75	0.577	(0.438)	0.194	(0.375)	0.211	(0.435)	0.160	(0.365)
age02==76	0.187	(0.491)	0.289	(0.410)	0.494	(0.463)	-0.414	(0.449)
age02==77	0.526	(0.493)	-0.459	(0.428)	0.348	(0.480)	0.197	(0.421)
age02==78	0.698	(0.534)	-0.0502	(0.509)	0.558	(0.523)	0.320	(0.458)
age02>=79	0.695	(0.470)	0.433	(0.423)	0.380	(0.459)	-0.287	(0.419)
Sub. Health	-0.424	(0.0856)	-0.541	(0.0835)	-0.170	(0.0843)	0.00488	(0.0955)
Cog. Skills	-0.0297	(0.0389)	0.0240	(0.0547)	0.0130	(0.0405)	0.0350	(0.0490)
Wealth	-0.139	(0.0923)	0.219	(0.0853)	0.245	(0.0863)	0.106	(0.0977)
Risk Tol.	-0.0915	(0.0841)	0.0999	(0.0803)	0.138	(0.0813)	0.110	(0.0934)
Hospital 98	0.186	(0.0983)	0.0319	(0.103)	-0.0220	(0.0991)	-0.0326	(0.111)
Doctor 98	0.00676	(0.00372)	0.0285	(0.00994)	0.00426	(0.00437)	0.00349	(0.00378)
Out.Serv.98	0.191	(0.102)	0.155	(0.103)	0.298	(0.0997)	0.0146	(0.118)
Ins. 98	0.0788	(0.105)	-0.0156	(0.106)	-0.0285	(0.101)	0.699	(0.103)
Hospital 00	0.486	(0.0952)	0.168	(0.0994)	-0.0617	(0.0994)	0.125	(0.109)
Doctor 00	-0.000177	(0.00431)	0.0414	(0.0116)	0.0114	(0.00437)	-0.00942	(0.00542)
Out.Serv.00	0.117	(0.0962)	-0.201	(0.0960)	0.505	(0.0920)	-0.0774	(0.106)
Ins.00	0.161	(0.108)	-0.0475	(0.108)	-0.00468	(0.104)	1.342	(0.105)
Constant	-0.772	(0.554)	-0.695	(0.602)	-1.328	(0.561)	-1.849	(0.562)

Standard errors in brackets.

Finally, we can compare our results with those obtained performing the FMP test suggested by Finkelstein and Poterba [56] and applied by Finkelstein and McGarry [58] and Cutler *et al.* [32], where the auxiliary indicators of unobservable types are used as proxies in the probit regressions for the four response variables of main interest. The estimated coefficients of the four probits are reported in Table 3.7 below. While many of the proxies used to detect residual

[25]Even though Fang *al.* [53] uses the HRS dataset, strictly speaking their results cannot be directly compared with ours since they use a different variable for risk occurrence (in particular, they use medical expenses rather than our three binary measures), and employ imputation techniques to derive it.

heterogeneity are significant in each of the four equations, which is a clear indication of the existence of private information in this Medigap market, their pattern does not give any clear indication of the nature of the selection effects; for example, no single proxy seems to be significant in both the insurance and the medical care equations.

3.7.3 The unconditional model

Along the same lines as the previous application, we finally estimate the unconditional (saturated) model. Table 3.8 reports the estimated conditional probabilities by types, and shows a broad agreement with the estimated probabilities in Table 3.5.[26]

Table 3.8: Estimated probabilities in the unconditional model

	T=1	T=2	T=3	T=4
Panel A: Main Equations				
Hospital 02	0.4034	0.1317	0.1701	0.5650
Doctor 02	0.8092	0.2937	0.2909	0.8806
Out. Services 02	0.4000	0.1674	0.2133	0.4294
Insurance 02	0.0725	0.1067	0.7483	0.8423
Panel B: Auxiliary Indicators				
Hospital 98	0.3567	0.1101	0.1300	0.4255
Doctor 98	0.7531	0.2157	0.2924	0.8384
Out. Services 98	0.2871	0.0835	0.1504	0.3498
Insurance 98	0.1040	0.1206	0.7952	0.8287
Hospital 00	0.3784	0.0857	0.1357	0.4249
Doctor 00	0.8617	0.2349	0.2667	0.8632
Out. Services 00	0.3345	0.1278	0.2095	0.3315
Insurance 00	0.0453	0.0535	0.8544	0.8680
Sub. Health	0.2768	0.5427	0.6764	0.1765
Cog. Skills	0.5941	0.6448	0.7241	0.7174
Wealth	0.2887	0.2696	0.4263	0.4372
Risk Tolerance	0.3717	0.3397	0.3781	0.3254

[26]For completeness Table B.6 in the Appendix reports the estimated α coefficients and their s.e.

3.8 Conclusion

In this paper we study how to detect selection effects in insurance markets under multidimensional private information. We first discuss how the standard PC test performs in this setting. Since insurance contracts may be both adversely and favourably selected by different individuals, the PC test -relying on a single statistic to appraise the risk-coverage correlation- may run into serious difficulties if there is more than one source of private information.

We show how multidimensional unobserved heterogeneity can be modeled using a finite number of heterogeneous types, and extend the standard adverse and favourable selection definitions into *local* and *global* ones. We propose a finite mixture model which allows estimation of the insurance and loss probabilities of the unobserved types, and explain how these can be used to analyze selection effects and test for the multidimensionality of private information.

We apply our procedure to the US long-term care and Medigap insurance markets. In both markets we find that there is significant evidence of residual heterogeneity, and that the standard Rothschild-Stiglitz adverse selection model is not supported by the data. In the long-term care insurance market, data are compatible with the existence of a single private information variable, namely cautiousness, which yields a negative risk-coverage correlation. In the Medigap market, data are compatible with the existence of two unobservable dimensions of private information, namely actual risk and risk preference, which yield very strong local adverse and favourable selection of Medigap insurance by different unobserved types. Notice that since on average these selection effects roughly offset each other, the PC test yields insignificant risk-coverage correlation even in the presence of huge residual heterogeneity in the data. Since the analysis is carried conditional on variables used by insurers to price risks in the underwriting process, it is not surprising that the extent of residual heterogeneity seems substantially larger in the Medigap market which is heavily regulated than in the rather competitive long-term care insurance market. Notice however that our results do not provide exhaustive evidence of the role provided by regulation in comparing the Medigap and the long-term care insurance markets. In particular the effect of regulation we capture is mainly due to the additional noise in private information due to limitation in the information gathered by insurer (see Einav *et al.* [47]). Clearly regulation's effects in these two markets may also involve several other dimensions - such as price and insurance coverage definition, etc. - that need further investigations. Another important source explaining the differences on results related to asymmetric information between the long-term care and the Medigap application is the differences in the structure of the insurance contract. In particular Medigap offers supplemental coverage for services that are already covered by Medicare, while long-term care insurance offers additional coverage that it is not provided by any other insurance. Thus, while in the long-term application risk preferences play the main role in defining the insurance purchase decision for additional services, it is possible that the part of private information due to adverse selection in the Medigap application is mainly related to moral hazard. In fact it could be possible that individuals are selecting insurance on the basis of their anticipated behavioural response to it. However this is another possible explanation of the differences we found in the dimensionality of private information that should be combined with the particular structure of the markets.

A couple of caveats are in order. First our approach, while providing useful descriptive evidence on the existence and the extent of selection effects under multidimensional private information in a given insurance market, does not explain the structural forces which determine insurance demand and market equilibrium, and thus is of limited direct use in appraising market efficiency and welfare effects of policy interventions. Some recent work in this direction is discussed by Einav *et al.* [47]. Second, as we discuss in the paper, moral hazard may limit the interpretation of the empirical results, so that care should be exercised when it is not reasonable to ignore incentive effects, or selection effects are marginal.

Appendix A

Proof of the Proposition

To show that 3 is equivalent to 4, rewrite 4 as $P(O=1, I=1) \cdot P(O=0, I=0) > P(O=1, I=0) \cdot P(O=0, I=1)$, substitute $P(O,I)$ with $P(O \mid I)P(I)$, then use $P(O=1 \mid I) = 1 - P(O=0 \mid I)$, and simplify; to show that 1 is equivalent to 2, substitute $\bar{P_O}$ and $\bar{P_I} = P(R=0)P(I=1 \mid R=0) + P(R=1)P(I=1 \mid R=1)$ in 2 and simplify.

To show the equivalence between 3 and 1 under Assumption (3.1), use the Law of Total probability

$$P(O \mid I) = P(R=0 \mid I) \cdot P(O \mid I, R=0) + P(R=1 \mid I) \cdot P(O \mid I, R=1)$$

and (3.1) to get

$$P(O \mid I) = P(R=0 \mid I) \cdot P(O \mid R=0) + P(R=1 \mid I) \cdot P(O \mid R=1).$$

Thus,

$$P(O=1 \mid I=1) - P(O=1 \mid I=0) =$$

$$\big(P(R=1 \mid I=1) - P(R=1 \mid I=0)\big) \cdot \big(P(O=1 \mid R=1) - P(O=1 \mid R=0)\big)$$

using $P(R=0 \mid I) = 1 - P(R=1 \mid I)$. By the same argument used above to show the equivalence of 3 and 4, it is easily seen that $P(R=1 \mid I=1) > P(R=1 \mid I=0)$ if and only if $P(I=1 \mid R=1) > P(I=1 \mid R=0)$.

Appendix B

Tables

Table B.1: Long-Term Care Insurance: Estimated α parameters

	$\alpha(T=1)$	$\alpha(T=2)$
Seat Belt	2.7118 (0.5392)	0.5543 (0.3222)
Subjective Riskness	-0.0005 (0.1074)	-0.6094 (0.1946)
Preventive Care	-0.1748 (0.1163)	-0.9530 (0.2321)
No Smoking and Drinking	3.0033 (0.4021)	1.4048 (0.2552)
Long-term Care Insurance	-1.4843 (0.2971)	-3.1169 (0.7182)
Nursing Home	-3.5377 (0.5043)	-2.4243 (0.4787)

Standard errors in brackets.

Table B.2: Long-Term Care Insurance: Estimated β Parameters

	LTCI	NH
Risk Classification 1	0.2606 (0.3419)	-0.7322 (0.6656)
Risk Classification 2	0.6510 (0.3318)	-0.2574 (0.5958)
Risk Classification 3	0.2460 (0.3510)	0.1542 (0.5616)
Risk Classification 4	0.3489 (0.3534)	0.2120 (0.5664)
Risk Classification 6	0.0624 (0.3637)	0.6418 (0.5261)
Risk Classification 7	0.3640 (0.3447)	1.0169 (0.4987)
Risk Classification 8	0.6142 (0.3355)	0.8207 (0.5067)
Risk Classification 9	0.3231 (0.3567)	1.6761 (0.4842)
Risk Classification 10	-0.0625 (0.3717)	2.2320 (0.4728)

Standard errors in brackets.

Table B.3: Long-Term Care Insurance: Estimated α parameters of the unconditional model

	$\alpha(T=1)$	$\alpha(T=2)$
Seat Belt	2.698 (0.5091)	0.6582 (0.2874)
Subjective Riskness	0.0554 (0.1194)	-0.667 (0.1999)
Preventive Care	-0.139 (0.1241)	-0.9634 (0.2265)
No Smoking and Drinking	2.9151 (0.3535)	1.5236 (0.232)
Long-term Care Insurance	-1.1189 (0.1607)	-2.8907 (0.7049)
Nursing Home	-2.7749 (0.2956)	-1.611 (0.223)

Standard errors in brackets.

Table B.4: Medigap: Estimated α parameters

	$\alpha(T=1)$	$\alpha(T=2)$	$\alpha(T=3)$	$\alpha(T=4)$
Sub. Health	-0.9589 (0.1233)	0.1718 (0.1055)	0.7226 (0.1789)	-1.6019 (0.3088)
Cog. Skills	0.3834 (0.108)	0.5952 (0.1077)	0.9585 (0.1765)	0.9329 (0.2278)
Wealth	-0.9007 (0.1174)	-1.0044 (0.1174)	-0.2853 (0.1588)	-0.2555 (0.2056)
Risk Tolerance	-0.5234 (0.1097)	-0.6688 (0.109)	-0.4923 (0.1614)	-0.7363 (0.219)
Hospital 98	-0.5971 (0.1118)	-2.0873 (0.1768)	-1.8925 (0.2437)	-0.26 (0.208)
Doctor 98	1.1202 (0.1417)	-1.2979 (0.1452)	-0.868 (0.1845)	1.6962 (0.3266)
Out. Services 98	-0.9111 (0.1172)	-2.4008 (0.2028)	-1.7239 (0.2229)	-0.5988 (0.2124)
Insurance 98	-2.1555 (0.1926)	-1.9794 (0.1832)	1.3346 (0.2367)	1.5653 (0.3115)
Hospital 00	-0.5078 (0.1115)	-2.3542 (0.2041)	-1.8336 (0.2366)	-0.2711 (0.2078)
Doctor 00	1.8375 (0.1977)	-1.1896 (0.1484)	-0.981 (0.1935)	1.887 (0.3602)
Out. Services 00	-0.6913 (0.1127)	-1.9222 (0.164)	-1.3317 (0.1941)	-0.6735 (0.2147)
Insurance 00	-3.0892 (0.3323)	-2.8669 (0.3262)	1.755 (0.3237)	1.8919 (0.3903)
Hospital 02	-1.0278 (0.6781)	-2.4702 (0.6909)	-2.1611 (0.7059)	-0.3128 (0.7009)
Doctor 02	1.2053 (0.6406)	-1.1684 (0.6372)	-1.1327 (0.6519)	1.8056 (0.7339)
Out. Services 02	-0.9376 (0.6622)	-2.1448 (0.671)	-1.8023 (0.6825)	-0.7974 (0.685)
Insurance 02	-2.8742 (0.8632)	-2.4718 (0.8518)	0.815 (0.8514)	1.3904 (0.8893)

Standard errors in brackets.

Table B.5: Medigap: Estimated β Parameters

	Hospital 02	Doctor 02	Out. Services 02	Insurance 02
age02==69	0.3875 (0.6812)	0.2801 (0.6352)	0.5392 (0.6644)	0.3020 (0.8409)
age02==70	0.4893 (0.6774)	0.2916 (0.6316)	0.5405 (0.6613)	0.1731 (0.8368)
age02==71	0.7077 (0.6842)	0.3753 (0.6407)	0.8979 (0.6669)	-0.3093 (0.8522)
age02==72	0.8861 (0.7126)	0.3170 (0.6763)	0.2818 (0.7025)	0.1765 (0.8989)
age02==73	0.4409 (0.7380)	0.2221 (0.6976)	0.5020 (0.7159)	0.3500 (0.9234)
age02==74	0.3765 (0.7382)	0.0316 (0.6969)	0.2312 (0.7213)	0.3417 (0.9228)
age02==75	1.0867 (0.7472)	0.3994 (0.7220)	0.4716 (0.7376)	0.3316 (0.9590)
age02==76	0.3848 (0.8337)	0.4893 (0.7947)	0.8507 (0.7854)	-0.8240 (1.0905)
age02==77	1.1896 (0.8379)	-0.9171 (0.8568)	0.6003 (0.8313)	0.3634 (1.1104)
age02==78	1.4694 (0.8933)	0.0013 (0.9209)	1.0484 (0.8726)	0.7073 (1.1789)
age02>=79	1.5162 (0.8019)	0.9322 (0.7986)	0.8482 (0.7890)	-0.4868 (1.0880)
fem	0.0280 (0.1528)	-0.0346 (0.1575)	-0.1234 (0.1451)	0.4043 (0.2117)

Standard errors in brackets.

Table B.6: Medigap: Estimated α parameters of the unconditional model

	$\alpha(T = 1)$	$\alpha(T = 2)$	$\alpha(T = 3)$	$\alpha(T = 4)$
Sub. Health	-0.9603 (0.1235)	0.1712 (0.1053)	0.7371 (0.1819)	-1.5402 (0.2979)
Cog. Skills	0.3809 (0.1081)	0.5964 (0.1075)	0.9648 (0.1789)	0.9315 (0.2252)
Wealth	-0.9017 (0.1176)	-0.9965 (0.117)	-0.2968 (0.1608)	-0.2524 (0.2033)
Risk Tolerance	-0.5249 (0.1098)	-0.6647 (0.1087)	-0.4977 (0.1635)	-0.7292 (0.2163)
Hospital 98	-0.5897 (0.1119)	-2.0894 (0.1769)	-1.9008 (0.2476)	-0.3003 (0.206)
Doctor 98	1.1151 (0.1416)	-1.2912 (0.1445)	-0.8839 (0.1879)	1.6466 (0.3167)
Out. Services 98	-0.9093 (0.1173)	-2.3955 (0.202)	-1.7317 (0.2265)	-0.6198 (0.2107)
Insurance 98	-2.153 (0.1928)	-1.9864 (0.1839)	1.3563 (0.2425)	1.5764 (0.3103)
Hospital 00	-0.4964 (0.1116)	-2.3675 (0.2055)	-1.8513 (0.2414)	-0.3027 (0.2059)
Doctor 00	1.8296 (0.1969)	-1.1808 (0.1475)	-1.0113 (0.1984)	1.8422 (0.3502)
Out. Services 00	-0.6878 (0.1128)	-1.9203 (0.1637)	-1.3282 (0.1962)	-0.7014 (0.2134)
Insurance 00	-3.0475 (0.3244)	-2.8725 (0.3276)	1.7698 (0.331)	1.883 (0.3843)
Hospital 02	-0.3912 (0.1101)	-1.8862 (0.1638)	-1.5852 (0.2233)	0.2613 (0.2118)
Doctor 02	1.4451 (0.1564)	-0.8772 (0.1261)	-0.8912 (0.1924)	1.9985 (0.3773)
Out. Services 02	-0.4054 (0.1094)	-1.6045 (0.145)	-1.3054 (0.1976)	-0.2843 (0.2044)
Insurance 02	-2.5491 (0.2404)	-2.1248 (0.1915)	1.0893 (0.2167)	1.6755 (0.3362)

Standard errors in brackets.

Chapter 4

Incentive and selection effects of Medigap insurance on inpatient care

4.1 Introduction

Medicare is a public program which provides health insurance for the elderly (aged 65 or older) and some disabled non elderly. As many other standard health insurance plans, Medicare relies deeply on mechanisms such as coinsurance, deductibles and copayments to control health care expenditure for many covered services. This insurance structure leaves beneficiaries at risk for large out-of-pocket expenses. As a result, many beneficiaries purchase voluntary supplemental private policies, such as Medigap, to fill Medicare's gaps in non-covered health care services and limit cost sharing.

Medicare cost-sharing structure reflects the belief that health insurance, by lowering the price per services, gives individuals' an *incentive* to increase the demand for health care. Although the presence of an incentive effect - usually called *ex-post moral hazard* - is very well known by the theoretical literature on contract theory (Arrow[6], Pauly [97] and Zweifel and Manning [122]), its empirical relevance is still debated in the literature see Abbring *et al.* [1] - [2], Buchmuller *et al.* [17], Cardon and Henderl [21], Cohen [28], Schellhorn [106], and Cohen and Spiegelman [30] for a review. A major difficulty in estimating the presence of moral hazard in Medigap insurance is the existence of *self-selection*, since individuals who expect high health care costs may choose a more generous coverage and then ex-post purchase more services.

In an important body of literature following the seminal paper by Chiappori and Salanié [25], the presence of asymmetric information in insurance markets is appraised using the so called "positive correlation" (PC) test (see two recent reviews by Cohen and Spiegelman [30] and Einav *et al.* [47]), which rejects the null hypothesis of no asymmetric information when there is a positive correlation between insurance purchases and risk occurrence, conditional on the individual characteristics used by insurers to price contracts. The PC test, however, cannot disentangle incentive and selection effects, since finding a positive insurance coverage-risk occurrence correlation in the data does not provide conclusive evidence whether there is adverse selection into insurance contracts, moral hazard, or both.

There are different ways to distinguish empirically selection from incentive effects. A strategy is to use experimental data such as the RAND Health Insurance Experiment (RHIE), where to identify the incentive effect controlling for self-selection individuals were randomly assigned to plans with different coverages so that insurance choice becomes exogenous. Another strategy is to exploit (quasi) natural experiments where insurance choice or the incentive structure has been modified exogenously (Chiappori *et al.* [26] and Eichner [45]). In observational studies, the standard approach to evaluate incentive effects controlling for self-selection is to model endogenously the insurance choice and estimate a bivariate probit model with a recursive structure between insurance and health care utilization (Holly, *et al.* [73], Jones *et al.* [80], Buchmuller *et al.* [17]).

The aim of this paper is to evaluate how different empirical strategies perform in trying to separate incentive and selection effects of Medigap supplemental insurance on inpatient care. The econometric approach relies on a recursive bivariate probit model and on a multiresponse discrete finite mixture model. We use data from the Health and Retirement Study (HRS), which contains information on a rich set of variables concerning health status and individual preferences for risk, Medigap purchase and hospital admissions as binary dependent variables representing insurance purchase and health care utilization.

The paper is organized as follows. In the next section we report a brief overview of Medicare and Medigap insurance programs; section 3 reviews the main empirical contribution in the related literature; we then discuss the model to be estimated (section 4) using the data described in section 5. Finally section 6 and 7 report estimation results and some concluding remarks respectively.

4.2 Health insurance and access to care for elderly in US

4.2.1 Medicare

Medicare is probably the main source of health insurance for all individuals aged 65 in US and the coverage is near universal (about 97% of the elderly have Medicare)[1].

The Medicare program consists mainly of two plans in which people may be enrolled. The first plan, named Medicare Part A, is also known as "Hospital Insurance" since it covers the basic hospital's health care services such as inpatient's admissions. Most of beneficiaries, who have paid Medicare taxes for at least 10 years, are automatically enrolled with their spouse in Part A when they turn 65. Part A plan pays almost the entire medical expenditure (except a deductible) for the first 60 nights of inpatient hospital staying and imposes an increasing cost sharing structure if hospital admission lasts over this first period.

The second plan is Medicare Part B. Most of beneficiaries choose to extend Medicare Part A insurance coverage to Part B because it covers several medicare services such as doctors' services, outpatient care and some preventive services. Part B enrollment requires the payment of a monthly premium which may depend on income. Part B's deductible and co-payment amount respectively to $110 and to 20% of expenses.

4.2.2 Supplemental insurance coverage and medigap policy

There are several limitations of Medicare original plans: limitation in the coverage of health care services, high out-of-pocket expenses to beneficiaries and lack of a catastrophic cap expenditure. These induce seniors to seek additional coverage provided by private insurance.

There are three main sources of supplemental private insurance which pay

[1] Current Population Reports (2005) "Income, Poverty and Health Insurance Coverage in the United States: 2004"

for some additional (to Medicare) services or help pay the share of the costs of Medicare-covered services. The first one is the employer-sponsored supplemental insurance and it is purchased usually by a former employer or union. The second one is represented by Tricare (available only to military personal) and the Medicare Advantage plans (Part C) provided by private health insurance.

The third one and also the most common source of supplemental coverage comes from Medigap-private health insurance which are specifically designed to cover those "gaps" of coverage left by original Medicare plans. Since 1990 the Medigap insurance market is highly regulated by Federal law. Medigap plans are standardized into ten plans,"A" through "J", which cover a single individual, offer certain additional services and help beneficiaries pay health care cost (deductibles and co-payment) that the original Medicare plan does not cover. This means that if individuals are enrolled in Medicare plus a supplemental Medigap insurance, health care cost is covered by both plans. For example the basic plan, A, covers the entire coinsurance or copayments for hospital stays, physician visits and outpatient care.

Federal regulation of the Medigap market designed a particular mechanism favoring the insured: Medigap insurance companies must offer the basic plan "A" if they offer any other more generous plan. In addition, there is a free enrolment period which lasts for six months from the first month in which people are both 65 years old and enrolled in Medicare Part B. During this period Medigap cannot refuse any insurer even if there are pre-existing conditions. Legal restrictions involve also the pricing criteria, which are mainly based on individual's age and gender.

In the supplemental health insurance market the most popular Medigap plans are C and F, because they cover major benefits and are less expensive than other plans. For example, plan C offers coverage for skilled-nursing-facility coinsurance, foreign-travel emergencies, deductibles that are required under traditional Medicare and other basic benefits like hospital and outpatient coinsurance. All these plans include Medigap A, which is the basic one, the least expensive and least comprehensive. This plan covers several losses; for example it provides (increasing) coverage for daily medicare copayment per day for hospitals stays; it reimburses the full cost of up to 365 additional hospitals days and the partial cost of other services related to doctor's (outpatient) visits, preventive health screening and outpatient prescription drug.

4.3 Related Literature

The empirical literature on the incentive and selection effects in health insurance is growing at a fast pace and controversial since disentangling the two effects is not straightforward because the unobserved nature of individual preferences and health status pose serious endogeneity problems.

A "radical" solution is to exploit experiments or some particular features of the data which make insurance choice exogenous. The best known study is the RAND Health Insurance Experiment conducted in 1974. To control for self-selection, individuals were randomly assigned to insurance plans with different coinsurance rate. Manning *et. al.* [91] show that patients ensured by a plan with first dollar coverage had 37% more physician visits than those facing co-insurance rates of 25% suggesting strong evidence of *ex-post* moral hazard, but they found no significant differences among the alternative coinsurance plans in the use of inpatient services; individuals with free insurance plan tended to use slightly more inpatient services than individuals with coinsurance.

In non-experimental settings most of the studies use large observational data sets which include information on individuals, health care services and insurance status. There are different econometric strategies to empirically appraise this issue. The first approach considers insurance choice as exogenous in the health care utilization equation, and estimates health care utilization with a probit model (Hurd and McGarry [76]) or a two-parts model (Ettner [51], Khandker and McCormack [86]) using health indicators to mitigate the presence of unobserved heterogeneity in health status.

Another approach is to model endogenously insurance choice considering both selection on observable and unobservable factors. In this framework many studies conducted in the European health insurance market exploited a recursive bivariate probit to model simultaneously the probability to have at least one inpatient stay and purchasing supplemental insurance (Holly *et al.* [73], Jones *et al.* [80], Buchmueller *et al.* [17]). In general the most common finding in these empirical studies exploiting the bivariate probit model is to find a positive (direct) effect of insurance on health care demand and no positive (statistically significant) correlation between residuals of the insurance and the health risk occurrence equations.

In general results from observational data show that Medicare enrollees with supplemental insurance (Medigap or employer plans) have higher levels of total spending, though differences are small for inpatient care, and that individuals reporting better health are significantly more likely to enroll in private supplemental plans (see e.g. Ettner [51], Cartwright *et al.* [22], Hurd and McGarry [76]).

These findings are arousing a great deal of interest among researchers. In particular Finkelstein and McGarry [58] and Cutler *et al.* [32] take an innovative approach based on insurance company unused variables to test the (positive) correlation between health care utilization and insurance coverage. Using one wave of the Health Dynamics Among Oldest (AHEAD) Cutler *et al.* [32] identify two groups of individuals who purchase supplemental insurance and use health care services: those who prefer insurance for cautionary reasons and ex-post are less likely to use health care, and those who are subjectively riskier and ex post have higher risk occurrence.[2] Their findings show that individuals who engage in risky behavior are systematically less likely to hold Medigap. Moreover people with higher preferences for insurance appear to have lower expected claims, creating offsetting advantageous selection.[3]

Using the the same dataset of Cutler *et al.* [32], Fang *et al.* [53] provide strong evidence of advantageous selection in the Medigap market and find cognitive ability is an important factor influencing selection. They conclude that this reflects the idea that senior citizens may have difficulties in understanding Medicare and Medigap rules. Therefore, the existence of multiple sources of private information depending for example on cognitive skills (Fang *et al.* [53]) or actual risk and risk preferences (Cutler *et al.* [32]), may seriously affect the interpretation of zero correlation between insurance and health care utilization in the bivariate probit model, since these different unobserved forces may wash each other out (see also Dardanoni and Li Donni [35]).

A rather different approach is to control for unobserved heterogeneity using LC analysis. Deb and Trivedi [38]-[39] develop a finite mixture negative binomial model and estimate health care demand for several health care measures. They

[2] AHEAD is a cohort of the Health and Retirement Study (HRS) from which our sample is drawn.

[3] The first paper to notice a *negative* risk/coverage correlation is Hemenway [70], which called it *propitious* selection. A theoretical paper which analyzes advantageous selection is De Meza and Webb [37]. A review of the extensive empirical literature who found advantageous selection in insurance markets is Cohen and Spiegelman [30] and Einav *et al.* [47]. Evidence of positive and statistical significant relationship between risk aversion and health attitudes have been found in the US health insurance market by Vistnes and Banthin [8] and Landerman *et al.* [89]

distinguish two unobserved groups in the population: the "healthy" and the "ill". After controlling for these two "types" of people, they find that individuals with supplementary private health insurance tend to seek care from physicians and non-physicians more often than the uninsureds, while this effect is not significant for inpatient staying.[4]

Our paper could be located ideally in this framework, drawing inspiration from the recursive bivariate probit literature, from the studies on multiple dimension of private information of Fang et. al [53], Cutler et al. [32] and Finkelstein and McGarry [58], and the LC models of Deb and Trivedi [38]-[39].

4.4 Modeling incentive and selection effects

Suppose at time $t = T$ we observe a binary variable $S_T \in \{0, 1\}$ which takes value 1 if an individual has bought a supplemental insurance contract which protects from a fixed loss, and a binary variable $M_T \in \{0, 1\}$ which takes value 1 if the individual incurs the loss. In general, S_T and M_T need not be binary variables, but frequently the researcher can only observe whether the individual occurred in the risk or whether she is covered by an insurance plan.

Standard economic theory predicts that risk occurrence and insurance coverage are positively correlated and this relationship depends on two sources (Rothschild and Stiglitz [104] and Arnott and Stiglitz [5]). On the one hand when individuals have private information about their actual risk, the insurance contract will be *adversely selected*, with high risk individuals choosing higher insurance coverage. On the other hand insurance contract may give the *incentive* to increase risk occurrence by increasing the probability to incur in the risk (ex-ante moral hazard) or by increasing utilization (ex post moral hazard). Both sources of asymmetric information (adverse selection and moral hazard) will cause the same observed positive correlation in the data.

The predicted positive correlation between risk and coverage has inspired

[4]A general limitation affecting many studies on incentive and selection effects in this context is to leave out a third possible "health improving" effect of insurance purchase. If Medigap increases outpatient visits or expands drugs coverage, this can increase health status and decrease subsequent medical care use. Some studies have investigated the effects of supplementary insurance on the health status of the elderly (see. e.g. Doescher *et al.* [42] and Dor *et al.* [43]) supporting this possibility. In our application, we have tried to control for this effect by using various aggregate health status measures. We thank an anonymous referee for this remark.

the seminal contribution by Chiappori and Salanié [25], who considered the testable implications of asymmetric information in insurance markets, and proposed the so called Positive Correlation (PC) test. The PC test rejects the null of absence of private information in a given insurance market when, conditional on consumers' characteristics used by insurance companies to price contracts, individuals with more coverage experience more of the insured risk.

4.4.1 The insurer's model

Let $\boldsymbol{w}$ denote the set of variables used by insurance companies to price a given insurance contract. Notice that conditioning on $\boldsymbol{w}$ is crucial to properly identify selection effects, since without conditioning on $\boldsymbol{w}$ it would not be possible to know whether a correlation arises because individuals which are offered the same contract have different risk (adverse selection) or rather because they face contracts at different prices (Chiappori and Salanié [25], Einav and Finkelstein [46]).

Since individuals with the same value of $\boldsymbol{w}$ face the same insurance contract, one can study incentive and selection effects using the same logic as the PC test by considering the following recursive model:

$$M_T = 1\big(a + bS_T + \boldsymbol{c}'\boldsymbol{w} + \eta_M\big) \tag{4.1}$$

$$S_T = 1\big(d + \boldsymbol{e}'\boldsymbol{w} + \eta_S\big) \tag{4.2}$$

(with $1(.)$ denoting the indicator function), where the residual heterogeneity which is induced by all variables not used by insurance companies is collected in two random variables (η_M, η_S). Within model (4.1)–(4.2), adverse selection implies a positive correlation between η_M and η_S, while the incentive effect implies a positive value of the coefficient b. Model (4.1)–(4.2) can be estimated with standard maximum likelihood methods by assuming that (η_M, η_S) are normally distributed with standardized margins and correlation coefficient equal to, say, ρ, in which case it is called a recursive bivariate probit model.

Estimating (4.1)–(4.2) with a recursive bivariate probit to disentangle incentive and selection effects has some theoretical and econometric difficulties. From an econometric point of view, identification of model (4.1)–(4.2) relying on the normality assumption, though theoretically feasible as long as data on $\boldsymbol{w}$ are of full rank (see Wilde [118]), it is quite fragile in the absence of exclusion restrictions. On the other hand, from the theoretical perspective, the bivariate normality assumption can be severely limiting when residual heterogeneity is actually multidimensional. For example, as explained by Dardanoni and Li Donni [35], if there are two conflicting sources of private information (following Finkelstein and McGarry [58], say individual's actual risk and risk preference), one can observe that the insurance contract could be at the same time *both* adversely *and* favorably selected by different types (where a type is defined as a given combination of actual risk and risk preference), but these selection effects could go undetected by using the single statistic ρ.[5]

A standard practice in the applied economics literature is to check the robustness of recursive bivariate probit estimates by comparing them to the estimates obtained by performing separate probit regressions for the two response variables. It is not uncommon, however, that the two sets of estimates differ substantially, even in the case when a standard LR test does not reject the null of $\rho = 0$. When this happens, it can be taken as evidence of potential identification problems in the bivariate probit model estimates.

Discrete multivariate finite mixture model

An alternative procedure to estimate equations (4.1)–(4.2) is by using a latent class model, trying to control for residual heterogeneity by identifying a finite number of "types" which differ with respect to their attitude to buy insurance and to use medical care. In particular, we can assume that U is a discrete random variable taking values in, say, $\{1, \ldots, m\}$, which define m unobservable heterogeneous "types"; in practice, U can be seen as a cross-classification of underlying unobservable individual characteristics, such as risk tolerance and attitudes to use medical care. Equations (4.1)–(4.2) can be rewritten as

$$M_T = 1\big(\sum_{u=1}^{m} \alpha_{M_T}(u)U(u) + bS_T + \boldsymbol{c}^{'}\boldsymbol{w} + \epsilon_M\big) \tag{4.3}$$

$$S_T = 1\big(\sum_{u=1}^{m} \alpha_{S_T}(u)U(u) + \boldsymbol{e}^{'}\boldsymbol{w} + \epsilon_S\big) \tag{4.4}$$

[5]Dardanoni and Li Donni [35] give various theoretical and empirical examples with multidimensional private information implies the presence of selection effects which are however undetected by ρ.

where $U(1), \ldots, U(m)$ denote the set of m dummy variables indicating "latent type" membership, so that the coefficients $\alpha_{M_T}(u)$ and $\alpha_{M_T}(u)$ can be interpreted as random intercepts with a nonparametric discrete specification, like in Heckman and Singer [68]. Since no underlying structure is imposed on U, it can capture in a unrestrictive way any variable which affects medical care use and insurance demand after conditioning on $\boldsymbol{w}$, and thus take into account the potential multidimensionality of residual heterogeneity. (ϵ_M, ϵ_S) on the other can be seen as uncorrelated idiosyncratic errors. Estimation of equations (4.3)–(4.4) involves first the choice of an appropriate distribution for (ϵ_M, ϵ_S); we assume a standard logistic in order to use logit parameters for estimation. However, while relying on the parametric choice of the distribution of the idiosyncratic errors, no parametric structure is imposed on the residual heterogeneity variable U; for this reason model (4.3)–(4.4) is known in the literature as a *semiparametric* finite mixture model. Well known applications are Deb and Trivedi ([38] and [39]) in health economics and Cameron and Heckman [20] in education.

The semiparametric finite mixture model is well established and achieves nonparametric estimation of residual heterogeneity. However, conditionally on $\boldsymbol{w}$ we observe only two binary variables (that is, 3 parameters), while even with two mixture types there are 5 parameters to estimate (four α's and a mixture probability). Thus, it must rely on covariates' variation to achieve parameters' identification. In practice, it typically achieves identification of very few unobserved types; in many applications only two types are identified.

To achieve sharper identification of the heterogeneous types we may use a set of observable *indicators* of the variables which affect insurance choice and loss occurrence after conditioning on $\boldsymbol{w}$. To identify U, we exploit the panel nature of the dataset, and in particular lagged values of M and S, to define a set of *auxiliary* equations which are used as *indicators* of U and help identifying the mixture components probabilities in (4.3)–(4.4), which are of primary interest:

$$\begin{aligned} Pr(M_{T-t} = 1 \mid u, \boldsymbol{w}) &= \Lambda(\alpha_{M_{T-t}}(u) + \boldsymbol{\beta}'_{M_{T-t}}\boldsymbol{w}),\ t = 1, \ldots, T-1 \\ Pr(S_{T-t} = 1 \mid u, \boldsymbol{w}) &= \Lambda(\alpha_{S_{T-t}}(u) + \boldsymbol{\beta}'_{S_{T-t}}\boldsymbol{w}),\ t = 1, \ldots, T-1 \end{aligned} \tag{4.5}$$

where Λ denotes the binary logit link.[6]

[6] In the auxiliary system (4.5) one could also condition the probabilities of medical care and insurance choice on lagged values, in which case the latent types would be defined relatively to past choices. For example, if one conditions current insurance choice on its lagged values, one identifies risk attitudes relatively to past choices, while if no conditioning is made, one tends to identify unadjusted risk preferences. In our experience, for the purpose of estimating the

Our discrete multiresponse latent class model is completed by the types' membership probabilities $P(u)$. To force the types probabilities to lie between zero and one and sum to one, it is convenient use a multinomial logit parameterization:

$$Pr(U = u) = \frac{\exp(\alpha_U(u))}{\sum_{u=1}^{m} \exp(\alpha_U(u))}, \ \alpha_U(m) = 0 \tag{4.6}$$

so that the $m - 1$ logit parameters α_U are simply reparametrization of the membership probabilities, and do not impose any parametric restriction on the distribution of U.

The discrete multiresponse finite mixture model is defined by the main equations (4.3)–(4.4), the auxiliary equations (4.5), and the membership probability equation (4.6). It can be seen as an instance of a discrete multivariate MIMIC (Joreskog and Goldberger[81]) model (see Goodman [63] for the seminal paper on finite mixture models with multivariate binary responses, and Huang and Bandeen-Roche [75] for a recent general treatment). Differently to the MIMIC model, the residual heterogeneity U is not a continuous univariate variable on the real line, but an unstructured nonparametric variable. Notice that, contrary to the semiparametric finite mixture model which does not use the auxiliary system of equations (4.5), the discrete multiresponse finite mixture model can achieve identification of a considerable number of heterogeneous types without relying on covariates variation.[7]

Within this model the absence of incentive and selection effects can be tested by imposing appropriate restrictions on the model parameters: one rejects the null of no incentive effects when b in equation (4.3) is positive, while the null of the absence of selection effects is rejected when the equality constraints $\alpha_{M_T}(1) = \cdots = \alpha_{M_T}(m)$ and $\alpha_{S_T}(1) = \cdots = \alpha_{S_T}(m)$ are rejected.

---

parameters of the main system of interest, in practice there is little difference in the results obtained under the two approaches.

[7]For an intuitive explanation on how this can be done, notice that conditionally on $\boldsymbol{w}$, if for example $T = 3$, we observe the joint distribution of 6 response variables with $2^6 - 1 = 63$ observable parameters, with $7 \times m - 1$ unknown parameters. Thus, in principle up to 9 different types can be identified. However, this is only a necessary condition for the identification of the unobservable parameters; there are well known pathological examples in the literature on finite mixture models which show that this counting condition is not sufficient.

4.4.2 An extended model

Following the seminal paper by Chiappori and Salanié [25], current practice on testing for asymmetric information in insurance markets appraises the existence of asymmetric information after conditioning on individual characteristics which are used by insurance companies to set premiums and price contracts. For the purpose of detecting selection effects, this approach seems by now universally accepted (see e.g. Cohen and Spiegelman [30], Einav *et al.* [47]). However, this approach contrasts with previous studies which have tried to estimate the incentive effect of medical care insurance.

Research on moral hazard distinguishes between ex ante moral hazard, which refers to the attitudes of insured to take reduced precautions, and ex post moral hazard, which refers to actions engaged by insured after a loss occurs. The main issue to measure this second source of moral hazard consists of disentangling the variation of medical care consumption corresponding to a given level of information (e.g. related to individual tastes, preferences beliefs, etc.) from the variation of consumption due to the structure of the contract. For this reason the standard applied approach is to include in addition to variables used to price insurance contracts, also additional controls devoted to capture potential variation in consumption.

To study asymmetric information, it is not clear whether conditioning on all the possible determinants of demand for medical care should be necessarily preferred to conditioning only on variables used by insurers, since both moral hazard and adverse selection directly depend on the structure of the insurance contract and then on the variables used to price individual risk (see e.g. Gardiol *et al.* [61] and Chiappori *et al.* [26]). Thus, for completeness we present both model specifications by including all controls traditionally employed in the empirical literature on insurance and demand for medical care (see e.g. Buchmuller *et al.* [17], Holly *et al.* [73], Jones *et al.* [80]).

In this section we will thus extend model (4.1)–(4.2) by considering

$$M_t = 1\big(a + bS_t + \boldsymbol{c}'\boldsymbol{w}_M + \eta_M\big) \tag{4.7}$$

$$S_t = 1\big(d + \boldsymbol{e}'\boldsymbol{w}_S + \eta_S\big) \tag{4.8}$$

where $\boldsymbol{w}_M$ and $\boldsymbol{w}_S$ denote all observable determinants of individual's medical care and insurance choices, and η_M and η_S collect all residual unobservable heterogeneity which affects medical care use and insurance demand.

Recursive bivariate probit

In equations (4.7)-(4.8), $\boldsymbol{w}_M$, $\boldsymbol{w}_S$, η_M and η_S denote all observable and unobservable determinants of individual's medical care and insurance choices, reflecting individuals' preferences and constraints, and are likely to include individuals' characteristics such as risk tolerance, attitude towards medical care use and insurance purchase, actual (health) riskiness and so on. Following standard models in the literature (Ettner [51], Holly *et al.* [73], Buchmuller *et al.* [17], Jones *et al.* [80]), we assume that $\boldsymbol{w}_M$ and $\boldsymbol{w}_S$ include demographic, socio-economic and health status observables, which we denote $\boldsymbol{x}_M$ and $\boldsymbol{x}_S$, and lagged medical care and insurance partecipation which may act as proxies for unobservable attitudes to buy insurance and use health care. We then rewrite equations (4.7) and (4.8) as

$$M_T = 1\big(a + bS_T + \boldsymbol{c}'\boldsymbol{x_M} + \sum_{t=1}^{T-1} d_{T-t}M_{T-t} + \sum_{t=1}^{T-1} e_{T-t}S_{T-t} + \eta_M\big) \tag{4.9}$$

$$S_T = 1\big(f + \boldsymbol{g}'\boldsymbol{x_S} + \sum_{t=1}^{T-1} h_{T-t}M_{T-t} + \sum_{t=1}^{T-1} k_{T-t}S_{T-t} + \eta_S\big) \tag{4.10}$$

Assuming that (η_M, η_S) are distributed as a bivariate normal with standard margins and correlation coefficient ρ, equations (4.9)-(4.10) define a standard recursive bivariate probit.

The recursive bivariate probit model above is simple to estimate and to interpret, allows estimation of the effect of supplementary insurance on medical care use and testing for selection effects with standard software, and has been much used in this context (Buchmueller *et al.* [17], Holly *et al.* [73], and Jones *et al.* [80]). Furthermore, by appropriate inclusion/exclusion restriction, it can allow a much more robust parameters' identification compared to model (4.1)–(4.2) above. However, it still relies on bivariate normality to achieve parameters' identification, and, as explained above, it can still run into difficulties in detecting selection effects when residual unobserved heterogeneity is multidimensional.

Discrete multivariate finite mixture model

Following the same logic as above, our strategy is to control for residual unobserved heterogeneity by identifying a finite number of "types" which differ with respect to their attitude to buy insurance and to use medical care, by letting again U denote a discrete random variable taking values in $\{1, \ldots, m\}$. Equations (4.7)–(4.8) can then be rewritten as

$$M_T = 1\Big(\sum_{u=1}^{m} \alpha_{M_T}(u)U(u) + bS_T + \boldsymbol{c}'\boldsymbol{x_M} + \sum_{t=1}^{T-1} d_{T-t}M_{T-t} + \sum_{t=1}^{T-1} e_{T-t}S_{T-t} + \epsilon_M\Big) \tag{4.11}$$

$$S_T = 1\Big(\sum_{u=1}^{m} \alpha_{S_T}(u)U(u) + \boldsymbol{g}'\boldsymbol{x_S} + \sum_{t=1}^{T-1} h_{T-t}M_{T-t} + \sum_{t=1}^{T-1} k_{T-t}S_{T-t} + \epsilon_M\Big) \tag{4.12}$$

where again the idiosyncratic errors (ϵ_M, ϵ_S) are assumed logisticly distributed in order to use logit parameters for estimation.

To achieve sharper identification of the heterogeneous types we define again a set of *auxiliary equations*:

$$\begin{aligned} Pr(M_{T-t} = 1 \mid u, \boldsymbol{x}_M) &= \Lambda(\alpha_{M_{T-t}}(u) + \boldsymbol{\beta}'_{M_{T-t}}\boldsymbol{x}_M),\ t = 1, \ldots, T-1 \\ Pr(S_{T-t} = 1 \mid u, \boldsymbol{x}_S) &= \Lambda(\alpha_{S_{T-t}}(u) + \boldsymbol{\beta}'_{S_{T-t}}\boldsymbol{x}_S),\ t = 1, \ldots, T-1 \end{aligned} \tag{4.13}$$

and complete the model by modeling the types membership probabilities $P(u)$ as in equation (4.6).

An interesting feature of this model is that past medical care use and insurance choice variables perform a 'double duty': they help identification of residual heterogeneity with the system of auxiliary equations (4.13) (as in the model of section 4.4.1 above), but they also enter directly the conditional distribution of current medical care and insurance choice in equations (4.11) and (4.12). Absence of selection effects and moral hazard can be tested by imposing appropriate restrictions on the α parameters.

4.4.3 Estimation of the discrete multiresponse finite mixture model

Estimation of the model parameters can be obtained by the EM algorithm, which is the standard approach for maximum likelihood estimation of finite mixture models, and has been shown (Dempster *et al.* [40]) to converge to the maximum of the true likelihood. Given the binary nature of the response variables, the E-step is equivalent to compute, for each subject, the posterior probability of belonging to each unobservable type; the M-step requires maximization of a multinomial likelihood with individual covariates. A detailed discussion of the EM algorithm in a system of non linear structural equations with latent classes can be found in Bergsma *et al.* [14].[8]

It is well known that the EM algorithm may converge even if the model is not identified, which is a crucial issue for finite mixture models. Conditions for parametric identification are discussed in Theorem 1 of Huang and Bandeen Roche [75]; local identification can be checked using the the numerical test described by Forcina [59], which consists in checking that the Jacobian of the transformation between the parameters of the observable responses and the mixture model parameters is of full rank for a wide range of parameter values.

Despite the usefulness of finite mixture models to detect underlying residual heterogeneity, one unresolved issue in their application is how to determine the number of unobserved types m. The currently preferred approach suggests the use of Schwartz's Bayesian Information Criterion (BIC) to guide this choice, which in certain conditions is known to be consistent and generally helps preventing overparametrization (see McLachlan and Peel [93] for a thorough introduction to finite mixture models and a review of existing criteria for the choice of the number of types). BIC is calculated from the maximized log-likelihood L by penalizing parameters' proliferation, $BIC = -2L + \upsilon log(n)$, where n denotes sample size and υ the number of parameters; the model with the lowest BIC is preferred.

Finally, it may be worth noticing that, since the number of types m is not predetermined, formal hypotheses tests performed in finite mixture models are in fact conditional on m, and pre-testing for the number of types may invalidate distributional results of the test statistics employed. Of course, while pre-testing is a common problem in most applied research whenever final estimates are obtained after searching for appropriate specification, this is an issue which should be kept in mind whenever test results differ significantly when performed under different values of m.

[8]We are grateful to Antonio Forcina for kindly providing the Matlab code for the EM estimation.

4.5 Data and Descriptive Statistics

We use data from the Health and Retirement Study (HRS). Since 1992 the HRS is a biennial survey targeting elderly Americans over the age of 50 sponsored by the National Institute on Aging. Although the survey is not conducted on an yearly basis, from 1998 it provides longitudinal data for an array of information, consistently administrated, on several different fields such as health and health care utilization, type of insurance coverage, socioeconomic condition, retirement plans and family structure and transfers.

For our purpose we use the last available wave on 2006 (that is, $T = 2006$) as reference point to collect information on insurance status, health care utilization, health and socio-economic characteristics from the previous two waves (2002 and 2004). To evaluate more closely the effect of asymmetric information on Medicare expenditure, we consider a sample restricted to Medicare Part A or B enrollees over the last wave. This means that we consider only individuals older than 65 in 2002. Information about Medicare is binary coded and it is clearly reported in the survey as the first question asked in the insurance section.

Since we study the effect of supplemental insurance (Medigap) on health care, we also exclude those individuals that received additional coverage through a former employer, spouse or some other government agency. Following Fang *et al.* [53] we define an individual as having additional health insurance coverage (Medigap) if she purchased directly health insurance policy in addition to Medicare. The HRS asked respondents whether they are also covered by Medicaid, CHAMPUS or CHAMPVA (Tri-care) and who payed for the supplemental insurance. This detailed information allows to identify any source of (additional) coverage available to individual. Since our focus is to disentangle incentive and selection effects at time 2006 for those individual who deliberately choose to purchase additional coverage, our sample should be limited only to people who are covered by Medicare part A or B, are not covered by additional public insurance and pay personally the required monthly premium. Therefore we exclude those who are enrolled in any other public program different from Medicare or that are covered at year 2006 by Medigap insurance plan provided by own or spouse's former employer.

Our sample is composed by 3368 individuals and descriptive statistics are reported in table A.1. Supplemental insurance status at 2002 (spins02), 2004 (spins04) and 2006 (spins06) is coded as binary variable which takes 1 if respondent has any (no long-term care) supplemental private Medigap insurance coverage. Almost 50 percent of Medicare beneficiaries in the sample has a supplemental insurance, and 87 percent of them are continuously covered since they turn 65; thus, most of Medicare beneficiaries purchase a supplemental insurance coverage as soon as they are enrolled in Medicare Part A or B. In addition to these variables we also use information on whether additional coverage was provided in the previous years by a former employer (iemp04) or by the spouse (iemps04). These variables have been used in the literature to explain individual choice to take out voluntary supplemental coverage, and can be considered as appropriate instruments since they are likely to affect insurance choice but they can be excluded from the inpatient care equation (Ettner [51] finds that selection into additional coverage can be driven by employer-provided plans since high-risk (or risk averse) workers may self-select into jobs that provide better retirement medical benefits).

The HRS offers detailed information on health care consumption. We focus on hospital staying over the three waves (h02, h04 and h06). Hospital admissions account for 29% of the Medicare's total expenditure and it is an important part of the Medicare total expenditure (see CMS, Office of the Actuary, National Health Expenditure Accounts, 2007). These variables are binary and take value 1 if individual had at least one hospital admission, and 0 otherwise. In the Medigap market insurance companies are constrained by Federal law to use only age and gender to price contracts. Therefore as variables used by insurer to predict risk we only use these two measures; for age we use four five-years dummies for added flexibility.[9]

In the extended model we follow previous studies on health care demand (Cameron *et al.* [19], Jones *et al.* [80], Deb and Trivedi [39] Vera-Heranandez [113]) and insurance choice (Propper [100]-[101], Cameron and Trivedi [18]) as guidance in selecting five groups of variables describing individual socioeconomic characteristics, insurance status, health care consumption, health status and individual risk preferences. Health conditions, in particular, have an important influence both on the decision to subscribe supplementary insurance as well as on the utilization of health care services. In the HRS, health condition is measured

[9] Using dummy variables is a common strategy to increase the computational speed of maximization methods, such as the the EM algorithm, that tend to be slow. However for robustness we performed all estimations using different age-band definitions and found that estimated relationships did not vary substantially.

along different measures. We include in the analysis the number of self-reported doctor diagnosed longstanding or chronical diseases (DIS) - such as high blood pressure, hypertension, cardiovascular disease, lung disease, kidney conditions, emotional and psychiatric problems - and the index of Activities of Daily Living (ADL), which measures difficulties in bathing, dressing, eating, getting in/out of bed and walking across a room and is defined on a discrete scale ranging between 0-5. Both indices are averaged out over the three years period to capture the persistency of need status and any preexisting condition which may affect insurance choice over the time.

Finally the last group of variables includes socio-economic characteristics, such as education and wealth. Education is measured using three dummy variables indicating whether individual i) is a high-school graduate, ii) has a degree which is less than a BA, or iii) has a college degree which is a BA or greater. The base category includes people with no qualification or lower than high-school. Individual wealth is measured using four dummy variables indicating the quartile of the total wealth (including the second home) distribution. The base category represents the poorest.

As reported in table A.1, comparing the sub-sample averages for the two groups of people with and without supplemental insurance, we find that people with additional insurance tend to have higher education, to be in the top wealth quartile, and to have a lower score of ADL, though the average number of diseases does not vary substantially.

4.6 Results

4.6.1 Insurer's model: single probit and bivariate probit results

Table A.2 reports in the first two columns the estimated parameters from the two single binary probit models for hospital admission and Medigap purchase, and in the third and fourth columns the estimated parameters of the bivariate recursive probit model (4.1)-(4.2). For all model specifications one observes that age has a significant positive effect on inpatient staying and females are more likely to be covered by Medigap.

Estimated coefficients for the recursive bivariate probit reveal a strongly significant incentive effect; the average marginal effect is unbelievable huge at 0.517.[10] Regarding selection effects, the estimated ρ coefficient is equal to -.880, with a s.e. equal to .137. The Wald test statistic is equal to 5.09 and is asymptotically distributed as a χ_1^2, so the null $\rho = 0$ is rejected with p-value 0.024. On the other hand, the likelihood ratio (LR) test statistic is equal to 1.26 and is also asymptotically distributed as a χ_1^2, so the null $\rho = 0$ is not rejected with p-value 0.262. Thus, the two testing procedures for the null of absence of selection effects give very contrasting results. Monfardini and Radice [95] show that the LR test in general performs significantly better compared to the Wald test, and argue that when the two test give contrasting results it usually signals identification problems. This is certainly possible in this case since no inclusion/exclusion restrictions are exploited to identify model parameters in the two equations.

As a robustness check, and since the LR test cannot reject the null hypothesis that conditional on $\boldsymbol{w}$ insurance choice is exogenous from hospital utilization, we may look at the results of the single equation probits. In this framework the key variable of interest is the incentive effect. Conditioning on buyer characteristics used by insurer to price contract, individuals with supplemental Medigap coverage are more likely to have any hospital inpatient staying, since b is positive and statistically significant at the 5% level. The average marginal effect is equal to 0.0382, which can be compared with the huge estimated effect of 0.517. The difference between these two estimates is a further indication of possible identification problems with the bivariate probit model. Summing up, it is probably safe to consider estimated parameters of the recursive bivariate probit with healthy skepticism.

4.6.2 Insurer's model: discrete multiresponse finite mixture model

We start by estimating the model under different numbers of latent classes. Parameters estimates for $m = 2, 3, 4, 5$ are reported in tables A.3–A.7 in the Appendix.[11] A glance at tables A.6-A.7 reveals that: i) the incentive effect param-

[10] Average marginal effects here and hereafter are computed averaging individuals marginal effects.

[11] We first performed the numerical test of Forcina [59] described in Section 4.4.3 to check model identifiability. The model passed such a test for each $m = 2, 3, 4, 5$ with samples of

eter b of main interest is quite robust to the number of types $m = 3, 4, 5$; ii) older individuals seem to significantly have more hospital admissions; iii) there is a substantial amount of heterogeneity as described by the α coefficients, which seem to vary very significantly across types.

Table 4.1 below reports the maximized log-likelihood L and Schwartz's Bayesian Information Criterion BIC. BIC indicates that four types are adequate to represent residual heterogeneity. We comment on incentive and selection effects focusing on the case $m = 4$.

Table 4.1: Model Selection Criteria for the insurer's model

	Number of Latent Classes			
	m=2	m=3	m=4	m=5
L	-11591.58	-11477.28	-11400.12	-11397.15
BIC	23540.54	23368.79	23271.32	23322.237
# of parameters	44	51	58	65

Let us first focus on the incentive effect parameter b reported in table A.6. This effect is positive and statistically significant at the 10% level. The average marginal effect is equal to 0.0383, which is practically identical to the average marginal effect calculated in the single equation probit model reported (0.0382). For robustness, we also calculated the marginal effect in the linear probability model (0.0382).

Regarding selection effects, we can derive the conditional probabilities of hospital admission and medigap purchase for the 4 types from the estimated coefficients α in Tables A.4 and A.5.

Table 4.2: Estimated conditional probabilities

	U=1	U=2	U=3	U=4
h06	0.2391	0.2923	0.6538	0.6860
h04	0.1575	0.1474	0.8338	0.7790
h02	0.2531	0.2115	0.5862	0.5691
spins06	0.8121	0.1327	0.1050	0.8050
spins04	0.9526	0.1012	0.0977	0.9167
spins02	0.8469	0.1235	0.1618	0.8784

Estimated probabilities are averaged out for $\boldsymbol{w}$.

10,000 draws. For robustness, we also estimated all versions of the discrete multiresponse finite mixture model from different starting points in order to check the presence of possible local maxima.

The table above shows very substantial heterogeneity in hospital admission and Medigap purchase by the four types: types 3-4 tend to have a much higher probability of hospital admission than types 1-2, while types 1-4 have a much higher probability of buying Medigap than types 2-3. Thus, it appears that there are at least two different sources of residual heterogeneity: the attitude to buy insurance (which is high for types one and four and low for types two and three), and the propensity to use medical care (which is high for types three and four and low for types one and two). Comparing to the average hospital admission and Medigap purchase probabilities, it emerges that types two and four adversely select Medigap insurance, while types one and three favorably select Medigap insurance. Thus the overall picture which emerges from our estimates is strongly suggestive of multidimensional residual private information, and there appears to be *both* favorable and adverse selection in this market, since the contract is favourably selected by individuals who have the same (high or low) propensity for insurance purchase and hospital admission, and adversely selected by individuals who have opposite attitudes.

While it is apparent that residual heterogeneity coefficients α's (and thus the conditional probabilities above) vary substantially across types, we can still check whether this simply reflects sampling variation by testing the equality constraint $\alpha_{h06}(1) = \cdots = \alpha_{h06}(4)$ and $\alpha_{spins06}(1) = \cdots = \alpha_{spins06}(4)$. As expected, the LR test statistic overwhelmingly rejects the null with p-values less than 10^{-4}.

4.6.3 Extended model: binary probit and bivariate probit results

Table A.8 shows the single equation probits and the recursive bivariate probit estimated coefficients of insurance choice and inpatient stay for the extended model (4.9)-(4.10).

Estimated coefficients from the recursive bivariate probit model reveal that the probability of a hospital staying increases with age, but not with education and wealth, and it is positively associated with low persistent health status and past inpatient stay. On the other hand, the probability of enrolling in a supplementary insurance plan increases with past insurance status, wealth and education. In addition, having better coverage in the past (provided either individually, spins02 and spins04, or by a former employer, iemp04 and iemps04) significantly increases the probability of having supplementary insurance. These results seem to indicate that individuals with a persistent coverage over the years tend to be more likely to purchase supplemental insurance.

Estimated coefficients for the recursive bivariate probit reveal again a large but insignificant incentive effect; the average marginal effect is rather large at 0.183. Regarding selection effects, the estimated ρ is equal to 0.248, with a s.e. equal to 0.281. The Wald test statistic and LR test statistics are equal to 0.71 and 0.92 and are asymptotically distributed as χ^2_1. Both statistics cannot reject the null $\rho = 0$ with a p-values of 0.397 and 0.336 respectively. Thus, the two testing procedures for the null of absence of selection effects in this case give very similar conclusions.

Although Wilde [118] shows that identification can be obtained even when the same exogenous regressors appear in both equations (provided that regressors satisfy an appropriate full rank condition), a more robust identification can be obtained by appropriate exclusion/inclusion restrictions. In our model it implies excluding variables from the outcome (hospital) equation which are correlated with insurance choice but, conditional on exogenous variables, uncorrelated with hospital utilization. Following Ettner [51], in our specification we have adopted as instruments past supplemental insurance coverage provided by own (iemp04) or spouse's (iemps04) former employer.

To show sensitivity of the estimated coefficients to the assumed exclusion restrictions, we estimated the bivariate probit model under different specifications of the outcome equation, analogously to Buchmuller *et al.* [17]. Table A.9 reports results on the two parameters of interest: the incentive effect b and the selection effect ρ for several versions of the model. Although the point estimates vary only slightly depending on which variables are excluded from the utilization equation, the qualitative results are similar across the different specifications. For all versions of the recursive bivariate probit model, the insurance coefficient is larger than the one from the single equation probit model, and in all cases the correlation coefficient is negative and never statistically significant. Therefore it is plausible to conclude that insurance choice is exogenous in the hospital utilization equation, and then single equation probit model gives a more reliable estimate of the incentive effect parameter b, with an average marginal effect equal to 0.0409.

4.6.4 Extended model: discrete multiresponse finite mixture model

We first estimate the model under different numbers m of latent classes. Results for $m = 2, 3, 4, 5$ are reported in Appendix, while table 4.3 below reports the maximized log-likelihood L, the BIC and the number of parameters in each LC model specification.[12] BIC seems to indicate that three LC are adequate to represent unobserved heterogeneity. A glance at tables A.12-A.15 reveals also that estimated coefficients do not seem to vary substantially with respect to the number of LC specifications. We comment on parameters' estimates focusing on the case $m = 3$.

Table 4.3: Model Selection Criteria for the extended model

	Number of Latent Classes		
	$m = 2$	$m = 3$	$m = 4$
L	-11104.58	-11072.95	-11058.65
BIC	23037.62	23031.21	23059.47
$\#ofparmaters$	102	109	116

Estimated coefficients reveal a very similar picture to the probit models above; the probability of a hospital staying increases with age, but not with education and wealth, and is positively associated with past utilization and low persistent health status; the probability of enrolling in a supplementary insurance plan increases with past insurance status and wealth. The variables iemp04 and iemps04 are again very strongly significant, so that generally individuals with a persistent coverage over the years tend to be more likely to purchase supplemental insurance.

Regarding incentive effects, the coefficient b is positive and significant at the 10% significance level. The average marginal effect is equal to 0.0439, which is slightly higher than the probit marginal effect of 0.0409. Again, for robustness, we also calculated the marginal effect in the linear probability model (0.041).

Finally, looking at the estimated α coefficients for the main system (Table A.11), it emerges that, contrary to the model estimated in section 4.6.2, estimated coefficients of the residual heterogeneity U do not vary significantly

[12] We also performed the numerical test of Forcina [59] and the model passed such a test for each $m = 2, 3, 4$ with samples of 10,000 draws. For robustness, we also estimated all versions of the discrete multiresponse finite mixture model from different starting points in order to check the presence of possible local maxima.

across types. Thus, we first can check whether the observed pattern of α's in the main equation system is simply due to sampling variation by imposing equality constraints. The LR test statistics when $\alpha_{h06}(1) = \alpha_{h06}(2) = \alpha_{h06}(3)$ and $\alpha_{spins06}(1) = \alpha_{spins06}(2) = \alpha_{spins06}(3)$ are equal to 4.71 and 3.91 respectively. These are distributed as χ^2_2, so that the null hypothesis of no selection effects cannot be rejected at 5% significance level with p-values of 0.094 and 0.141.

Our results suggest that after controlling for all relevant observable determinants of inpatient staying and Medigap purchase, there seems to be no significant residual private information.

4.7 Final remarks

In the health insurance market consumers have private information about their health status (actual risk) and preferences. As a result, insurance contracts may be affected by incentive and/or selection effects. A standard way to study these effects is to measure the impact of insurance purchase on health care use and their residual association, after conditioning on variables used by insurance companies to price contracts. The standard econometric approach for this purpose relies on a recursive bivariate probit model.

In this paper, we explore the extent to which supplemental health insurance (Medigap) affects inpatient care in two distinct cases: using as conditioning variables those used by insurers' to price individual's risk, or alternatively using the rich set of observable characteristics which are available in the HRS data. While in highly competitive and unregulated markets the two sets of variables may coincide, in heavily regulated markets (such as Medigap) the two models may give very different results.

In both cases, we estimated two alternative econometric models: a standard recursive bivariate probit, and a discrete multiresponse finite mixture model. The main picture which emerges from our estimation is the following:

- Estimated incentive effects are quite similar across models, with average marginal effects ranging from 0.0382 to 0.0439, which are slightly smaller than the difference, in the sample, between the probability of hospital admission of individuals who are covered (0.36) and those who are not covered (0.31) by Medigap. Estimated coefficients of the incentive effects are generally significant between the 5% and 10% significance level. These results are broadly consistent with previous studies on Medigap (Ettner [51], Cartwright *et al.* [22], Hurd and McGarry [76]).

- There seems to be very significant selection effects when one conditions only on variables used by Medigap insurers, with the presence of *both* adversely and favorably selected individuals. This stems from the multidimensional nature of residual heterogeneity.

- On the other hand, when a rich set of observable variables, including past insurance decisions and past inpatient stay, are employed for conditioning, there seems to be no statistically significant private information in this market. Thus, one may conclude that selection effects in Medigap are mainly due to regulatory constraints. This suggests that future research may fruitfully investigate the welfare implication of regulatory Medigap constraints. A very simple intuitive explanation of the welfare loss due to adverse and favorable selection is contained in Einav and Finkelstein [46].

- In this setting, the recursive bivariate probit model runs into theoretical problems to detect selection effects when residual heterogeneity is multidimensional. Furthermore, in this data, the model seems to run into empirical problems probably due to identification issues.

Appendix A

Tables

Table A.1: Sample characteristics and variables definition

Variable	Definition of Binary Variables	Full Sample	No Ins	With Ins
Insurance Status				
spins06	1 = enrolled in Medigap at 2008.	0.42	0.00	1.00
spins04	1 = enrolled in Medigap at 2004.	0.45	0.21	0.80
spins02	1 = enrolled in Medigap at 2002.	0.43	0.22	0.72
iemp04	1 = additional coverage from former emp. at 2004.	0.08	0.10	0.05
iemps04	1 = additional coverage from spouse emp. at 2004.	0.04	0.05	0.04
Hospital Admission				
h06	1 = entered a hospital in 2005-2006.	0.33	0.31	0.36
h04	1 = entered a hospital 2003-2004.	0.29	0.29	0.29
h02	1 = entered a hospital in 2001-2002.	0.27	0.26	0.28
Variables Used by insurer to price Medigap plan				
age75	1 = aged between 76 and 80 years.	0.24	0.26	0.22
age80	1 = aged between 81 and 85 years.	0.18	0.17	0.20
age85	1 = aged between 86 and 90 years.	0.11	0.09	0.13
age90	1 = older than 90 years.	0.05	0.05	0.05
fem	1 = female.	0.61	0.59	0.63
Other Controls unused by insurer				
edu3	1 = if individual is high-school graduate.	0.35	0.33	0.39
edu4	1 = if individual has a degree lower than BA.	0.19	0.17	0.20
edu5	1 = if individual has college degree or greater.	0.16	0.16	0.15
wealth2	1 = if individual is in the second wealth quartile.	0.25	0.26	0.24
wealth3	1 = if individual is in the third wealth quartile.	0.25	0.24	0.27
wealth4	1 = if individual is in the top wealth quartile.	0.25	0.23	0.28
dis	1 = average number of disease over 2002-2006.	2.23	2.24	2.21
adl	1 = average ADL over 2002-2006.	0.29	0.31	0.26

Table A.2: Insurer's probit models for hospital admission and insurance choice

	Probit Model		Bivariate Probit Model	
Independent Variables	Hospital 2006	Insurance 2006	Hospital 2006	Insurance 2006
spins06	0.105	.	1.435	.
	(0.0456)	.	(0.161)	.
age75	0.216	-0.0401	0.177	-0.0307
	(0.0579)	(0.0561)	(0.0546)	(0.0552)
age80	0.356	0.127	0.184	0.129
	(0.0626)	(0.0610)	(0.0765)	(0.0610)
age85	0.396	0.238	0.148	0.239
	(0.0752)	(0.0736)	(0.0950)	(0.0739)
age90	0.413	0.0743	0.256	0.0793
	(0.103)	(0.102)	(0.105)	(0.101)
fem	-0.00646	0.0915	-0.0551	0.0903
	(0.0463)	(0.0450)	(0.0424)	(0.0446)
Constant	-0.660	-0.312	-0.974	-0.314
	(0.0490)	(0.0433)	(0.0462)	(0.0433)
# of Obs.	3368			
Log-likelihood	-2111.97	-2275.96	-4387.31	

Note: Robust standard errors are reported in brackets

Table A.3: Estimated Class Membership Probabilities for the insurer's model

	$m = 2$	$m = 3$	$m = 4$	$m = 5$
$\alpha(1)$	0.4440	0.3872	0.2913	0.2705
$\alpha(2)$	0.5560	0.1771	0.3878	0.3827
$\alpha(3)$	.	0.4357	0.1680	0.1419
$\alpha(4)$	.	.	0.1530	0.0531
$\alpha(5)$	.	.	.	0.1517

Table A.4: Estimated intercepts for the insurer's model: main equations

	$m = 2$		$m = 3$		$m = 4$		$m = 5$	
	Coef.	St. Er.	Coef.	St. Er.	Coef.	St. Er.	Coef.	St. Er.
Hosp. Adm. 2006								
$\alpha_{h06}(1)$	-0.77	0.11	-1.40	0.12	-1.62	0.22	-1.81	0.29
$\alpha_{h06}(2)$	-0.78	0.05	0.31	0.17	-1.34	0.12	-1.35	0.12
$\alpha_{h06}(3)$	.	.	-0.86	0.12	0.22	0.17	0.22	0.18
$\alpha_{h06}(4)$	.	.	.	.	0.37	0.26	-0.04	0.35
$\alpha_{h06}(5)$	.	.	.	.	.	.	0.30	0.26
Sup. Ins. 2006								
$\alpha_{spins06}(1)$	1.37	0.08	-2.03	0.11	1.38	0.11	1.38	0.13
$\alpha_{spins06}(2)$	-2.08	0.09	-2.03	0.19	-2.00	0.11	-2.10	0.15
$\alpha_{spins06}(3)$	.	.	1.42	0.08	-2.26	0.23	-1.84	0.16
$\alpha_{spins06}(4)$	.	.	.	.	1.34	0.18	0.79	1.92
$\alpha_{spins06}(5)$	.	.	.	.	.	.	1.36	0.17

Table A.5: Estimated intercepts for the insurer's model: auxiliary equations

	$m=2$		$m=3$		$m=4$		$m=5$	
	Coef.	St. Er.	Coef.	St. Er.	Coef.	St. Er.	Coef.	St. Er.
Hosp. Adm. 2002								
$\alpha_{h02}(1)$	-0.93	0.06	-1.78	0.14	-1.49	0.14	-1.62	0.17
$\alpha_{h02}(2)$	-1.07	0.06	0.03	0.16	-1.73	0.14	-1.73	0.14
$\alpha_{h02}(3)$	.	.	-0.95	0.06	-0.02	0.17	-0.02	0.18
$\alpha_{h02}(4)$	.	.	.	.	-0.10	0.16	-0.33	0.28
$\alpha_{h02}(5)$	.	.	.	.	.	.	-0.15	0.16
Hosp. Adm. 2004								
$\alpha_{h04}(1)$	-0.90	0.06	-2.39	0.35	-2.36	0.40	-2.54	0.49
$\alpha_{h04}(2)$	-0.92	0.05	1.07	0.34	-2.44	0.39	-2.57	0.43
$\alpha_{h04}(3)$	.	.	-0.95	0.06	1.13	0.40	1.25	0.46
$\alpha_{h04}(4)$	.	.	.	.	0.76	0.34	0.60	0.50
$\alpha_{h04}(5)$	.	.	.	.	.	.	0.39	0.27
Sup. Ins. 2002								
$\alpha_{spins02}(1)$	1.59	0.09	-2.23	0.13	1.50	0.12	1.49	0.13
$\alpha_{spins02}(2)$	-2.09	0.09	-1.66	0.16	-2.19	0.13	-2.21	0.15
$\alpha_{spins02}(3)$	.	.	1.63	0.09	-1.88	0.19	-2.01	0.60
$\alpha_{spins02}(4)$	.	.	.	.	1.77	0.21	-0.18	0.92
$\alpha_{spins02}(5)$	.	.	.	.	.	.	1.76	0.32
Sup. Ins. 2004								
$\alpha_{spins04}(1)$	2.50	0.15	-2.48	0.16	2.75	0.25	2.84	0.34
$\alpha_{spins04}(2)$	-2.54	0.14	-2.30	0.23	-2.54	0.18	-2.57	0.19
$\alpha_{spins04}(3)$	.	.	2.57	0.15	-2.58	0.30	-2.66	0.33
$\alpha_{spins04}(4)$	.	.	.	.	2.14	0.26	-2.09	2.58
$\alpha_{spins04}(5)$	.	.	.	.	.	.	-1.69	0.17

Table A.6: Estimated parameters for the insurer's model: main equations

	$m=2$		$m=3$		$m=4$		$m=5$	
	Coef.	St. Er.	Coef.	St. Er.	Coef.	St. Er.	Coef.	St. Er.
Hosp. Adm. 2006								
spins06	0.16	0.11	0.24	0.12	0.25	0.13	0.23	0.17
age75	0.36	0.10	0.39	0.10	0.41	0.11	0.41	0.11
age80	0.59	0.10	0.61	0.11	0.69	0.12	0.68	0.12
age85	0.65	0.12	0.65	0.13	0.74	0.15	0.73	0.15
age90	0.68	0.17	0.73	0.18	0.79	0.20	0.80	0.20
fem	-0.01	0.08	0.00	0.08	-0.02	0.09	-0.01	0.09
Sup. Ins. 2006								
age75	-0.23	0.14	-0.24	0.14	-0.23	0.14	-0.25	0.15
age80	0.26	0.16	0.22	0.16	0.26	0.16	0.25	0.17
age85	0.50	0.19	0.46	0.19	0.51	0.19	0.61	0.22
age90	-0.09	0.26	-0.12	0.26	-0.09	0.26	-0.19	0.27
fem	0.15	0.11	0.17	0.11	0.14	0.11	0.16	0.12

Table A.7: Estimated parameters for the insurer's model: auxiliary equations

	$m=2$		$m=3$		$m=4$		$m=5$	
	Coef.	St. Er.	Coef.	St. Er.	Coef.	St. Er.	Coef.	St. Er.
Sup. Ins. 2002								
age75	0.14	0.15	0.13	0.15	0.14	0.15	0.14	0.15
age80	0.28	0.17	0.23	0.16	0.28	0.17	0.25	0.16
age85	0.56	0.20	0.50	0.20	0.57	0.20	0.57	0.21
age90	0.12	0.28	0.09	0.28	0.12	0.28	0.04	0.27
fem	0.24	0.12	0.25	0.12	0.22	0.12	0.25	0.12
Sup. Ins. 2004								
age75	-0.20	0.24	-0.21	0.23	-0.19	0.24	-0.25	0.32
age80	0.46	0.25	0.36	0.24	0.45	0.25	0.52	0.30
age85	0.83	0.29	0.72	0.28	0.86	0.29	1.01	0.31
age90	0.54	0.41	0.48	0.39	0.56	0.41	0.48	0.48
fem	0.18	0.19	0.21	0.18	0.16	0.19	0.32	0.23
Hosp. Adm. 2002								
age75	0.27	0.10	0.29	0.11	0.29	0.11	0.29	0.11
age80	0.53	0.11	0.55	0.12	0.60	0.12	0.59	0.12
age85	0.66	0.13	0.67	0.14	0.72	0.14	0.71	0.15
age90	0.65	0.17	0.69	0.19	0.73	0.20	0.73	0.20
fem	-0.05	0.08	-0.04	0.09	-0.06	0.09	-0.05	0.09
Hosp. Adm. 2004								
age75	0.25	0.10	0.34	0.13	0.42	0.18	0.39	0.17
age80	0.42	0.11	0.52	0.14	0.75	0.20	0.69	0.19
age85	0.69	0.12	0.81	0.16	1.15	0.25	1.09	0.25
age90	0.77	0.17	1.00	0.22	1.34	0.32	1.34	0.32
fem	-0.07	0.08	-0.08	0.10	-0.15	0.14	-0.13	0.13

Table A.8: Extended probit models for Hospital Admission and Insurance Choice at 2006

	Probit Model		Bivariate Probit Model	
Independent Variables	Hospital 2006	Insurance 2006	Hospital 2006	Insurance 2006
spins06	0.125	.	0.559	.
	(0.0601)	.	(0.483)	.
spins04	0.0678	1.401	-0.140	1.402
	(0.0672)	(0.0635)	(0.217)	(0.0635)
spins02	-0.0367	0.736	-0.136	0.733
	(0.0613)	(0.0595)	(0.127)	(0.0598)
h04	0.530	-0.0175	0.525	-0.0179
	(0.0522)	(0.0599)	(0.0544)	(0.0597)
h02	0.148	0.0202	0.145	0.0255
	(0.0541)	(0.0601)	(0.0541)	(0.0605)
dis	0.169	-0.0150	0.169	-0.0140
	(0.0198)	(0.0218)	(0.0199)	(0.0220)
adl	0.180	-0.0249	0.180	-0.0240
	(0.0360)	(0.0395)	(0.0353)	(0.0391)
age75	0.139	-0.0763	0.146	-0.0711
	(0.0607)	(0.0661)	(0.0603)	(0.0668)
age80	0.255	0.107	0.241	0.112
	(0.0664)	(0.0737)	(0.0689)	(0.0741)
age85	0.165	0.186	0.142	0.190
	(0.0792)	(0.0879)	(0.0830)	(0.0882)
age90	0.158	0.00672	0.156	0.0111
	(0.112)	(0.129)	(0.112)	(0.128)
fem	-0.0213	0.0799	-0.0252	0.0810
	(0.0512)	(0.0554)	(0.0514)	(0.0553)
edu3	-0.0330	0.0751	-0.0401	0.0789
	(0.0594)	(0.0651)	(0.0595)	(0.0654)
edu4	-0.0498	0.0900	-0.0593	0.0943
	(0.0727)	(0.0803)	(0.0730)	(0.0806)
edu5	0.0523	0.0629	0.0443	0.0742
	(0.0794)	(0.0860)	(0.0797)	(0.0880)
wealth2	-0.0433	-0.0393	-0.0378	-0.0379
	(0.0677)	(0.0767)	(0.0671)	(0.0765)
wealth3	-0.0843	0.114	-0.0933	0.113
	(0.0702)	(0.0773)	(0.0720)	(0.0775)
wealth4	-0.0715	0.145	-0.0815	0.146
	(0.0783)	(0.0862)	(0.0812)	(0.0863)
iemp04	0.0567	0.536	.	0.535
	(0.0888)	(0.0957)	.	(.0.0951)
iemps04	0.140	0.617	.	.0.627
	(0.110)	(0.117)	.	(.0.116)
Constant	-1.205	-1.468	-1.213	-.1.478
	(0.0914)	(0.102)	(0.0901)	(.0.103)
# of Obs.	3368			
Log-likelihood	-1932.60	-1560.62	-3492.76	

Note: Robust standard errors are reported in brackets

Table A.9: Partial results for alternative bivariate probit specifications

Variables excluded in the utilization equation	Insurance Coefficient	ρ	Likelihood Ratio
Probit Model	0.125	-	-
	(0.0601)		
iemps04	0.579	-0.266	0.69
	(0.581)	(0.362)	
iemp04	0.307	-0.103	0.09
	(0.769)	(0.449)	
iemp04+iemps04	0.559	-0.254	0.92
	(0.483)	(0.300)	

Note: Robust standard errors are reported in brackets.
The sample size for all models is 3368

Table A.10: Estimated Class Membership Probabilities for the extended model

	$m = 2$	$m = 3$	$m = 4$
$\alpha(1)$	0.4806	0.1412	0.2961
$\alpha(2)$	0.5194	0.3701	0.2150
$\alpha(3)$	.	0.4887	0.2762
$\alpha(4)$	.	.	0.2127

Table A.11: Estimated intercepts for the extended model: main equations

	$m = 2$		$m = 3$		$m = 4$	
	Coef.	St. Er.	Coef.	St. Er.	Coef.	St. Er.
Hosp. Adm. 2006						
$\alpha_{h06}(1)$	-1.07	0.40	-1.81	0.30	-1.05	0.15
$\alpha_{h06}(2)$	-1.20	0.07	-1.07	0.10	-1.54	0.45
$\alpha_{h06}(3)$	.	.	-1.13	0.39	-0.90	0.51
$\alpha_{h06}(4)$	.	.	.	.	-1.01	0.48
Sup. Ins. 2006						
$\alpha_{spins06}(1)$	-1.40	0.42	-2.48	0.41	-2.57	0.37
$\alpha_{spins06}(2)$	-2.19	0.10	-2.16	0.13	-1.45	0.42
$\alpha_{spins06}(3)$	.	.	-1.28	0.39	-2.89	0.79
$\alpha_{spins06}(4)$	.	.	.	.	-1.09	0.52

Table A.12: Estimated intercepts for the extended model: auxiliary equations

	$m=2$		$m=3$		$m=4$	
	Coef.	St. Er.	Coef.	St. Er.	Coef.	St. Er.
Hosp. Adm. 2002						
$\alpha_{h02}(1)$	-0.99	0.06	0.14	0.37	-2.67	0.56
$\alpha_{h02}(2)$	-1.15	0.06	-1.94	0.29	-0.26	0.27
$\alpha_{h02}(3)$	.	.	-0.98	0.06	-1.88	0.36
$\alpha_{h02}(4)$	.	.	.	.	0.04	0.32
Hosp. Adm. 2004						
$\alpha_{h04}(1)$	-0.94	0.06	0.39	0.41	-2.13	0.40
$\alpha_{h04}(2)$	-1.02	0.06	-1.81	0.29	-0.15	0.23
$\alpha_{h04}(3)$	.	.	-0.95	0.06	-1.44	0.21
$\alpha_{h04}(4)$	.	.	.	.	-0.31	0.21
Sup. Ins. 2002						
$\alpha_{spins02}(1)$	1.73	0.34	-2.62	0.47	-2.71	0.42
$\alpha_{spins02}(2)$	-2.98	0.57	-2.83	0.41	-2.74	0.57
$\alpha_{spins02}(3)$	.	.	1.54	0.19	2.24	0.50
$\alpha_{spins02}(4)$	.	.	.	.	0.87	0.24
Sup. Ins. 2004						
$\alpha_{spins04}(1)$	1.77	0.27	-4.07	1.45	-2.09	0.35
$\alpha_{spins04}(2)$	-2.45	0.45	-2.49	0.37	1.11	0.28
$\alpha_{spins04}(3)$	.	.	1.87	0.24	2.34	0.43
$\alpha_{spins04}(4)$	.	.	.	.	1.42	0.36

Table A.13: Estimated parameters for the extended model: main equations

	$m = 2$		$m = 3$		$m = 4$	
	Coef.	St. Er.	Coef.	St. Er.	Coef.	St. Er.
Hosp. Adm. 2006						
spins06	0.22	0.11	0.21	0.11	0.27	0.18
spins04	0.03	0.21	-0.02	0.22	-0.15	0.30
spins02	-0.14	0.24	-0.12	0.20	-0.21	0.25
h04	0.87	0.08	1.02	0.12	0.97	0.19
h02	0.24	0.09	0.37	0.11	0.35	0.26
dis	0.29	0.03	0.27	0.04	0.28	0.04
adl	0.29	0.06	0.28	0.06	0.28	0.06
age75	0.24	0.10	0.24	0.10	0.25	0.10
age80	0.43	0.11	0.43	0.11	0.43	0.11
age85	0.29	0.14	0.28	0.14	0.30	0.14
age90	0.28	0.19	0.28	0.19	0.30	0.19
fem	-0.01	0.09	-0.01	0.09	0.00	0.09
edu3	-0.04	0.10	-0.04	0.10	-0.03	0.10
edu4	-0.08	0.12	-0.10	0.12	-0.09	0.12
edu5	0.08	0.13	0.08	0.13	0.08	0.13
wealth2	-0.05	0.12	-0.05	0.12	-0.03	0.12
wealth3	-0.10	0.12	-0.09	0.12	-0.08	0.13
wealth4	-0.07	0.13	-0.05	0.14	-0.04	0.14
Sup. Ins. 2006						
spins04	2.02	0.23	1.87	0.24	2.66	0.45
spins02	0.84	0.28	0.84	0.23	1.49	0.34
h04	-0.03	0.11	0.01	0.13	-0.42	0.24
h02	0.02	0.11	0.05	0.13	-0.60	0.32
dis	-0.01	0.04	-0.01	0.04	0.04	0.05
adl	-0.07	0.07	-0.07	0.07	0.01	0.08
age75	-0.12	0.12	-0.12	0.12	-0.12	0.13
age80	0.21	0.14	0.22	0.14	0.24	0.15
age85	0.38	0.17	0.39	0.17	0.38	0.18
age90	0.06	0.23	0.08	0.23	0.08	0.25
fem	0.17	0.11	0.17	0.11	0.10	0.11
edu3	0.21	0.13	0.22	0.13	0.16	0.14
edu4	0.20	0.15	0.21	0.15	0.22	0.16
edu5	0.09	0.16	0.09	0.16	0.14	0.17
wealth2	0.03	0.15	0.04	0.15	-0.10	0.16
wealth3	0.31	0.15	0.34	0.15	0.20	0.16
wealth4	0.35	0.17	0.37	0.17	0.22	0.18
iemp04	0.97	0.17	0.99	0.17	1.02	0.18
iemps04	1.11	0.21	1.13	0.22	1.19	0.23

Table A.14: Estimated parameters for the extended model: auxiliary equations, insurance choice

	$m = 2$		$m = 3$		$m = 4$	
	Coef.	St. Er.	Coef.	St. Er.	Coef.	St. Er.
Sup. Ins. 2002						
dis	0.23	0.08	0.20	0.06	0.21	0.07
adl	-0.28	0.13	-0.25	0.10	-0.27	0.11
age75	0.28	0.21	0.29	0.19	0.34	0.20
age80	0.39	0.25	0.35	0.21	0.28	0.21
age85	0.80	0.36	0.69	0.28	0.86	0.30
age90	0.53	0.41	0.52	0.36	0.67	0.38
fem	0.45	0.20	0.35	0.16	0.41	0.17
edu3	0.77	0.31	0.65	0.22	0.69	0.23
edu4	0.40	0.28	0.38	0.23	0.43	0.24
edu5	-0.55	0.26	-0.45	0.23	-0.40	0.24
wealth2	0.84	0.31	0.74	0.23	0.86	0.26
wealth3	0.76	0.31	0.68	0.24	0.76	0.26
wealth4	0.79	0.32	0.69	0.25	0.73	0.27
Sup. Ins. 2004						
dis	0.12	0.06	0.14	0.07	0.14	0.07
adl	-0.20	0.11	-0.22	0.12	-0.25	0.12
age75	0.03	0.18	0.08	0.20	0.15	0.21
age80	0.46	0.23	0.51	0.24	0.45	0.24
age85	0.82	0.30	0.89	0.32	1.14	0.35
age90	0.89	0.39	1.08	0.43	1.44	0.48
fem	0.39	0.18	0.38	0.18	0.45	0.19
edu3	0.81	0.25	0.89	0.25	0.95	0.27
edu4	0.56	0.25	0.64	0.27	0.69	0.28
edu5	-0.13	0.24	-0.10	0.26	-0.01	0.27
wealth2	0.89	0.27	0.98	0.26	1.11	0.29
wealth3	1.13	0.30	1.31	0.31	1.41	0.34
wealth4	1.07	0.31	1.17	0.30	1.21	0.32

Table A.15: Estimated parameters for the extended model: auxiliary equations, hospital utilization

	$m = 2$		$m = 3$		$m = 4$	
	Coef.	St. Er.	Coef.	St. Er.	Coef.	St. Er.
Hosp. Adm. 2002						
dis	0.39	0.03	0.42	0.04	0.49	0.06
adl	0.21	0.05	0.24	0.06	0.27	0.08
age75	0.16	0.11	0.16	0.12	0.18	0.13
age80	0.42	0.11	0.42	0.12	0.47	0.14
age85	0.39	0.14	0.40	0.15	0.45	0.17
age90	0.42	0.19	0.39	0.20	0.47	0.24
fem	-0.13	0.09	-0.14	0.10	-0.17	0.11
edu3	0.21	0.10	0.23	0.11	0.24	0.13
edu4	0.33	0.12	0.35	0.13	0.40	0.15
edu5	0.02	0.14	0.03	0.15	0.04	0.17
wealth2	-0.07	0.12	-0.07	0.13	-0.10	0.14
wealth3	0.05	0.12	0.06	0.13	0.05	0.15
wealth4	-0.20	0.13	-0.22	0.15	-0.26	0.16
Hosp. Adm. 2004						
dis	0.42	0.03	0.47	0.04	0.48	0.04
adl	0.23	0.05	0.27	0.06	0.27	0.06
age75	0.15	0.10	0.15	0.11	0.15	0.12
age80	0.31	0.11	0.31	0.12	0.32	0.13
age85	0.43	0.13	0.45	0.15	0.46	0.15
age90	0.56	0.18	0.55	0.20	0.59	0.21
fem	-0.12	0.09	-0.14	0.09	-0.15	0.10
edu3	0.03	0.10	0.04	0.11	0.02	0.11
edu4	0.22	0.12	0.24	0.13	0.24	0.14
edu5	-0.04	0.14	-0.03	0.15	-0.03	0.15
wealth2	0.02	0.11	0.03	0.13	0.02	0.13
wealth3	0.05	0.12	0.06	0.13	0.06	0.13
wealth4	-0.05	0.13	-0.06	0.14	-0.07	0.15

Chapter 5

Risk preference heterogeneity and multiple demand for insurance

5.1 Introduction

There is an emerging economic literature which examines the relationship between risk tolerance, insurance demand and attitude to risky behaviours (see Cutler *et al.* [32], Einav *et al.* [49], Barseghyany *et al.* [9]). Importantly, there is little consensus among these studies on how general are individual's financial and non-financial risk preferences to predict insurance demand.

Classical economic theory assumes that individuals have the same attitude to bear risk in different contexts, and then models all risky individual decisions using the same value (utility) function over wealth.

This implies that multiple choices over different risk dimensions (such as different insurance markets) taken by the same individual should reflect the same degree of risk aversion even if the contexts of decisions are different. Although there are evidence of positive correlation between financial and no financial risk aversion which may support the domain-general component of risk preference hypothesis (DGC), there is a large and important literature mostly related to behavioral economics which poses serious concerns on the internal validity of this assumption (Rabin [102] and Rabin and Thaler [103]). They argue that individuals' decision to take risk is influenced by the context of choice. This idea is supported by several findings obtained by exploiting lab experiments which show little or even no significant commonality between risky choice in different domains. As a result one would need to impose more theoretical assumptions to extend risk preference parameter estimated for one market to another one (Cohen and Einav [29]). The existence of this debate does not pose a clear view on where the reality lies especially when survey data, mainly employed in empirical research in economics, are used.

Some recent papers consider this issue in insurance markets and evaluate whether risk preferences are general. Cohen and Einav [29] and Barseghyan *et al.* [9] model individual choice following the standard expected utility theory and use insurance data on deductible choices to estimate risk aversion parameters in the sample by comparing the variation in the deductible menus across individuals and their choices from these menus. Their results show the existence of substantial heterogeneity in risk preferences and in general data do not support the context-invariant risk preferences hypothesis. Clearly since this approach estimates the distribution of risk aversion in the sample from individuals' deductible choices and claims, it requires a domain-specific model of ex-ante heterogeneity in risk. Einav *et al.* [49] propose another approach which focuses on within-person correlation between risky choices an individual makes across different domains. The idea is that under the no DGC hypothesis, individuals have different attitudes to bear risk among domains and then insurance decisions should not be inter-related after conditioning on individual characteristics. They reject the null that there is no domain-general component of preferences and find that the common element of an individual's preferences may be stronger among domains that are "closer" in context.

In this paper we propose an alternative framework to examine how general are risk preferences in the multiple demand of insurance using survey data. Specifically we extend previous setting focused on residual correlation across insurance (Einav *et al.* [49]) by identifying unobserved "types" with different risk preferences and examining the effect of these "types" on insurance purchase decision. We use data from the Health and Retirement Study (HRS) on four insurance purchase decisions: life insurance, Medicare supplemental insurance (Medigap), long-term care insurance and annuity. Using these data we investigate the stability of unobserved individual risk preferences across insurance choices and whether the context-specific differences are relevant. Our results show the existence of a stable pattern of individual risk preferences over different insurance domains, which supports the idea of domain-general component of preference. In addition we also provide further evidence, as found by Einav *et al.* [49], that context plays an important role in determining insurance choices particularly when insurance coverage decisions involve similar specific contexts.

The paper is organized as follows. The next section reviews the main empirical literature; section 3 reports a brief overview of insurance markets we are analysing and describes the data; we then discuss the model to be estimated (section 4). Section 5 and 6 report respectively the main findings and some concluding remarks.

5.2 Literature Review

The paper is related to three literatures that cut across insurance economics, health economics and experimental economics. The first stream of literature studies the determinants of the demand for insurance and has been mainly developed in the context of the analysis of asymmetric information. Friedman and Warshawsky [60] study the selection effect in the annuity market, which is mainly related to the existence of unobserved heterogeneity in risk preferences and risk aversion. In a more recent series of papers Finkelstein and Poterba [55]-[56] and McCarthy and Mitchell [92] using data from different countries provide more evidence of the existence of unobservables in the decision to purchase annuity and suggest the possible existence of risk preference-based selection effect.

In contrast to the papers on the demand for annuity, those studying selection in life insurance markets reach generally puzzling conclusions, since data do not show clear conclusion on how heterogenous private information affects the purchase decision (see Cawly and [23]). Browne and Kim [16] study the demand for life insurance across different countries and find the religion being an important determinant. They claim that the degree of risk aversion in a country could be related to the predominant religion, and therefore, religion affects the demand for life insurance.

Long-term insurance combines elements of both annuity and life insurance. Finkelstein and McGarry [58] study the US market using data from the Asset and Health Dynamics (AHEAD) that is part of the HRS. They find that demand for coverage is substantially related to risk aversion. In particular they use as proxy of risk preferences the share of preventive care activities undertaken by a subject and whether individual always wear seat belt, and assume that who take more of these actions are more risk-averse. Their results show that insurance purchase decision is positively associated with preventive care and the use of seat belts suggesting that risk aversion is an important factor affecting insurance demand.

In another paper Cutler *et al.* [32] use data from the HRS and examine the relationship between risk reducing behaviours (such as smoking, drinking, job-mortality risk, etc.), risk occurrence and five insurance purchase decisions in the Unites States. They consider each market separately and find that people who engage in risky behavior, and then who are more risk tolerant, are systematically less likely to hold life insurance, acute private health insurance, annuities, long-term care insurance, and Medigap. Moreover, they show that this preference effect has different sign across markets, suggesting that heterogeneity in risk preference may be important in explaining the differential patterns of insurance coverage in various insurance markets.

The second related literature focuses on estimating risk preferences from observed choices. This is a vast and constantly growing literature which is hard to fully summarize here - for a review see Blavatskyy and Pogrebna [15]. In general these studies use individual observed choice obtained from survey data - sometimes with experimental module (e.g., Viscusi and Evans [116]; Evans and Viscusi [52]; Barksy *et al.* [10]; Dohmen *et al.* [111]) or laboratory or natural experiment experiment (Holt and Laury [74], Jullien and Salanié [82], Guiso and Paiella [67]) to estimate risk preference.

Barsky *et al.* [10] use survey responses to hypothetical situations from the HRS to construct a measure of risk preferences. They compare the measured risk tolerance with a set of risky behaviours and find that smoking, drinking, failing to have insurance, and holding stocks rather than Treasury bills are positively related with risk tolerance. Dohmen *et al.* [111] also find statistically significant evidence of relationship between financial and non-financial risk aversion on the basis of survey data. Guiso and Paiella [66] use household survey data to construct a direct measure of absolute risk aversion and find individual risk aversion having a considerable predictive power for a number of key household decisions such as choice of occupation, portfolio selection, moving decisions and exposure to chronic disease. Cutler and Glaeser [33] used a similar approach to investigate what the extent health-related behaviours are correlated and find that those individuals who choose to follow an healthy life style are also more likely to behave healthier in another context.

Another group of studies use data on insurance choice to analyse individual risk aversion (Cicchetti and Dubin [27], Sydnor [109]). In a recent paper Cohen and Einav [29] develop a structural econometric model to estimate risk preferences from data on deductible choices in auto insurance contracts. Their empirical strategy relies on modelling individual insurance purchase decision following the expected utility theory in which risk aversion parameter depends on unobserved characteristics and then compare variation in the deductible menus across individuals and their choices from these menus to estimate risk aversion in the sample. They find the existence of heterogeneity in risk preferences and that risk aversion is also related to sex and age. Each of these studies, however, examine risk aversion in a single insurance context. More recently another group of studies examined the insurance multicontext choice and focused on the stability of risk preferences across contexts.

This is the third stream of literature which studies multiple demand for insurance and whether risk preferences are invariant across risk domains. In general the principle of general component of risk preference has received considerable attention in the economic literature and in particular in behavioral economic studies, which mainly involve laboratory or natural experiments (for reviews, Kahneman [83]-[84]). Standard economic theory predicts that individual risk preferences are stable across decision contexts. This principle of invariance of risk preferences implies that multiple risky choices by the same economic agent should reflect the same degree of risk aversion even when decision is taken in different contexts. This principle has motivated a vast empirical research. Many studies found the existence of a common, but small, element of domain-general risk preferences (see for example Barsky *et al.* [10], Dohmen *et al.* [111], Kimball *et al.* [87]), while several other studies based on laboratory experiments and hypothetical money gables showed that context is the most important factor (Wolf and Pohlman [120]) or even that choice depends on whether questions are framed as a "gamble" or as "insurance" (Hershey *et al.* [72], Johnson *et al.* [78]).

Recently Barseghyan *et al.* [9] take an innovative approach to test generality of individual risk preference. Following Cohen and Einav [29] they use insurance company data to examine whether risk preferences are stable over a set of multiple insurance choices. In particular they test whether individuals' deductible choices in automobile and home insurance are consistent with the context-invariant risk preferences hypothesis. They find that some individuals are more risk averse in their home deductible choices than their auto deductible choices. Therefore, the hypothesis of stable risk preferences across domain is rejected by their data.

Einav *et al.* [49] focus on within-person correlation in the ordinal ranking of the riskiness of the choice an individual makes across different domains.

They use data on employee benefit choices for the U.S. workers at Alcoa.Inc regarding the 401(k) asset allocation and five different employer-provided insurance domains, that include health and disability insurance. Since they are mainly interested on the rank correlation within individuals across domains in their choice among options in a domain, their econometric strategy relies on a multivariate regression to estimate residual correlation between domains conditional on individual characteristics. Since they are mainly focused on risk preferences across domains, they use observable characteristics capturing individual predicted (by insurer) and ex-post risk to control whether conditional on these variables there is no residual correlation between insurance choices. However proxies may not capture perfectly individual risk and then the residual correlation could also indicate correlation in the unobserved risk rather than commonality of risk preferences. To address this issue they focus not only on residual correlations between insurance choices, but also on the correlation between insurance coverage and 401(k) portfolio allocation, which they claim to be uncorrelated with individual risk. They found a small effect of individual risk controls on the correlation pattern as well as a statistically significant residual correlation between 401(k) and insurance. Thus, they conclude that correlations are more likely to capture correlation in underlying risk aversion and that risk preferences are likely to be stable across domains.

Although our paper is closer in spirit with those of Barseghyan *et al.* [9] and Einav *et al.* [49], since we model multiple insurance purchase decisions and estimate how stable are risk preferences across these contexts, our approach differs substantially from two perspectives. First we study risk preference stability using survey data on insurance choices. Although information on insurance plans are more detailed in insurance company data, survey data offer a wide set of information over individual risk attitudes to bear risk in several contexts. Moreover survey data are more often employed by applied economists and it could be interesting to examine how an empirical appraisal based on residual correlation across insurance choices perform to study the stability of risk preferences. Second we exploit latent variable techniques, which allow to interpret and identify directly the residual correlation related to individual risk preference and that one potentially introduced by non-preference factors (such as context specificity, unpriced risk, etc.).

5.3 Data and Institutional background

Our analysis uses individual-level data from the fifth wave of the Health and Retirement Study (HRS). The HRS is a biennial survey targeting elderly Americans over the age of 50 and provide detailed information on insurance coverage, health status, life style and financial and socioeconomic status. We use these data to study four insurance purchase decisions among people older than 65 in 2002: in particular we study whether the individual has: a term life insurance, a Medicare supplemental coverage (Medigap), a long-term care insurance and an annuity. Previous theoretical and empirical studies model the demand for insurance as a function of individual risk aversion and individual risk. Since our main focus is to study how risk tolerance is related to the decision of holding any of these insurance plans and whether there exists an heterogenous patter of risk preferences across domains, we need to control for both predicted (by insurer) and unobserved heterogeneity in risk (adverse selection). Conditioning on the characteristics used in pricing insurance, which is the risk classification of insurer, and on the ex-post risk is crucial to identify the effect of risk aversion on the decision to purchase an insurance. For this purpose we follow previous studies on demand for insurance (see Cutler *et al.* [32], Finklstein and McGarry [58]) and exploit the dynamic structure of the data to track both predicted and actual individual riskiness in each domains (such as mortality, subsequent health care utilization, etc.). In addition since risk tolerance is not directly observed, we use a rich set of indicators on individual's characteristics and behaviours that has been shown being likely to capture individual risk aversion (see Barsky *et al.* [10], Kimball *et al.* [87]). After cleaning for missed (or inconsistent) observation and considering only those individuals who are at least 65 years old, the remaining sample size consists of 2488 observations. Descriptive statistics of the sample and variables' definition are reported in table A.1, while in the following subsections we describe the variables used to measure insurance coverage, individual risk and risk preferences.

5.3.1 Insurance

The first measure of insurance refers to whether an individual has a Medicare supplemental health insurance in 2002. This supplemental insurance is often named Medigap, since it is specifically designed to cover "gaps" of coverage left by Medicare public plans. These gaps include for example limitations in the coverage of health care services, high out-of-pocket expenses to Medicare beneficiaries and lack of a catastrophic cap expenditure. Since Medigap-private health insurance plan offer coverage only when people turn elder, we exclude from the sample all individuals who are younger than 65 in 2002. In addition we focus on individual who have deliberately purchased supplemental insurance as our interest is mainly on the demand for insurance (see Fang *et al.* [53]). Therefore we define an individual as having additional health insurance coverage (Medigap) if they purchased directly health insurance policy in addition to Medicare. As result we exclude those who received coverage by a former employer or spouse and who have free access by other public founded program such as Medicaid, CHAMPUS or CHAMPVA (Tri-care).

The second measure of insurance purchase decision we consider is the long-term insurance. Long-term care expenditure risk is one the greatest financial risks faced by the elderly in the US. This markets, differently from the Medigap insurance markets, is not subject to heavy regulation and then insurance companies are free to price contracts according to individual riskiness. We define an individual as having long-term insurance if the declare to be covered by long-term insurance during the year 2002.

Finally ours third and fourth insurance purchase decision are life insurance and annuity. We define an individual as having a life insurance or an annuity in the 2002 HRS if they answer positively to the question about these two coverage options. In the sample there is about 52% holding a supplemental health insurance, about 15% is covered by a long-term insurance, about 63% and 46% has respectively a life insurance and an annuity.

5.3.2 Risk Occurrence

The corresponding measures to control for predicted and ex-post risk occurrence change according to the insurance risk domain one considers.

Consider first our measures of predicted (by insurer) risk. These are controls for risk that we use in each insurance market. Which factors to include depends on the information insurers collect and use in pricing premiums. Clearly the insurance company defines the premium according to the predicted risk. We follow previous studies on demand for insurance to better define which variables to use as controls (see for example Cutler *et al.* [32], Cohen and Einav [29], Cohen and Spiegelman [30]).

In the supplemental health insurance market, Medigap companies use only individual age and sex to price contracts. This is so because by law there is a free enrolment period which lasts for six months from the first month in which people are both 65 years old and enrolled in Medicare. During this period Medigap cannot refuse any person even if there are pre-existing conditions and pricing is allowed only on the basis of age and sex. We therefore include only individual gender and age as dummy variables to control for predicted risk. In particular gender is measured by *fem* which takes 1 if individual is a female, while age is decomposed in four dummies, one for each five-years age band from 65 to 80. In the sample there is about 50% of female and on average individuals are 72 years old.

In the long-term care insurance market insurers collect with age and sex also many information on health status. Using a rich set of health related variables such as the number of diseases, the total number of limitations in the activities of daily living (ADL), the number of limitations with respect to instrumental activities of daily living (IADL) and a mental health index which measure any cognitive impairments,[1]we construct a synthetic binary indicators (*health status*) which takes 1 if individual has both a number of disease, ADL, IADL and impairments greater then the median individual.

In the life insurance market the premium depends mainly on age, gender and health status and on the size of policy the applicant is considering. Unfortunately we cannot observe the size of the policy and we include as control in addition to age and sex dummies mentioned above, a binary indicator of health status. Finally annuity classification risk is based solely on age and sex and therefore only these two variables are included as controls.

Let consider now our measures of ex-post risk. These measures should capture the residual unobserved heterogeneity which remains after conditioning on risk classification made by insurer. This residual association between risk oc-

[1]This mental health index is based on a score developed by the Center for Epidemiologic Studies Depression (CESD) and it is given by the differences between five "negative" indicators and two "positive" indicators. The negative indicators measure whether the respondent experienced depression or other mental impairments status. The positive indicators measure whether the respondent felt happy and enjoyed life, all or most of the time. Mehta *et al.* [94]) showed that this measure is associated with the existence of psychiatric problems.

currence and insurance purchase decision is often mentioned as source of adverse selection (Cohen and Spiegelman [30] and Einav *et al.* [47]). A standard measure of risk occurrence in the analysis of health insurance market is health care utilization. We employs the subsequent two waves (from 2004 and 2006) to track utilization. This is measured as the average number of hospital inpatients staying, doctor visits and outpatient services an individual used during the periods 2003-2006. Since the sample is based on elders, which are expected to register high level of health care utilization, and we want to capture the relative individual riskiness as compared with the sample, we construct a binary variable (*health care*) which takes 1 if the average number of services used by the individual is greater than the number of services used by the median. Clearly ex-post moral hazard can affect this measure, however it should be less effective when one considers subsequent utilization over a longer period and use it to model previous individuals' insurance choice decisions (see Cohen and Einav [29]).

For the life insurance market we use whether an individual is still alive in the subsequent two waves. The variable *mortality* equals 1 if the individual is deceased in the following waves, 0 otherwise. The ex-post risk measure for the annuity is clearly the opposite of that for life insurance, specifically whether the individual survives in the subsequent years. In the sample 6% of individual died in the subsequent years. Finally for the long-term insurance our measure is whether the individual had any nursing home entry in the following waves. The variable *nursing home* takes 1 if individual entered a nursing home, 0 otherwise. In the sample about 26% had a health care utilization greater than the median, about 8% of people used a nursing home and about 6% of individual died between years 2002 and 2006.

5.3.3 Risk Tolerance Indicators

Since individual risk tolerance is not directly observable, it is also not easy to measure. A standard strategy is to use proxy based on individual characteristics and behaviours which are likely to capture risk aversion. Thus we use the following set of indicators: job-based mortality risk, receipt of preventive health care, no risky portfolio choice, number of jobs the respondent reports having through job history, the subjective probability to leave over a certain age, wealth and a composite indicator of health related behaviours based on drinking, smoking and the body mass index. Barsky *et al.* [10] and Cutler and Glaeser [33] showed that most of these variables are significantly associated with individual risk aversion and then they can be effective to identify unobserved heterogeneity in risk preferences.

The first indicator is the job-based mortality risk. Following Cutler *et al.* [32] we derive the mortality rates from Viscusi [115]. He used data from the U.S. Bureau of Labor Statistics Census of Fatal Occupational Injuries to estimate job mortality rates by industry. We assign mortality rates in our HRS sample using industry-occupation cells (or occupation alone) and current job (if any), including self employment. If the respondent is not employed in the 2002 HRS, we then use the last available job information. Missing values for this variable are assigned if the individual has never held a job or if it is not possible to identify either job or industry code. Job mortality (*job-mort*) is then set equal to 1 if individual has job-mortality rate lower than the median.

Portfolio decision and the demand for risky assets are important dimensions of risk aversion. We define an individual as holding less risky assets if he/she has a total positive financial assets and the share of portfolios invested in Treasury bills and savings accounts is greater than those invested in stock. Therefore we set *norass* equal to 1 if individual has no risky assets, 0 otherwise. Notice that, since information on financial assets are collect at the household level and no information on asset ownership within the household are available, this measure could reflect risk preferences of the household rather than the individual. Although Barsky *et al.* [10] show that risk tolerance measure is positively, but not strongly, correlated within couples. In particular when the most knowledgable respondents is less risk averse than the second respondent in the couple, the share of portfolio in risky asset is lower, but the differences are not statistically significant.

Our third risk aversion indicator is derived by looking at the individual job history. Guiso and Paiella [66]-[67] show the existence of a negative relationsihip between the decision to leave a job and risk aversion. They argue that leaving a sure and known prospect for a new one unknown could imply incurring in new risks. Therefore we define our variable (*job-num*) equal to 1 if individual had a number of jobs lower than the median during his/her job history.

The fourth indicator refers to the self-reported probability of leaving to a given age. In the HRS the question varies according with the individual age. If the respondent is 75 or younger, than s/he is asked to report the probability to live to 75, while if he is older than 75, he/she is asked to report the probability of leaving to 100. Our indicator (*prlife*) is a binary variable which equals 1 if individual reports a probability greater than the median. Risk aversion could also be related with individual wealth since being more risk-averse can be translated into lower expected labour income (see for example Guiso and Paiella [66]-[67]). Individual wealth indicator is defined as a binary variable (*wealth*) which takes 1 if individual is in the top wealth quartile.

Finally we construct two binary indicators of individual health behaviours. The first one measures individual attitudes to health-related life styles. This indicator (*healthb*) takes 1 if the respondent has a normal body mass index (namely the BMI should have a score between 30 and 18), has less than three drinks per day and does not smoke. The second indicator which has been used in many other studies on risk and insurance (see Cutler*et al.* [32] and Finkelstein and McGarry [58]) refers to the fraction of gender-appropriate preventive health activity undertaken by individual. Preventive activities include: a flu shot, a blood test for cholesterol, a check of her breasts for lumps, a mammogram or breast x-ray, a Pap smear and a prostate screen. Our binary indicator (*preventive*) takes one if individual undertakes a fraction of gender-appropriate preventive health activity greater than the median. In the sample there are about 52% who does not smoke, drink and have a normal BMI; about 55% received sex-adjusted preventive care; about 54% has a job-based mortality risk lower than the median; 63% changed jobs less often than the median during the job history; 31% holds a share of no risk asset greater than the share of portfolio in stock; about 30% is in the top wealth quartile and 46% reports a subjective probability of leaving to a certain age greater than the median.

5.4 The Model

Our aim is to study the extent to which choices across insurance domains display a common risk aversion and test whether there is a residual correlation across domains related to non-preference factors. To this aim we use some recent developments in latent class analysis to model multiple choices, and test the residual association among choices after conditioning on covariates and latent variable (Huang and Bandeen-Roche [75], Bartolucci and Forcina [13] and Dardanoni, Forcina and Modica [36]).

Let I_j denote a binary variable which takes value 1 if an individual has purchased insurance in the risk domain j, with $j = 1, \ldots, J$. We want to study the following conditional expectations:

$$\begin{gathered} Pr(I_1 = 1 \mid \boldsymbol{w}_1, P) \\ \vdots \\ Pr(I_J = 1 \mid \boldsymbol{w}_J, P) \end{gathered} \tag{5.1}$$

where $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_J$ are vectors of individual observable and unobservable characteristics (such as individual risk) which affect insurance purchase decision in each of the J domains; while P represents individual risk preferences.

Clearly if one would control properly for $\boldsymbol{w}_j$ and P would be directly observable, then one could test directly the hypothesis of domain-general component (DGC) of risk preferences by examining any variations in the direct effect of P on the insurance purchase decision across domains. Suppose now that individual risk may be captured relatively well by observables proxy (e.g. insurer risk classification, subsequent risk occurrence rate, etc.). Since P is not observable, *how can we detect whether individual risk preferences are general?*

Consider that if risk preferences are specific and then depends mainly on the insurance context involved in the decision, then there is no unique underlying unobservable P affecting choices across domains. Thus P varies across domains and the system of equations (5.1) can be written as:

$$\begin{gathered} Pr(I_1 = 1 \mid \boldsymbol{w}_1, P_1) \\ \vdots \\ Pr(I_J = 1 \mid \boldsymbol{w}_J, P_J) \end{gathered} \tag{5.2}$$

This means that individual's willingness to bear risk in one insurance domain is different from his/her willingness to bear risk in another contexts. Einav *et al.* [49] propose to test the null of DGC of preferences by looking at the residual correlation between risk domains conditional on observables. Following this approach if the null of no correlation is reject then there are evidence of a sort of common element in the unobserved risk preferences.

An alternative is to assume that $P_1, \ldots, P_J$ are *discrete*, with P_j taking say m_k levels, $k = 1, \ldots, K$. This is a fairly innocuous assumption since any continuous variable can be approximated arbitrarily well by a discrete one. It implies that we can cross-classify $P_1, \ldots, P_K$ into a single discrete unobservable variable

U which takes say $m = m_1 \times \cdots \times m_K$ values, which identifies m heterogeneous "types". Differences among "types" are driven by different attitudes to bear risk across contexts.

To test then the DGC hypothesis suppose that for some arrangement of the M types U we have

$$\begin{gathered} Pr(I_1 = 1 \mid \boldsymbol{w}_1, U = 1) \leq \cdots \leq Pr(I_1 = 1 \mid \boldsymbol{w}_1, U = m) \\ \vdots \\ Pr(I_J = 1 \mid \boldsymbol{w}_J, U = 1) \leq \cdots \leq Pr(I_J = 1 \mid \boldsymbol{w}_J, U = m) \end{gathered} \tag{5.3}$$

This means that each variable P_j, with $(j = 1, \ldots, J)$, has a monotonic effect on the insurance purchase decision across domains. Note that if equalities do not hold for some unobserved "types", say for example that $Pr(I_1 = 1 \mid \boldsymbol{w}_1, U = 1) \leq \cdots \geq Pr(I_1 = 1 \mid \boldsymbol{w}_1, U = M))$, then individual has different attitude to bear risk in a context as compared with his/her peer in another context. The simple idea is the following. If risk preferences are general then there is a one-dimensional latent variable, representing unobserved types with different attitudes toward risk, which affect each insurance purchase decisions. Note in fact that types represent different attitudes to buy insurance and then different risk preferences. Under the null of DGC each type should always buy the same amount of insurance in each context as compared with another type, and then the same pattern on insurance purchase decision should be observed. Suppose for example three unobserved types with type one buying more health insurance than type two and the same between type two and three. Suppose also that this pattern holds also for the life insurance. If it so then there is an unidimensional latent variable, representing the order between types, having a monotonic effect on the two insurance purchase decisions.[2] Let to analyse how this procedure can be implemented empirically.

[2]This strategy relies on the idea that proxy variables of risk capture relatively well insurance purchase attitudes related to individual risk. To the extent that unobserved risk is not captured, abstracting from it will likely introduce bias in the identification of P that needs to be controlled. However applied economic literature studying domain-generality of an individual's risk preferences and insurance markets (see Cutler *et al.* [32], Cohen and Einav [29] and Einav *et al.* [49]) showed that using individual predict (by insurer) risk and subsequent risk occurrence are effective in capturing unobserved individual risk. However a possible solution in our framework, which still needs to be further investigated, is to set a model with two distinct unobservables, say U_1 and U_2, capturing individual risk preferences and the residual unobserved heterogeneity in risk occurrence.

5.4.1 Empirical strategy

Following standard models in the literature on insurance demand (Cohen and Einav [29], Cutler *et al.* [32], Einav *et al.* [49]), $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_J$ include observable characteristics designed to capture the risk classification used by insurers, which we denote with $\boldsymbol{x}_j$, and a set of variables ($\boldsymbol{r}_j$) which proxy individual subsequent risk. This set of covariates is an important confounding factor, since insurance demand is usually driven by both risk and risk aversion and then actual risk may cause potential residual correlation across domains. Assuming additive separability we can rewrite the equation system (5.1) as:

$$\begin{aligned} Pr(I_1 = 1 \mid \boldsymbol{w}_1, P) &= F(\boldsymbol{x}'_1\boldsymbol{\beta}_1 + \boldsymbol{r}'_1\boldsymbol{\gamma}_1 + \boldsymbol{v}'_1\boldsymbol{\delta}_1) \\ &\;\;\vdots \\ Pr(I_J = 1 \mid \boldsymbol{w}_1, P) &= F(\boldsymbol{x}'_J\boldsymbol{\beta}_J + \boldsymbol{r}'_J\boldsymbol{\gamma}_J + \boldsymbol{v}'_J\boldsymbol{\delta}_J) \end{aligned} \tag{5.4}$$

where F denotes the appropriate link function and $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_J$ are vectors of unobservables capturing residual heterogeneity in risk preferences. To estimate the equation system (5.4) and test the hypothesis of DGC which is the focus of the analysis, we consider two possible models: a multivariate regression model as proposed by Einav *et al.* [49] and extended LCA model.

5.4.2 Multivariate probit regression

In a recent paper Einav *et al.* [49] study the DGC hypothesis examining the correlation structure of the error terms in a multivariate regression. Following this approach, let the link function F be standard normal, so that we can equivalently rewrite the system (5.4) as:

$$\begin{aligned} I_1 &= 1\big(\boldsymbol{x}'_1\boldsymbol{\beta}_1 + \boldsymbol{r}'_1\boldsymbol{\gamma}_1 + \boldsymbol{v}'_1\boldsymbol{\delta}_1 + \epsilon_1\big) \\ &\;\;\vdots \\ I_J &= 1\big(\boldsymbol{x}'_J\boldsymbol{\beta}_J + \boldsymbol{r}'_J\boldsymbol{\gamma}_J + \boldsymbol{v}'_J\boldsymbol{\delta}_J + \epsilon_J\big) \end{aligned} \tag{5.5}$$

where $\epsilon_1, \ldots, \epsilon_J$ are independent standard normal errors. If we let $\eta_j = \boldsymbol{v}'_j\boldsymbol{\delta}_j + \epsilon_j$ in each domain and assume that $(\eta_1, \ldots, \eta_J)$ are distributed as a multivariate normal with standard margins and correlation coefficient equal to ρ, we get the multivariate probit:

$$
\begin{aligned}
I_1 &= 1\big(\boldsymbol{x}'_1\boldsymbol{\beta}_1 + \boldsymbol{r}'_1\boldsymbol{\gamma}_1 + \eta_1\big) \\
&\vdots \\
I_J &= 1\big(\boldsymbol{x}'_J\boldsymbol{\beta}_J + \boldsymbol{r}'_J\boldsymbol{\gamma}_J + \eta_J\big)
\end{aligned}
\tag{5.6}
$$

The multivariate probit is relatively easy to estimate and provide the baseline correlations to evaluate how general are risk preferences across insurance purchase decisions. However it does rely on multivariate normality to achieve parameters' identification, and does not allow to control directly whether conditional on individual risk preferences there exists a residual correlation between choices indicating the residual role played by the specific context.

5.4.3 Extended LCA

As mentioned above an alternative way to control for the residual unobserved heterogeneity in risk preference U is by identifying a finite number of unobservable "types" M, which differ in their attitudes to bear risk in different contexts. Thus, the equation system (5.4), which account for the unobserved U can be written as:

$$
\begin{aligned}
I_1 &= \textstyle\sum_{u=1}^{m} \alpha_u^{I_1} U_u + \boldsymbol{x}'_1\boldsymbol{\beta}_1 + \boldsymbol{r}'_1\boldsymbol{\gamma}_1 + \eta_1 \\
&\vdots \\
I_J &= \textstyle\sum_{u=1}^{m} \alpha_u^{I_J} U_u + \boldsymbol{x}'_J\boldsymbol{\beta}_J + \boldsymbol{r}'_J\boldsymbol{\gamma}_J + \eta_J
\end{aligned}
\tag{5.7}
$$

where $U_1, \ldots, U_m$ denote the set of m dummy variables indicating "latent type" membership. Thus, the coefficients $\alpha_u^{I_j}$ in each equations can be interpreted as random intercepts with a nonparametric discrete specification.

To identify unobserved risk preferences U, we exploit in addition to observed individual purchase decisions, which are of main interest in our framework, a set of auxiliary equations that are used as indicators of U and then capture individual attitudes to bear risk. Using a standard logit link in equations (5.7), we estimate the model:

$$
\begin{aligned}
\lambda^{I_1} &= \textstyle\sum_{u=1}^{m} \alpha_u^{I_1} U_u + \boldsymbol{x}'_1\boldsymbol{\beta}_1 + \boldsymbol{r}'_1\boldsymbol{\gamma}_1 \\
&\vdots \\
\lambda^{I_J} &= \textstyle\sum_{u=1}^{m} \alpha_u^{I_J} U_u + \boldsymbol{x}'_J\boldsymbol{\beta}_J + \boldsymbol{r}'_J\boldsymbol{\gamma}_J
\end{aligned}
\tag{5.8}
$$

together with the class membership probabilities $Pr(U = u)$ which can be written in terms of *adjacent logits* as

$$\log\left(\frac{Pr(U=u+1)}{Pr(U=u)}\right) = \lambda_u^U = \alpha_u^U \qquad u = 1, \ldots, m-1 \tag{5.9}$$

and the following system which can be considered instrumental for identifying U:

$$\begin{aligned} \lambda^{H_1} &= \textstyle\sum_{u=1}^m \alpha_u^{H_1} U_u \\ &\vdots \\ \lambda^{H_T} &= \textstyle\sum_{u=1}^m \alpha_u^{H_T} U_u \end{aligned} \tag{5.10}$$

Note that the the system of equations (5.10) is used to capture and identify individual unobserved types which differ in terms of risk preferences. Thus it can be considered auxiliary to the simultaneous equation system (5.8).

In addition to equations (5.8-5.10) we also allow residual correlation among insurance purchase decisions to capture conditional on U potential non-preference factors - such as context-specificity - which may introduce correlation between choices. This can be written as:

$$\log\left(\frac{Pr(I_j=0,I_k=0)Pr(I_j=1,I_k=1)}{Pr(I_j=1,I_k=0)Pr(I_j=0,I_k=1)}\right) = \lambda_{I_j,I_k} = \alpha_{I_j,I_k} \tag{5.11}$$

with $j \neq k$ and $j,k = 1,\ldots,J$. This means to estimate one parameter for each of the $\binom{J}{2}$ combinations of insurance purchase decision. Thus (5.11) allows to control for residual correlation among risk domains introduced by non-preference factors - for example some choices may be"closer" in context, such as health and disability insurance purchase decision (Einav *et al.* [49]). Note that U is of main focus to test the DGC hypothesis since it represents individual risk preference. On the contrary λ is only included to capture any residual association unrelated with U.

Within the model defined by equations ((5.8)-(5.11)),

- the null hypothesis of DGC of individual risk preferences (that is equation (5.3)) can be viewed as testing the null hypothesis that there is a underlying unidimensional unobservable variable U such that choices are monotonically dependent on it. This can be implemented by setting a system of linear inequalities as explained for example in Bartolucci and Forcina [12]. Techniques of order restricted inference can be used to show that the likelihood ratio test statistic for the monotonicity null is asymptotically distributed as a mixture of chi-squared distributions (see Gourieroux and Monfort [64] for a general exposition, Dardanoni and Forcina [34] for an explanation of how the mixing weights can be calculated by simulations, and Kodde and Palm [88] for bounds on the test distribution).

- the null hypothesis of absence of residual heterogeneity related to potential non-preference factors (since they are unrelated to U) can be tested by imposing for each of the $\binom{J}{2}$ α parameters the restriction that α_{I_j,I_k} is not statistically different from zero. This can be implemented with a standard t-test statistic.

5.5 Results

In this section we first examine results from a multivariate binary probit model for the probability of purchase Medicare supplemental health insurance, life insurance, long-term care insurance and annuity. We then analyse in the subsequent section result from the extend LCA which both identifies unobserved types with different attitudes to bear risk across domains and allow residual correlation between insurance choices to capture non-preference factors.

5.5.1 Multivariate Regression

Tables A.2 and A.3 present respectively the estimated coefficients of controls and correlation terms from the baseline multivariate probit regression suggested by Einav *et al.* [49] and described above in equation (5.6). Let consider first the determinants of supplemental health insurance purchase decision. Table A.2 reveals that the probability of enrolling in a supplementary insurance plan increases with age and sex. Not surprisingly people who are more risky and then tend to use more health care resources - for example hospital inpatient stays, doctor visits and outpatient services - are also significantly more likely to buy additional coverage. Therefore our result on ex-post risk occurrence confirms previous analysis, which found the existence selection effect in the Medigap market related also to private information on individual actual risk (see for example Fang *et al.* [53], Ettner [51]).

The probability to purchase a long-term care insurance is also increasing with individual age, but the effect is not statistically significant, and with health status. In particular those who report having more diseases and physical impairments in the daily living activities (measured by ADL and IADL) are also more likely to hold a long-term insurance plan. As expected ex-post utilization of any nursing home in the two waves following 2002 HRS increases the probability to buy insurance, but surprisingly this effect is not statistically significant.

Taking a glance at life insurance results, table A.2 shows that people who are female and married are also more likely to purchase this type of insurance. On the contrary ex-post measured risk does not seem to have a statistically significant effect although the estimated coefficient has the expected sing.

Finally annuity purchase decision is positively related with age, but negatively with individual gender. Although there is not a clear effect between gender and the probability of having an annuity, in a recent paper Agnew *et al.* [3] find that women are more likely to buy annuity than man, since gender differences may indicate also differences in risk aversion. However, if risk aversion and predicted risk are driving the decision to choose annuities, after controlling for these two factors, gender differences should not affect the annuity decision. Ex-post measured risk in this market has a negative and statistically significant effect. In particular those who are more likely to live longer are also more likely to hold an annuity, suggesting that individual private information on mortality risk is an important sources of asymmetric information in this market after conditioning on predicted (by insurer) individual risk (Cohen and Spiegelman [30]).

Consider now the estimated correlations between insurance purchase decisions. In all of the pairs reported in table A.3, we can reject - at least at 10% statistical significance level - the null hypothesis of correlation being zero, except for correlations between health and long-term care insurance with life insurance. Following Einav *et al.* [49], this result can be interpreted as evidence that we can reject the null of no domain general component of choice. Viewed alternatively, this means that one's coverage choice in any of the other domains is predictive of individual choice in a given domain. In particular the magnitude of the correlations generally seems to be higher for those insurance purchase decision which seems to be "closer", for example long-term care is more correlated with Medicare supplemental health insurance rather than annuity, and on the contrary life insurance is correlated with annuity. A possible limitation of this approach when only insurance choices are considered is that correlations across domains could reflect not just unobserved risk preference, but also unobserved correlation introduced by unpriced risk. Note that predicted and realized (ex-post) risk may not perfectly capture heterogenous individual actual risk and then it could be hard to interpret whether these correlation between insurance (risk) domains reflect systematic differences in each of these domains or rather unobserved preferences.

5.5.2 Results from the Extended LCA Model

We start by estimating the system of equations (5.8)-(5.11) under different numbers m of latent classes. Maximum likelihood estimation is performed by a *EM* algorithm. In particular while in the E step the posterior probability of latent class M given the observed configuration of insurance choices and auxiliary indicators is computed, in the M-step the likelihood function is maximized and further refined in each iteration by the E-step. More details on estimation procedure of parameters α and β can be derived by looking at Dardanoni, Forcina and Modica [36] and at Bartolucci and Forcina [13].[3] For completeness we report in tables A.4-A.10 model's estimated parameters under different number of latent classes, namely $m = 2, 3, 4$. Table A.4 reports the maximized log-likelihood $L(\psi)$, the Schwartz's Bayesian Information Criterion $BIC(\psi) = -2L(\psi) + vlog(n)$, where n denotes sample size and v is the number of parameters. BIC seems to indicate that three LC are adequate to represent the unobserved heterogeneity U. A glance at all tables reveals also that estimated α, β and correlation coefficients do not seem to vary substantially with respect to the number m of latent classes specifications. For sake of brevity we will discuss mainly results obtained under $m = 3$ latent classes. Calculating the types membership probabilities reported in table A.5, about 50% of individuals are of type 1, while 30% and 20% are of type 2 and 3 respectively.

To understand what these types indicate, let consider the estimated probabilities reported in table A.6, obtained using the α parameters of tables A.7 and A.8. Type 3 individuals are those who are on average about three times more likely to buy any Medicare supplemental health insurance, long-term insurance, life insurance and annuity than type 1. The picture does not change substantially comparing type 3 with type 2, although the latter seems to be more likely to hold long-term insurance and annuity than type 3 individuals. Therefore a first glance at Panel A of table A.6 shows that types differ in the attitudes to purchase insurance. In particular conditional on predicted and ex-post realized risk, type 3 individuals are more risk averse that type 1 since they are always less

[3] We are grateful to Antonio Forcina for kindly providing the Matlab code for the estimation.

prone than type 1 individuals to bear risk in any of the four insurance domains.

This result is also supported by looking at Panel B of table A.6, which reports the relationships between no risky behaviours and unobserved types. The table reveals that estimated probabilities to perform risky behaviours or characteristics increase with m. In particular people who hold T-bills rather than stock in their own financial portfolio, who change job less frequently, have a mortality rate of the individual's industry-occupation cell lower than the median rate, who have a normal body mass index and do not smoke and drink, who invest into health risk prevention activities and have a life expectation greater than the median are more likely to be of type 3 rather than any other unobserved types. Not surprisingly type 3 individuals are less "wealthy" than type 2, which is in line with the idea that more risk averse individuals are relatively less wealthy than others (see for example Barsky *et al.* [10] Guiso and Paiella [67]). The pattern we find is consistent with other studies, such as Barsky *et al.* [10], who checked the external validity of some risk tolerance measures using risky behaviours indicators. Therefore results indicate two main conclusions. First, the picture which emerges from the estimated probabilities is that, after conditioning on individual predicted and ex-post realized risk there exists an important source of heterogeneity in the underlying risk preferences represented by the latent types, which plays an important role in the insurance purchase decisions. This result is consistent with recent studies (Cohen and Einav [29], Barseghyan *et al.* [9]and Einav *et al.* [49]), which found heterogeneity in risk preferences being more important than heterogeneity in risk to explain how heterogenous are insurance coverage choices.

Second the three unobserved types which differ in their attitudes to bear risk, and then in how individual are risk adverse seem to follow the same pattern across domains. In particular those individuals who are less risk averse in one domain are also more likely to bear risk in any other domains. For example, type 1 is on average less likely to perform risk reducing behaviours than type 2, who is at the same time less likely than type 3 individuals. This pattern between no risky behaviours and unobserved types seems to hold also for insurance choice, providing evidence against the hypotheses of no domain-general component if the insurance choices. In other words, after conditioning on predicted and realized risk, it seems there is a single latent variable which is common to each insurance choice domains.

The question naturally arises then whether this pattern in insurance choices is due to sampling variations, or rather to the presence of a single latent variable that conditional on predicted and realized risk has a common effect on insurance choice domains. The testing procedures described by equation (5.3) can however be employed to formally test the unidimensionality of latent variable. The LR test statistic for the model under the null that $\alpha_1^{I_1} \leq \alpha_2^{I_1} \leq \alpha_3^{I_1}$ is equals to 9.15. Since U has three levels ($m = 3$) and the insurance choices we consider are four, the conservative 1% critical value with 8 df is equal to 25.370 (Kodde and Palm ([88], page 1246); thus, the null of domain general component cannot be rejected indicating the existence of a single underlying unobservable variables which each insurance purchase decisions.

Although the existence of a general commonality of domain risk preferences is not really surprisingly, it is interesting to note that after conditioning on individual unobserved types and individual risk, there still exists a sort of non-preference based correlation (related for example to context specificity), which renders some insurance choices more related than others. In fact taking a quick glance at table A.10 reveals that correlations are statistically different from zero in most of the cases and that are greater in magnitude when choices are "closer" - for example long-term insurance is more correlated to Medicare supplemental insurance rather than annuity, while life insurance is mainly correlated with annuity. This result has also been found by Einav *et al.* [49] and support the idea that choice is driven both by context and by how individuals are risk averse in general. However the existence of this residual correlation between responses can also indicate the existence of unpriced risk not captured by risk occurrence proxies. The simple idea is that if risk proxies do not fully capture individual risk then individuals, say, with higher health risk tend to purchase more health related insurance coverage. Notice that these findings are compatible with the DGC, since risk preferences can be common across domains, but some choices can be more correlated each other due to the presence of non risk preference factors (e.g. similarity in the decision context, unpriced risk factors, etc.).

Finally let us to consider the effect of predicted (by insurer) and realized ex-post risk in each insurance equation. Table A.9 shows a similar pattern of the effects of risk controls on insurance purchase decisions. In particular age and gender have a positive and statistically significant effect on the decision to buy Medicare additional coverage and annuity. Ex-post risk has always the expected sign. Interestingly if compared with the multivariate probit the dummy variable indicating whether an individual died in the succeeding two waves has positive and now statistically significant effect in the decision to purchase life insurance and negative for annuity. Therefore conditional on risk preferences, ex-post realized risk proxies indicate how important could be the role of private information on individual risk to determine the insurance choices which as been documented in several other studies (see for a review Cohen and Spiegelman [30] and Einav *et al.* [47]).

5.6 Conclusion

In this paper we examined the relationship between unobserved risk preferences and insurance purchase decision and in particular how general are preferences for risk across domains. Standard economic theory generally assumes that individuals take decisions over a set of risky domains according to their own risk preference which is stable across decision contexts. This assumption of context-invariant risk preference has motivated a large literature in microeconometrics and has caused debate in the literature concerning its validity. There is a large literature in psychology and behavioral economics which uses experimental lab test to claim that risk preferences are mainly related to context, and that decisions are not related to each other by any general risk domain components. To study this issue in the framework of multiple demand for insurance, we follow a recent stream of papers by Cohen and Einav [29], Barseghyan *et al.* [9] and Einav *et al.* [49] which focus on how general are risk individual preferences.

In particular we start following an innovative approach proposed by Einav *et al.* [49] that used residual correlation across insurance domains. Conditioning on predicted (by insurer) and ex-post risk to test whether individuals show the same willingness to bear risk across domains.

In our setting we model the correlations between insurance choices using a latent class analysis. Conditioning on predicted and realized risk we exploit LCA to identify individual risk aversion throughout a set of auxiliary variables which are likely to capture individual risk preferences. In addition we also allow for residual correlation between insurance choices in order to capture any residual correlation related to non-preference factors.

Using data from the Health and Retirement Study and a rich set of information on individual about risk and life-style behaviours, we study four insurance purchase decision: Medicare supplemental health insurance, long-term insurance, life insurance and annuity. In our data we identify three unobserved types which differ in terms of risk aversion. We find that individual who tend to buy a certain type a of insurance, say health insurance, are also more likely to buy insurance in another context, for example long-term care insurance. This can be interpreted as source of commonality in how individuals bear risk across domains. Thus our results provide an additional piece of evidence against the absence of domain general component of risk preferences, although context plays an important role in risky decision since insurance choices who are "closer" in context are also more correlated conditional on unobserved risk preferences. Therefore heterogeneity in risk preferences is also an important factor to consider in addition to heterogeneity in risk when individual choices on insurance coverage are examined. The question of what drives this heterogeneity and why the residual domain-specificity correlation still plays an substantial role remains an interesting question for further exploration.

Appendix A

Tables

Table A.1: Sample Characteristics and Variable Definition

Variable	Definition of Binary Variables	Mean
Insurance Status		
Sup. Health Ins.	1 = enrolled in any health insurance (Medigap).	0.520
long-term Ins.	1 = enrolled in any long-term insurance.	0.148
Life Ins	1 = covered by life insurance.	0.636
Annuity	1 = has an annuity.	0.459
Controls used by insurer to assess risk		
age65	1 = aged between 66 and 70 years.	0.387
age70	1 = aged between 71 and 75 years.	0.277
age75	1 = aged between 76 and 80 years.	0.158
age80	1 = older than 80 years.	0.073
fem	1 = female.	0.553
mar	1 = married.	0.610
health status	1 = # of disease, ADL and IADL	0.493
Ex-post Risk Indicators		
mortality	1 = died in the subsequent years 2004-2006.	0.063
health care	1 = used health care service during years 2004-2006.	0.262
nursing home	1 = entered in any nursing home in the years 2004-2006.	0.078
Risk Preference Indicators		
healthb	1 = does not smoke, has a normal weight and no drinking problems.	0.518
preventive	1 = received sex-adjusted preventive care.	0.551
job-mort	1 = has a job-based mortality risk lower than the median.	0.531
job-num	1 = has a number of jobs lower than the median.	0.632
norass	1 = holds no risk asset such as T-bills.	0.312
weatlh	1 = in the top wealth quartile.	0.301
prlife	1 = subjective life expectation grater than the median.	0.464

Table A.2: Multivariate Probit Model's Estimated Parameters of predicted and realized risk

Variables	Sup. Health Ins.		Long-term Ins.		Life Ins.		Annuity	
	Coef.	St.Er.	Coef.	St.Er.	Coef.	St.Er.	Coef.	St.Er.
fem	0.1941	(0.0509)	0.0401	(0.0642)	-0.2140	(0.0553)	-0.1470	(0.0510)
age65	0.1012	(0.0883)	0.1630	(0.1110)	0.1240	(0.0900)	0.1970	(0.0888)
age70	0.1740	(0.0920)	0.0871	(0.1160)	-0.0487	(0.0934)	0.2090	(0.0925)
age75	0.2265	(0.1010)	0.2100	(0.1250)	0.0787	(0.1030)	0.0385	(0.1010)
age80	0.5253	(0.1230)	0.0601	(0.1570)	-0.1020	(0.1240)	-0.3150	(0.1240)
mard02			0.1520	(0.0665)	0.1390	(0.0568)		
health			0.1210	(0.0609)	0.1090	(0.0521)		
nursing home			0.1540	(0.1130)				
health care	0.5320	(0.1880)						
mortality					-0.1110	(0.1060)	-0.2748	(0.1041)
constant	-0.2370	(0.0850)	-1.3650	(0.1220)	0.3011	(0.1000)	-0.1410	(0.0845)

Robust standard errors in brackets.

Table A.3: Multivariate Probit Model's Estimated Correlation Terms Controlling for Predicted and Realized Risk

Variables	Sup. Health Ins.		Long-Term Ins.		Life Ins.	
Long-Term Ins.	0.3121	(0.0392)				
Life Ins.	0.0458	(0.0320)	0.0611	(0.0384)		
Annuity	0.2180	(0.0318)	0.2810	(0.0393)	0.0572	(0.0323)

Robust standard errors in brackets.

Table A.4: Model Selection Criteria for System of Equations (5.8)-(5.11)

	Number of Latent Classes		
	2LC	3LC	4LC
$L(\psi)$	-17166.44	-17110.02	-17092.10
$BIC(\psi)$	34778.580	34759.58	34817.57
$\#ofparmaters$	57	69	81

Table A.5: Estimated Class Membership Probabilities

	2LC	3LC	4LC
α_1^U	0.5051	0.4985	0.2242
α_2^U	0.4949	0.2970	0.2585
α_3^U	.	0.2045	0.2306
α_4^U	.	.	0.2867

Table A.6: Estimated Probabilities of Extended LC Model

	2LC		3LC			4LC			
	M=1	M=2	M=1	M=2	M=3	M=1	M=2	M=3	M=4
Panel A: Main Eq.									
Sup. Health Ins.	0.2464	0.5297	0.2625	0.5167	0.6437	0.7439	0.2181	0.9027	0.7434
Long-Term Ins.	0.0285	0.1411	0.0274	0.1588	0.0992	0.6379	0.5433	0.4501	0.5948
Life Ins.	0.6875	0.6627	0.7021	0.6269	0.7439	0.4851	0.6129	0.5675	0.5358
Annuity	0.1911	0.8681	0.2181	0.9027	0.7434	0.5371	0.6184	0.5991	0.5327
Panel B: Aux. Ind.									
norass	0.1682	0.4383	0.1542	0.3879	0.5368	0.3879	0.5368	0.4934	0.6151
job-mort	0.4901	0.5727	0.4934	0.5015	0.6150	0.5005	0.5974	0.5946	0.7699
job-num	0.6131	0.6511	0.5974	0.5946	0.7699	0.0438	0.6928	0.3573	0.5029
weatlh	0.0431	0.5637	0.0438	0.6928	0.3573	0.4052	0.7171	0.4581	0.6196
healthb	0.5236	0.5116	0.4052	0.5029	0.717	0.6781	0.3887	0.3061	0.8782
preventive	0.4627	0.6412	0.4581	0.6196	0.678	0.2625	0.5167	0.6437	0.0274
prlife	0.4287	0.5005	0.3887	0.3060	0.8782	0.1588	0.0992	0.7021	0.6269

Table A.7: Estimated Intercepts α of Equation System (5.8)

Insurance	2LC		3LC		4LC	
Choice	Coef.	St.Er.	Coef.	St.Er.	Coef.	St.Er.
Sup. Health Ins.						
$\alpha_1^{I_1}$	-1.1177	(0.1811)	-1.0333	(0.1828)	-1.0398	(0.231)
$\alpha_2^{I_1}$	0.119	(0.1757)	0.0668	(0.1930)	-1.0457	(0.223)
$\alpha_3^{I_1}$			0.5913	(0.2390)	0.0211	(0.206)
$\alpha_4^{I_1}$					0.4525	(0.218)
Long-Term Ins.						
$\alpha_1^{I_2}$	-3.5288	(0.3319)	-3.5690	(0.3350)	-3.5873	(0.494)
$\alpha_2^{I_2}$	-1.8071	(0.2828)	-1.6670	(0.2966)	-3.4400	(0.417)
$\alpha_3^{I_2}$			-2.2066	(0.3295)	-1.5152	(0.306)
$\alpha_4^{I_2}$					-2.2834	(0.322)
Life Ins.						
$\alpha_1^{I_3}$	0.7884	(0.1945)	0.8573	(0.1973)	0.5276	(0.267)
$\alpha_2^{I_3}$	0.6754	(0.1942)	0.5188	(0.2103)	1.3667	(0.276)
$\alpha_3^{I_3}$			1.0662	(0.2409)	0.4733	(0.230)
$\alpha_4^{I_3}$					1.2790	(0.238)
Annuity						
$\alpha_1^{I_4}$	-1.4429	(0.2878)	-1.2770	(0.2582)	-0.6759	(0.354)
$\alpha_2^{I_4}$	1.8836	(0.2937)	2.2273	(0.3295)	-3.0203	(1.496)
$\alpha_3^{I_4}$			1.0637	(0.2853)	2.6705	(0.485)
$\alpha_4^{I_4}$					0.7443	(0.295)

Table A.8: Estimated Intercepts α of Equation System (5.10)

Indicators	2LC		3LC		4LC	
	Coef.	St.Er.	Coef.	St.Er.	Coef.	St.Er.
norass						
$\alpha_1^{H_1}$	-1.5986	(0.0955)	-1.7019	(0.1068)	-1.6185	(0.222)
$\alpha_2^{H_1}$	-0.2481	(0.0685)	-0.4563	(0.1041)	-2.0098	(0.286)
$\alpha_3^{H_1}$			0.1475	(0.1529)	-0.5861	(0.13)
$\alpha_4^{H_1}$					0.1173	(0.147)
job-mort						
$\alpha_1^{H_2}$	-0.0397	(0.0634)	-0.0265	(0.0645)	-0.1626	(0.148)
$\alpha_2^{H_2}$	0.2928	(0.065)	0.0021	(0.1340)	0.0952	(0.123)
$\alpha_3^{H_2}$			0.4682	(0.1013)	0.6369	(0.131)
$\alpha_4^{H_2}$					-0.0227	(0.114)
job-num						
$\alpha_1^{H_3}$	0.4603	(0.0647)	0.3945	(0.0664)	-0.0041	(0.179)
$\alpha_2^{H_3}$	0.6233	(0.0672)	0.3831	(0.1033)	0.7116	(0.158)
$\alpha_3^{H_3}$			1.208	(0.1847)	0.4427	(0.119)
$\alpha_4^{H_3}$					0.9398	(0.131)
weatlh						
$\alpha_1^{H_4}$	-3.1034	(0.3334)	-3.0839	(0.3201)	-2.2608	(0.375)
$\alpha_2^{H_4}$	0.2561	(0.0883)	0.813	(0.1782)	-4.3692	(1.934)
$\alpha_3^{H_4}$			-0.5871	(0.1761)	1.0669	(0.251)
$\alpha_4^{H_5}$					-0.5535	(0.163)
healthb						
$\alpha_1^{H_5}$	0.0945	(0.0632)	-0.3837	(0.1213)	-0.3529	(0.174)
$\alpha_2^{H_5}$	0.0465	(0.0639)	0.0117	(0.0663)	0.2317	(0.139)
$\alpha_3^{H_5}$			0.9297	(0.2011)	-0.4121	(0.136)
$\alpha_4^{H_5}$					0.6618	(0.14)
preventive						
$\alpha_1^{H_6}$	-0.1493	(0.0649)	-0.1679	(0.0666)	-0.5144	(0.184)
$\alpha_2^{H_6}$	0.5805	(0.0689)	0.4879	(0.0988)	0.0425	(0.136)
$\alpha_3^{H_6}$			0.7446	(0.1486)	0.5798	(0.122)
$\alpha_4^{H_6}$					0.634	(0.12)
prlife						
$\alpha_1^{H_7}$	-0.2872	(0.0641)	-0.4527	(0.0763)	-1.9279	(0.809)
$\alpha_2^{H_7}$	0.002	(0.0641)	-0.8191	(0.1881)	0.2918	(0.268)
$\alpha_3^{H_7}$			1.9757	(0.6201)	-0.8432	(0.194)
$\alpha_4^{H_7}$					1.1622	(0.255)

Table A.9: Extend LC Model Estimated β Parameters of Predicted and Realized Risk

Variables	2LC		3LC		4LC	
	Coef.	St.Er.	Coef.	St.Er.	Coef.	St.Er.
Sup. Health Ins.						
age65	0.2783	(0.1683)	0.2718	(0.1701)	0.2746	(0.169)
age70	0.3302	(0.1743)	0.1436	(0.176)	0.1306	(0.175)
age75	0.3984	(0.1885)	0.1485	(0.1904)	0.1354	(0.189)
age80	0.7782	(0.2099)	0.4828	(0.2105)	0.478	(0.210)
fem	0.4053	(0.0878)	0.4018	(0.0888)	0.3921	(0.088)
health care	0.1334	(0.0984)	0.1309	(0.0994)	0.1357	(0.099)
Long-Term Ins.						
age65	0.4264	(0.2495)	0.4284	(0.2546)	0.4385	(0.255)
age70	0.2059	(0.2595)	0.2858	(0.2637)	0.2728	(0.264)
age75	0.3747	(0.2770)	0.5004	(0.2800)	0.4957	(0.281)
age80	0.1070	(0.3145)	0.2152	(0.3163)	0.2289	(0.316)
fem	0.1384	(0.1279)	0.1409	(0.1285)	0.1173	(0.128)
mard02	0.2297	(0.1328)	0.2179	(0.1334)	0.2219	(0.133)
health	0.2387	(0.1197)	0.2367	(0.1203)	0.2323	(0.120)
nursing home	0.3454	(0.2128)	0.3400	(0.2126)	0.3385	(0.213)
Life Ins.						
age65	-0.0446	(0.1672)	-0.0457	(0.1702)	-0.0535	(0.177)
age70	-0.3185	(0.1719)	-0.4209	(0.1745)	-0.6164	(0.181)
age75	-0.1118	(0.1874)	-0.2561	(0.1894)	-0.5060	(0.195)
age80	-0.5510	(0.2038)	-0.7161	(0.2064)	-0.9822	(0.214)
fem	-0.3621	(0.0903)	-0.3693	(0.0912)	-0.4011	(0.094)
mard02	0.2206	(0.0917)	0.2314	(0.0924)	0.2370	(0.095)
mort	0.1830	(0.0849)	0.1874	(0.0856)	0.1940	(0.088)
health	-0.1512	(0.1727)	-0.1462	(0.1739)	-0.1611	(0.178)
Annuity						
age65	0.0373	(0.2544)	0.0223	(0.2317)	0.0467	(0.246)
age70	-0.1736	(0.2640)	-0.1568	(0.2406)	0.1348	(0.255)
age75	-0.7565	(0.2874)	-0.584	(0.2632)	-0.2740	(0.275)
age80	-1.8209	(0.3164)	-1.5793	(0.3007)	-1.3150	(0.319)
fem	-0.2405	(0.1308)	-0.2362	(0.1215)	-0.2281	(0.129)
mort	-0.4118	(0.2608)	-0.4354	(0.2509)	-0.4471	(0.265)

Table A.10: Extend LC Model's Estimated Parameters of Equation System (5.11)

	Sup. Health Ins.		Long-Term Ins.		Life Ins.	
2LC						
Long-Term Ins.	0.5284	(0.1399)				
Life Ins.	0.1904	(0.0936)	0.3202	(0.1318)		
Annuity	-0.3966	(0.2206)	-0.0977	(0.2189)	0.3784	(0.1567)
3LC						
Long-Term Ins.	0.5661	(0.1423)				
Life Ins.	0.1735	(0.0962)	0.3840	(0.1354)		
Annuity	-0.2008	(0.1815)	-0.0601	(0.2141)	0.4594	(0.1515)
4LC						
Long-Term Ins.	0.6192	(0.1433)				
Life Ins.	0.1616	(0.1017)	0.4306	(0.1436)		
Annuity	-0.1669	(0.1947)	-0.0785	(0.2404)	0.7033	(0.1834)

Standard errors are reported in brackets.

Chapter 6

Conclusion

This thesis presents four studies that make use of latent class analysis to model unobserved heterogeneity in different empirical contexts applied to health and health care. The latent classes analysis is a particular way to model unobserved heterogeneity and it is exploited when both manifest dependent variables and latent variable are dichotomous and/or categorical. Therefore, latent class analysis requires computationally that the dependent variables are discrete. [1] Although this approach is extremely suitable to model multiple sources of unobserved heterogeneity for some particular economic context where the dependent is binary (e.g. having or not supplemental insurance, having any doctor visits or hospital admissions, etc.), there could be a loss of information when additional (continuous) indicators are discretized in order to be include in the model to identify unobserved heterogeneity. However as long as discretizations of continuous variables are reasonable results should no be affected.

In the previous chapters unobserved private information is examined in relation to 1) health production and self-reporting health behaviour, 2) the role of asymmetric information in health insurance markets and health care utilization, and 3) the generality of risk preferences in multiple demand for insurance. The work presented in the thesis provides a useful basis to develop new methodology and empirically evaluate policies in all of these contexts. The empirical appraisal of the effect of individual characteristics on health production and self-reporting behaviour introduced in Chapter 2 could be extended by modelling additional relationships between biomarkers in order to capture the effect of pre-existing health conditions on subjective health status. In addition, results provide evidence of the importance of using biomarkers in measuring health status. However,

[1] Note that in many cases binary variables are preferred to categorical variables in order to reduce the amount of time the model needs to converge throughout the EM algorithm (see. Forcina [59]).

more studies are required to determine the extent to which biomarkers are related to dimensions of health which differ from physical health.

The analysis of asymmetric information proposed in Chapters 3 and 4 suggests that multidimensionality of private information is an important issue when evaluating adverse selection in insurance markets. Extending the framework used in this thesis it would be interesting to understand whether there are other sources of multiple private information in addition to risk preferences and actual risk and their role in the selection effect. Moreover, a further stream of research might focus on understanding, both theoretically and empirically, the welfare implications of multidimensionality. For example, in the Medigap market the welfare effect of heavy cross subsidization of some types (who are high risk and risk averse) at the expense of others (who are low risk, but with low preferences for risk) has not been studied in detail. Finally, the analysis of the generality of risk preferences in Chapter 5 shows not only the existence of a common general component of risk preferences, but also the existence of residual specific heterogeneity related to the insurance choice context that are similar (e.g Medigap and long-term care insurance). Further research might extend several issues. From an econometric point of view it would be interesting to extend the estimation procedure to two or more latent variables in order to capture separately residual heterogeneity in risk and risk aversion. In addition it would be interesting to understand the drivers of this heterogeneity, what types of individuals domain-specificity are more relevant (e.g is context more important in the health insurance market than in the life insurance market?) and why should context still remain important after conditioning on individual risk and risk preferences.

Bibliography

[1] Abbring, Jaap H., Pierre-Andrè Chiappori, and Jean Pinquet (2003a) 'Moral hazard and dynamic insurance data.' *Journal of the European Economic Association* 1(4), 767–820

[2] Abbring, Jaap H., Pierre-Andrè Chiappori, James J. Heckman, and Jean Pinquet (2003b) 'Adverse selection and moral hazard in insurance: Can dynamic data help to distinguish?' *Journal of the European Economic Association* 1(2/3), 512–521

[3] Agnew, Julie R., Lisa R. Anderson, Jeffrey R. Gerlach, and Lisa R. Szykman (2008) 'Who chooses annuities? an experimental investigation of the role of gender, framing, and defaults.' *American Economic Review* 98(2), 418–22

[4] Agresti, A. (2002) *Categorical data analysis* (Wiley-Interscience)

[5] Arnott, Richard J., and Joseph E. Stiglitz (1990) 'The basic analytics of moral hazard.' NBER Working Papers 2484, National Bureau of Economic Research, Inc, February

[6] Arrow, Kenneth J. (1963) 'Uncertainty and the welfare economics of medical care.' *The American Economic Review* 53(5), 941–973

[7] Banks, James, Michael Marmot, Zoe Oldfield, and James P. Smith (2006) 'Disease and Disadvantage in the United States and in England.' *JAMA* 295(17), 2037–2045

[8] Banthin, Jessica S., and Jessica P. Vistnes (1997) 'The demand for Medicare supplemental insurance benefits: the role of attitudes toward medical care and risk.' *Inquiry: a journal of medical care organization, provision and financing* 34(4), 311

[9] Barseghyany, Levon, Jeffrey Princez, and Joshua C. Teitelbaum (forthcoming) 'Are risk preferences stable across contexts? evidence from insurance data.' *American Economic Review*

[10] Barsky, Robert B., F. Thomas Juster, Miles S. Kimball, and Matthew D. Shapiro (1997) 'Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study*.' *Quarterly Journal of Economics* 112(2), 537–579

[11] Bartholomew, David J., and Martin Knott (1999) *Latent variable models and factor analysis* (A Hodder Arnold Publication)

[12] Bartolucci, Francesco, and Antonio Forcina (2005) 'Likelihood inference on the underlying structure of irt models.' *Psychometrika* 70(1), 31–43

[13] Bartolucci, Francesco, and Antonio Forcina (2006) 'A class of latent marginal models for capture-recapture data with continuous covariates.' *Journal of the American Statistical Association* 101, 786–794

[14] Bergsma, Wicher, Marcel Croon, and Jacques A. Hagenaars (2009) *Marginal Models: For Dependent, Clustered, and Longitudinal Categorial Data* (Springer Verlag)

[15] Blavatskyy, Pavlo, and Ganna Pogrebna (2008) 'Risk aversion when gains are likely and unlikely: Evidence from a natural experiment with large stakes.' *Theory and Decision* 64(2), 395–420

[16] Browne, Mark J., and Kihong Kim (1993) 'An international analysis of life insurance demand.' *The Journal of Risk and Insurance* 60(4), 616–634

[17] Buchmueller, Thomas C., Agnès Couffinhal, Michel Grignon, and Marc Perronnin (2004) 'Access to physician services: does supplemental insurance matter? evidence from france.' *Health Economics* 13(7), 669–687

[18] Cameron, A. Colin, and Pravin K. Trivedi (1991) 'The role of income and health risk in the choice of health insurance : Evidence from australia.' *Journal of Public Economics* 45(1), 1 – 28

[19] Cameron, A. Colin, Pravin K. Trivedi, Frank Milne, and John Piggott (1988) 'A microeconometric model of the demand for health care and health insurance in australia.' *The Review of Economic Studies* 55(1), 85–106

[20] Cameron, Stephen V., and James J. Heckman (1998) 'Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of american males.' *The Journal of Political Economy* 106(2), 262–333

[21] Cardon, James H., and Igal Hendel (2001) 'Asymmetric information in health insurance: Evidence from the national medical expenditure survey.' *The RAND Journal of Economics* 32(3), 408–427

[22] Cartwright, William S., Teh-Wei Hu, and Lien-Fu Huang (1992) 'Impact of varying medigap insurance coverage on the use of medical services of the elderly.' *Applied Economics* 24(5), 529–39

[23] Cawley, John, and Tomas Philipson (1999) 'An empirical examination of information barriers to trade in insurance.' *The American Economic Review* 89(4), 827–846

[24] Chiappori, Pierre-André, Bruno Jullien, Bernard Salanié, and Francois Salanié (2006) 'Asymmetric information in insurance: General testable implications.' *The RAND Journal of Economics* 37(4), 783–798

[25] Chiappori, Pierre-André, and Bernard Salanié (2000) 'Testing for asymmetric information in insurance markets.' *The Journal of Political Economy* 108(1), 56–78

[26] Chiappori, Pierre-Andre, Franck Durand, and Pierre-Yves Geoffard (1998) 'Moral hazard and the demand for physician services: First lessons from a french natural experiment.' *European Economic Review* 42(3-5), 499–511

[27] Cicchetti, Charles J., and Jeffrey A. Dubin (1994) 'A microeconometric analysis of risk aversion and the decision to self-insure.' *Journal of Political Economy* 102(1), 169–86

[28] Cohen, Alma (2005) 'Asymmetric information and learning: Evidence from the automobile insurance market.' *Review of Economics and Statistics* 87(2), 197–207

[29] Cohen, Alma, and Liran Einav (2007) 'Estimating risk preferences from deductible choice.' *American Economic Review* 97(3), 745–788

[30] Cohen, Alma, and Peter Spiegelman (2010) 'Testing for adverse selection in insurance markets.' *Journal of Risk and Insurance* 77(1), 39–84

[31] Crossley, Thomas F., and Steven Kennedy (2002) 'The reliability of self-assessed health status.' *Journal of Health Economics* 21(4), 643 – 658

[32] Cutler, David M., Amy Finkelstein, and Kathleen McGarry (2008) 'Preference heterogeneity and insurance markets: Explaining a puzzle of insurance.' *American Economic Review* 98(2), 157–162

[33] Cutler, David M., and Edward Glaeser (2005) 'What explains differences in smoking, drinking, and other health-related behaviors?' *American Economic Review* 95(2), 238–242

[34] Dardanoni, Valentino, and Antonio Forcina (1998) 'A unified approach to likelihood inference on stochastic orderings in a nonparametric context.' *Journal of the American Statistical Association* 93(443), 1112–1123

[35] Dardanoni, Valentino, and Paolo Li Donni (2008) 'Testing for asymmetric information in insurance markets with unobservable types.' Health, Econometrics and Data Group (HEDG) Working Papers 08/26, HEDG, c/o Department of Economics, University of York, October

[36] Dardanoni, Valentino, Antonio Forcina, and Salvatore Modica (2008) 'Direct causal effects in education transmission.' mimeo, University of Palermo

[37] De Meza, David, and David C. Webb (2001) 'Advantageous selection in insurance markets.' *The RAND Journal of Economics* 32(2), 249–262

[38] Deb, Partha, and Pravin K. Trivedi (1997) 'Demand for medical care by the elderly: a finite mixture approach.' *Journal of Applied Econometrics* 12(3), 313–336

[39] Deb, Partha, and Pravin K. Trivedi (2002) 'The structure of demand for health care: latent class versus two-part models.' *Journal of Health Economics* 21(4), 601–625

[40] Dempster, Arthur P., N. M. Laird, and D. B. Rubin (1977) 'Maximum likelihood from incomplete data via the em algorithm.' *Journal of the Royal Statistical Society. Series B (Methodological)* 39(1), 1–38

[41] Dionne, Georges, Mathieu Maurice, Jean Pinquet, and Charles Vanasse (forthcoming) 'Incentive mechanisms for safe driving: A comparative analysis with dynamic data.' *The Review of Economics and Statistics*

[42] Doescher, Mark P., Peter Franks, Jessica S. Banthin, and Carolyn M. Clancy (2000) 'Supplemental Insurance and Mortality in Elderly Americans: Findings From a National Cohort.' *Arch Fam Med* 9(3), 251–257

[43] Dor, Avi, Joseph Sudano, and David W. Baker (2003) 'The effects of private insurance on measures of health: Evidence from the health and retirement study.' NBER Working Papers 9774, National Bureau of Economic Research, Inc, June

[44] d'Uva, Teresa Bago, Eddy Van Doorslaer, Maarten Lindeboom, and Owen O'Donnell (2008) 'Does reporting heterogeneity bias the measurement of health disparities?' *Health Economics* 17(3), 351–375

[45] Eichner, Matthew J. (1998) 'The demand for medical care: What people pay does matter.' *The American Economic Review* 88(2), 117–121

[46] Einav, L., and A. Finkelstein (2010) 'A Graphical Analysis of Selection in Insurance Markets.' *submitted*

[47] Einav, Liran, Amy Finkelstein, and Jonathan Levin (2009) 'Beyond testing: Empirical models of insurance markets.' Working Paper 15241, National Bureau of Economic Research

[48] Einav, Liran, Amy Finkelstein, and Paul Schrimpf (forthcoming) 'The welfare cost of asymmetric information: Evidence from the u.k. annuity market.' *Econometrica*

[49] Einav, Liran, Amy Finkelstein, Iuliana Pascu, and Mark R. Cullen (2010) 'How general are risk preferences? choices under uncertainty in different domains.' Working Paper 15686, National Bureau of Economic Research

[50] Etilé, Fabrice, and Carine Milcent (2006) 'Income-related reporting heterogeneity in self-assessed health: evidence from france.' *Health Economics* 15(9), 965–981

[51] Ettner, Susan L. (1997) 'Adverse selection and the purchase of Medigap insurance by the elderly.' *Journal of Health Economics* 16(5), 543–562

[52] Evans, William N., and W. Kip Viscusi (1991) 'Estimation of state-dependent utility functions using survey data.' *The Review of Economics and Statistics* 73(1), 94–104

[53] Fang, Hanming, Michael P. Keane, and Dan Silverman (2008) 'Sources of advantageous selection: Evidence from the medigap insurance market.' *Journal of Political Economy* 116(2), 303–350

[54] Finkelstein, Amy (2004) 'Minimum standards, insurance regulation and adverse selection: evidence from the Medigap market.' *Journal of Public Economics* 88(12), 2515–2547

[55] Finkelstein, Amy, and James Poterba (2002) 'Selection effects in the united kingdom individual annuities market.' *The Economic Journal* 112(476), 28–50

[56] Finkelstein, Amy, and James Poterba (2004) 'Adverse selection in insurance markets: Policyholder evidence from the u.k. annuity market.' *Journal of Political Economy* 112(1), 183–208

[57] Finkelstein, Amy, and James Poterba (2006) 'Testing for adverse selection with "unused observables".' Working Paper 12112, National Bureau of Economic Research

[58] Finkelstein, Amy, and Kathleen McGarry (2006) 'Multiple dimensions of private information: Evidence from the long-term care insurance market.' *The American Economic Review* 96(4), 938–958

[59] Forcina, Antonio (2008) 'Identifiability of extended latent class models with individual covariates.' *Computational Statistics and Data Analysis* 52(12), 5263–5268

[60] Friedman, Benjamin M., and Mark J. Warshawsky (1990) 'The cost of annuities: Implications for saving behavior and bequests.' *The Quarterly Journal of Economics* 105(1), 135–154

[61] Gardiol, Lucien, Pierre-Yves Geoffard, and Chantal Grandchamp (2005) 'Separating selection and incentive effects in health insurance.' Technical Report

[62] Gerdtham, U. G., M. Johannesson, L. Lundberg, and D. Isacson (1999) 'The demand for health: results from new measures of health capital.' *European Journal of Political Economy* 15(3), 501 – 521

[63] Goodman, Leo A. (1974) 'Exploratory latent structure analysis using both identifiable and unidentifiable models.' *Biometrika* 61(2), 215–231

[64] Gourieroux, Christian, and Alain Monfort (1995) *Statistics and econometric models* (Cambridge University Press)

[65] Groot, Wim (2000) 'Adaptation and scale of reference bias in self-assessments of quality of life.' *Journal of Health Economics* 19(3), 403 – 420

[66] Guiso, Luigi, and Monica Paiella (2005) 'The role of risk aversion in predicting individual behavior.' Temi di discussione (Economic working papers) 546, Bank of Italy, Economic Research Department, February

[67] Guiso, Luigi, and Monica Paiella (2008) 'Risk aversion, wealth, and background risk.' *Journal of the European Economic Association* 6(6), 1109–1150

[68] Heckman, James, and Burton Singer (1984) 'A method for minimizing the impact of distributional assumptions in econometric models for duration data.' *Econometrica* 52(2), 271–320

[69] Heckman, James J. (2001) 'Micro data, heterogeneity, and the evaluation of public policy: Nobel lecture.' *The Journal of Political Economy* 109(4), 673–748

[70] Hemenway, David (1990) 'Propitious selection.' *The Quarterly Journal of Economics* 105(4), 1063–1069

[71] Hernández-Quevedo, Cristina, Andrew Michael Jones, and Nigel Rice (2005) 'Reporting bias and heterogeneity in selfassessed health. evidence from the british household panel survey.' Health, Econometrics and Data Group (HEDG) Working Papers, HEDG, c/o Department of Economics, University of York

[72] Hershey, John C., Howard C. Kunreuther, and Paul J. H. Schoemaker (1982) 'Sources of bias in assessment procedures for utility functions.' *Management Science* 28(8), 936–954

[73] Holly, Alberto, Lucien Gardiol, Gianfranco Domenighetti, and Brigitte Bisig (1998) 'An econometric model of health care utilization and health insurance in switzerland.' *European Economic Review* 42(3-5), 513 – 522

[74] Holt, Charles A., and Susan K. Laury (2002) 'Risk aversion and incentive effects.' *American Economic Review* 92(5), 1644–1655

[75] Huang, Guan-Hua, and Karen Bandeen-Roche (2004) 'Building an identifiable latent class model with covariate effects on underlying and measured variables.' *Psychometrika* 69(1), 5–32

[76] Hurd, Michael D., and Kathleen McGarry (1997) 'Medical insurance and the use of health care services by the elderly.' *Journal of Health Economics* 16(2), 129–154

[77] Idler, Ellen L., and Yael Benyamini (1997) 'Self-rated health and mortality: A review of twenty-seven community studies.' *Journal of Health and Social Behavior* 38(1), 21–37

[78] Johnson, Eric J., John Hershey, Jacqueline Meszaros, and Howard Kunreuther (1993) 'Framing, probability distortions, and insurance decisions.' *Journal of Risk and Uncertainty* 7(1), 35–51

[79] Johnston, David W., Carol Propper, and Michael A. Shields (2009) 'Comparing subjective and objective measures of health: Evidence from hypertension for the income/health gradient.' *Journal of Health Economics* 28(3), 540 – 552

[80] Jones, Andrew Michael, Xander Koolman, and Eddy Van Doorslaer (2007) 'The impact of supplementary private health insurance on the use of specialists in European countries.' In 'Annales d'Economie et de Statistiques,' vol. 83

[81] Joreskog, Karl G., and Arthur S. Goldberger (1975) 'Estimation of a model with multiple indicators and multiple causes of a single latent variable.' *Journal of the American Statistical Association* 70(351), 631–639

[82] Jullien, Bruno, and Bernard Salanié (2000) 'Estimating preferences under risk: The case of racetrack bettors.' *Journal of Political Economy* 108(3), 503–530

[83] Kahneman, Daniel (2003a) 'Maps of bounded rationality: Psychology for behavioral economics.' *American economic review* 93(5), 1449–1475

[84] Kahneman, Daniel (2003b) 'A psychological perspective on economics.' *American Economic Review* 93(2), 162–168

[85] Kerkhofs, Marcel, and Maarten Lindeboom (1995) 'Subjective health measures and state dependent reporting errors.' *Health Economics* 4(3), 221–235

[86] Khandker, Rezaul K., and Lauren A. McCormack (1999) 'Medicare spending by beneficiaries with various types of supplemental insurance.' *Medical Care Research and Review* 56(2), 137

[87] Kimball, Miles S., Claudia R. Sahm, and Matthew D. Shapiro (2008) 'Imputing risk tolerance from survey responses.' *Journal of the American Statistical Association* 103(483), 1028–1038

[88] Kodde, David A., and Franz C. Palm (1986) 'Wald criteria for jointly testing equality and inequality restriction s.' *Econometrica* 54(5), 1243–48

[89] Landerman, Lawrence R., Gerda G. Fillenbaum, Carl F. Pieper, George L. Maddox, Deborah T. Gold, and Jack M. Guralnik (1998) 'Private health insurance coverage and disability among older Americans.' *The Journals of Gerontology: Series B* 53(5), S258

[90] Lindeboom, Maarten, and Eddy van Doorslaer (2004) 'Cut-point shift and index shift in self-reported health.' *Journal of Health Economics* 23(6), 1083 – 1099

[91] Manning, Willard G., Joseph P. Newhouse, Naihua Duan, Emmett B. Keeler, and Arleen Leibowitz (1987) 'Health insurance and the demand for medical care: Evidence from a randomized experiment.' *The American Economic Review* 77(3), 251–277

[92] McCarthy, David, and Olivia S. Mitchell (2003) 'International adverse selection in life insurance and annuities.' Technical Report

[93] McLachlan, Geoffrey J., and David Peel (2000) *Finite mixture models* (Wiley-Interscience)

[94] Mehta, Kala M., Kristine Yaffe, Kenneth M. Langa, Laura Sands, Mary A. Whooley, and Kenneth E. Covinsky (2003) 'Additive Effects of Cognitive Function and Depressive Symptoms on Mortality in Elderly Community-Living Adults.' *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 58(5), M461–M467

[95] Monfardini, Chiara, and Rosalba Radice (2008) 'Testing exogeneity in the bivariate probit model: A monte carlo study.' *Oxford Bulletin of Economics and Statistics* 70(2), 271–282

[96] Murray, CJL, A. Tandon, J. Salomon, and CD Mathers (2000) 'Enhancing cross-population comparability of survey results.' *Health systems performance: concepts, measurement and determinants. Geneva: World Health Organization*

[97] Pauly, Mark V. (1968) 'The economics of moral hazard: Comment.' *The American Economic Review* 58(3), 531–537

[98] Pearl, Judea (2000) *Causality: Models, reasoning, and inference* (Cambridge Univ Pr)

[99] Polborn, Mattias K., Michael Hoy, and Asha Sadanand (2006) 'Advantageous effects of regulatory adverse selection in the life insurance market.' *Economic Journal* 116(508), 327–354

[100] Propper, Carol (1989) 'An econometric analysis of the demand for private health insurance in England and Wales.' *Applied Economics* 21(6), 777–792

[101] Propper, Carol (1993) 'Constrained choice sets in the u.k. demand for private medical insurance.' *Journal of Public Economics* 51(3), 287 – 307

[102] Rabin, Matthew (2000) 'Risk aversion and expected-utility theory: A calibration theorem.' *Econometrica* 68(5), 1281–1292

[103] Rabin, Matthew, and Richard H. Thaler (2001) 'Anomalies: Risk aversion.' *The Journal of Economic Perspectives* 15(1), 219–232

[104] Rothschild, Michael, and Joseph Stiglitz (1976) 'Equilibrium in competitive insurance markets: An essay on the economics of imperfect information.' *The Quarterly Journal of Economics* 90(4), 629–649

[105] Sadana, R., C.D. Mathers, A.D. Lopez, C.J.L. Murray, and K. Iburg (2002) 'Comparative analysis of more than 50 household surveys of health status.' *Summary measures of population health: concepts, ethics, measurement and applications. Geneva, World Health Organization* pp. 369–386

[106] Schellhorn, Martin (2001) 'The effect of variable health insurance deductibles on the demand for physician visits.' *Health Economics* 10(5), 441–456

[107] Shmueli, Amir (2003) 'Socio-economic and demographic variation in health and in its measures: the issue of reporting heterogeneity.' *Social Science and Medicine* 57(1), 125 – 134

[108] Smart, Michael (2000) ‘Competitive insurance markets with two unobservables.’ *International Economic Review* 41(1), 153–69

[109] Sydnor, Justin (2006) ‘Abundant Aversion to Moderate Risk: Evidence from Homeowners Insurance’

[110] Terza, Joseph V. (1985) ‘Ordinal probit: a generalization.’ *Communications in statistics. Theory and methods* 14(1), 1–11

[111] Thomas, Dohmen, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G. Wagner (2009) ‘Individual risk attitudes: Measurement, determinants and behavioral consequences.’ Technical Report

[112] van Doorslaer, Eddy, and Andrew M. Jones (2003) ‘Inequalities in self-reported health: validation of a new approach to measurement.’ *Journal of Health Economics* 22(1), 61 – 87

[113] Vera-Hernandez, Ángel M. (2000) ‘Duplicate coverage and demand for health care. The case of Catalonia.’ *Health Economics* 8(7), 579–598

[114] Villeneuve, Bertrand (2003) ‘Concurrence et antisélection multidimensionnelle en assurance.’ *Annales d’Economie et de Statistique* (69), 06

[115] Viscusi, W. Kip (2004) ‘The value of life: Estimates with risks by occupation and industry.’ *Economic Inquiry* 42(1), 29–48

[116] Viscusi, W. Kip, and William N. Evans (1990) ‘Utility functions that depend on health status: Estimates and economic implications.’ *The American Economic Review* 80(3), 353–374

[117] Wambach, Achim (2000) ‘Introducing heterogeneity in the rothschild-stiglitz model.’ *The Journal of Risk and Insurance* 67(4), 579–591

[118] Wilde, Joachim (2000) ‘Identification of multiple equation probit models with endogenous dummy regressors.’ *Economics Letters* 69(3), 309 – 312

[119] Wingard, Deborah L. (1984) ‘The sex differential in morbidity, mortality, and lifestyle.’ *Annual Review of Public Health* 5(1), 433–458

[120] Wolf, Charles, and Larry Pohlman (1983) ‘The recovery of risk preferences from actual choices.’ *Econometrica* 51(3), 843–850

[121] Wooldridge, James M. (2002) *Econometric analysis of cross section and panel data* (The MIT press)

[122] Zweifel, Peter, and Willard G. Manning (2000) 'Moral hazard and consumer incentives in health care.' In *Handbook of Health Economics,* ed. A. J. Culyer and J. P. Newhouse, vol. 1 of *Handbook of Health Economics* (Elsevier) chapter 8, pp. 409–459