HEDG Health, Econometrics and Data Group

THE UNIVERSITY of York

HEDG Working Paper 10/17

RISK PREFERENCE HETEROGENEITY AND MULTIPLE DEMAND FOR INSURANCE

PAOLO LI DONNI

July 2010

york.ac.uk/res/herc/hedgwp

RISK PREFERENCE HETEROGENEITY AND MULTIPLE DEMAND FOR INSURANCE

PAOLO LI DONNI

Abstract. We examined the relationship between unobserved risk preferences and four insurance purchase decisions: health Medigap insurance, long-term insurance, life insurance and annuity. Standard economic theory assumes that individuals take decision 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. Using data from the Health and Retirement Study, we exploit latent class analysis to identify conditional on predicted and realized risk how heterogeneity in risk preferences affects multiple insurance demand. Our results provide evidence of the existence of domain general component of risk preferences, although non-preference factors - such as context specificity - play also an important role.

JEL Classification Numbers G11, D82, G22, I11

Keywords Risk Preferences, Multiple Demand for Insurance, Finite Mixture Model, Long-Term Care Insurance, Medigap, Annuity, Life Insurance.

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.* [12], Einav *et al.* [17], Barseghyany *et al.* [2]). 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 bearing 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 between a 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 [39] and Rabin and Thaler [40]). They argue that individuals'

Date: July 9, 2010.

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 [10]). 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 whather risk preferences are general. Cohen and Einav [10] and Barseghyan *et al.* [2] 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.* [17] 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.* [17]) 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), log-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.* [17], 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.

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 [24] 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 [21]-[22] and McCarthy and Mitchell [37] 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 [8]). Browne and Kim [7] 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.

Log-term insurance combines elements of both annuity and life insurance. Finkelstein and McGarry [23] 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.* [12] 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 [6]. In general these studies use individual observed choice obtained from survey data - sometimes with experimental module (e.g., Viscusi and Evans [44]; Evans and Viscusi [19]; Barksy *et al.* [3]; Dohmen *et al.* [42]) or laboratory or natural experiment experiment (Holt and Laury [29], Jullien and Salanié [32], Guiso and Paiella [27]) to estimate risk preference.

Barsky *et al.* [3] 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.* [42] also find statistically significant evidence of relationship between financial and non-financial risk aversion on the basis of survey data. Guiso and Paiella [26] 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 [13] 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 [9], Sydnor [41]). In a recent paper Cohen and Einav [10] 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 [33]-[34]). 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.* [3], Dohmen *et al.* [42], Kimball *et al.* [35]), while several other studies based on laboratory experiments and hypothetical money gables showed that context is the most important factor (Wolf and Pohlman [45]) or even that choice depends on whether questions are framed as a "gamble" or as "insurance" (Hershey *et al.* [28], Johnson *et al.* [31]).

Recently Barseghyan *et al.* [2] take an innovative approach to test generality of individual risk preference. Following Cohen and Einav [10] 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.* [17] 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.* [2] and Einav *et al.* [17], 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.).

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.* [12], Finklstein and McGarry [23]) 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.* [3], Kimball *et al.* [35]). 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 1, while in the following subsections we describe the variables used to measure insurance coverage, individual risk and risk preferences.

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.* [20]). 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.

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.* [12], Cohen and Einav [10], Cohen and Spiegelman [11]).

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 occurrence and insurance purchase decision is often mentioned as source of adverse selection (Cohen and Spiegelman [11] and Einav *et al.* [16]). 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 [10]).

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

[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.* [38]) showed that this measure is associated with the existence of psychiatric problems.

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.

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.* [3] and Cutler and Glaeser [13] 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.* [12] we derive the mortality rates from Viscusi [43]. 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.* [3] 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 [26]-[27] 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 leave 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 [26]-[27]). 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.* [12] and Finkelstein and McGarry [23]) 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.

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 [30], Bartolucci and Forcina [5] and Dardanoni, Forcina and Modica [15]).

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

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

where $\boldsymbol{w}_1, \dots, \boldsymbol{w}_J$ 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 (1) can be written as:

$$\begin{gathered} Pr(I_1 = 1 \mid \boldsymbol{w}_1, P_1) \\ \vdots \\ Pr(I_J = 1 \mid \boldsymbol{w}_J, P_J) \end{gathered} \tag{2}$$

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.* [17] 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 $P_1, \dots, P_J$ are *discrete*, with P_j taking say m_k levels, $k = 1, \dots, K$. 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 $P_1, \dots, P_K$ 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.

How can we test then the DGC hypothesis? Suppose that for some arrangement of the M types U we have

$$\begin{aligned} Pr(I_1 = 1 \mid \boldsymbol{w}_1, U = 1) \leq \cdots \leq Pr(I_1 = 1 \mid \boldsymbol{w}_1, U = M) \\ \vdots \\ Pr(I_J = 1 \mid \boldsymbol{w}_J, U = 1) \leq \cdots \leq Pr(I_J = 1 \mid \boldsymbol{w}_J, U = M) \end{aligned} \tag{3}$$

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.[2] Let to analyse how this procedure can be implemented empirically.

4.1. **Empirical strategy.** Following standard models in the literature on insurance demand (Cohen and Einav [10], Cutler *et al.* [12], Einav *et al.* [17]), $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_J$ 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 (1) as:

$$\begin{aligned} Pr(I_1 = 1 \mid \boldsymbol{w}_1, P) = F(\boldsymbol{x}'_1\boldsymbol{\beta}_1 + \boldsymbol{r}'_1\gamma_1 + \boldsymbol{v}'_1\boldsymbol{\delta}_1) \\ \vdots \\ Pr(I_J = 1 \mid \boldsymbol{w}_1, P) = F(\boldsymbol{x}'_J\boldsymbol{\beta}_J + \boldsymbol{r}'_J\boldsymbol{\gamma}_J + \boldsymbol{v}'_J\boldsymbol{\delta}_J) \end{aligned} \tag{4}$$

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 (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.* [17] and extended LCA model.

4.2. **Multivariate probit regression.** In a recent paper Einav *et al.* [17] study the DGC hypothesis examining the correlation structure of the error terms in

[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.* [12], Cohen and Einav [10] and Einav *et al.* [17]) 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.

a multivariate regression. Following this approach, let the link function F be standard normal, so that we can equivalently rewrite the system (4) as:

$$\begin{aligned} I_1 &= 1\big(\boldsymbol{x}'_1\boldsymbol{\beta}_1 + \boldsymbol{r}'_1\boldsymbol{\gamma}_1 + \boldsymbol{v}'_1\boldsymbol{\delta}_1 + \epsilon_1\big) \\ &\vdots \\ I_J &= 1\big(\boldsymbol{x}'_J\boldsymbol{\beta}_J + \boldsymbol{r}'_J\boldsymbol{\gamma}_J + \boldsymbol{v}'_J\boldsymbol{\delta}_J + \epsilon_J\big) \end{aligned} \tag{5}$$

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:

$$\begin{aligned} I_1 &= 1\big(\boldsymbol{x}'_1\boldsymbol{\beta}_1 + \boldsymbol{r}'_1\boldsymbol{\gamma}_1 + \eta_1\big) \\ &\vdots \\ I_J &= 1\big(\boldsymbol{x}'_J\boldsymbol{\beta}_J + \boldsymbol{r}'_J\boldsymbol{\gamma}_J + \eta_J\big) \end{aligned} \tag{6}$$

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.

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 (4), which account for the unobserved U can be written as:

$$\begin{aligned} I_1 &= \textstyle\sum_{u=1}^m \alpha_u^{I_1} U_u + \boldsymbol{x}'_1\boldsymbol{\beta}_1 + \boldsymbol{r}'_1\boldsymbol{\gamma}_1 + \eta_1 \\ &\vdots \\ I_J &= \textstyle\sum_{u=1}^m \alpha_u^{I_J} U_u + \boldsymbol{x}'_J\boldsymbol{\beta}_J + \boldsymbol{r}'_J\boldsymbol{\gamma}_J + \eta_J \end{aligned} \tag{7}$$

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 (7), we estimate the model:

$$\begin{aligned} \lambda^{I_1} &= \textstyle\sum_{u=1}^m \alpha_u^{I_1} U_u + \boldsymbol{x}'_1\boldsymbol{\beta}_1 + \boldsymbol{r}'_1\boldsymbol{\gamma}_1 \\ &\vdots \\ \lambda^{I_J} &= \textstyle\sum_{u=1}^m \alpha_u^{I_J} U_u + \boldsymbol{x}'_J\boldsymbol{\beta}_J + \boldsymbol{r}'_J\boldsymbol{\gamma}_J \end{aligned} \tag{8}$$

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

$$\log\left(\frac{Pr(U=u+1)}{Pr(U=u)}\right) = \lambda_u^U = \alpha_u^U \qquad u = 1, \ldots, m-1 \tag{9}$$

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

$$\begin{aligned} \lambda^{H_1} &= \sum_{u=1}^{m} \alpha_u^{H_1} U_u \\ \lambda^{H_2} &= \sum_{u=1}^{m} \alpha_u^{H_2} U_u \\ \lambda^{H_3} &= \sum_{u=1}^{m} \alpha_u^{H_3} U_u \\ \lambda^{H_4} &= \sum_{u=1}^{m} \alpha_u^{H_4} U_u \end{aligned} \tag{10}$$

Note that the the system of equations (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 (8).

In addition to equations (8-10) 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:

$$\lambda_{I_j, I_k} = \alpha_{I_j, I_k} \tag{11}$$

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 (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.* [17]).

Within the model defined by equations ((8)-(11)),

- the null hypothesis of DGC of individual risk preferences (that is equation (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 [4]. 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 [25] for a general exposition, Dardanoni and Forcina [14] for an explanation of how the mixing weights can be calculated by simulations, and Kodde and Palm [36] for bounds on the test distribution).
- the null hypothesis of absence of residual heterogeneity related to potential non-preference factors 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. 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, log-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.

5.1. **Multivariate Regression.** Tables 2 and 3 present respectively the estimated coefficients of controls and correlation terms from the baseline multivariate probit regression suggested by Einav *et al.* [17] and described above in equation (6). Let consider first the determinants of supplemental health insurance purchase decision. Table 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.* [20], Ettner [18]).

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 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.* [1] 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 [11]).

Consider now the estimated correlations between insurance purchase decisions. In all of the pairs reported in table 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.* [17], 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.

5.2. **Results from the Extended LCA Model.** We start by estimating the system of equations (8)-(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 [15] and at Bartolucci and Forcina [5].[3] For completeness we report in tables 4-10 model's estimated parameters under different number of latent classes, namely $m = 2, 3, 4$. Table 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 5, about 50% of individuals are of type 1, while 30% and 20% are of type 2 and 3 respectively.

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

To understand what these types indicate, let consider the estimated probabilities reported in table 6, obtained using the α parameters of tables 7 and 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 log-term insurance and annuity than type 3 individuals. Therefore a first glance at Panel A of table 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 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 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.* [3] Guiso and Paiella [27]). The pattern we find is consistent with other studies, such as Barsky *et al.* [3], 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 [10], Barseghyan *et al.* [2]and Einav *et al.* [17]), 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 (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 insurnace choices we consider are four, the conservative 1% critical value with 8 df is equal to 25.370 (Kodde and Palm ([36], 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 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 log-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.* [17] and support the idea that choice is driven both by context and by how individuals are risk averse in general. However it is possible that estimated residual correlations may also depend on unpriced risk not by risk occurrence proxies.

Finally let us to consider the effect of predicted (by insurer) and realized ex-post risk in each insurance equation. Table 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 [11] and Einav *et al.* [16]).

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 [10], Barseghyan *et al.* [2] and Einav *et al.* [17] which focus on how general are risk individual preferences.

In particular we start following an innovative approach proposed by Einav *et al.* [17] 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, log-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 log-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 no 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 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
Log-Term Ins.	1 = enrolled in any log-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 2. Multivariate Probit Model's Estimated Parameters of predicted and realized risk

Variables	Sup. Health Ins.		Log-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 3. Multivariate Probit Model's Estimated Correlation Terms Controlling for Predicted and Realized Risk

Variables	Sup. Health Ins.		Log-Term Ins.		Life Ins.	
Log-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 4. Model Selection Criteria for System of Equations (8)-(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 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 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
Log-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: Auxiliary 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 7. Estimated Intercepts α of Equation System (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)
Log-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 8. Estimated Intercepts α of Equation System (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 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)
Log-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 10. Extend LC Model's Estimated Parameters of Equation System (11)

	Sup. Health Ins.		Log-Term Ins.		Life Ins.	
2LC						
Log-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						
Log-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						
Log-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.

REFERENCES

[1] 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

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

[3] 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

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

[5] ——— (2006) 'A class of latent marginal models for capture-recapture data with continuous covariates.' *Journal of the American Statistical Association* 101, 786–794

[6] 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

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

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

[9] 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

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

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

[12] 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

[13] 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

[14] 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

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

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

[17] 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

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

[19] 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

[20] 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

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

[22] ——— (2004) 'Adverse selection in insurance markets: Policyholder evidence from the u.k. annuity market.' *Journal of Political Economy* 112(1), 183–208

[23] 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

[24] 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

[25] Gourieroux, Christian, and Alain Monfort (1995) *Statistics and econometric models* (Cambridge University Press)
[26] 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
[27] ——— (2008) 'Risk aversion, wealth, and background risk.' *Journal of the European Economic Association* 6(6), 1109–1150
[28] 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
[29] Holt, Charles A., and Susan K. Laury (2002) 'Risk aversion and incentive effects.' *American Economic Review* 92(5), 1644–1655
[30] 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
[31] 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
[32] Jullien, Bruno, and Bernard Salanié (2000) 'Estimating preferences under risk: The case of racetrack bettors.' *Journal of Political Economy* 108(3), 503–530
[33] Kahneman, Daniel (2003a) 'Maps of bounded rationality: Psychology for behavioral economics.' *American economic review* 93(5), 1449–1475
[34] ——— (2003b) 'A psychological perspective on economics.' *American Economic Review* 93(2), 162–168
[35] 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
[36] Kodde, David A., and Franz C. Palm (1986) 'Wald criteria for jointly testing equality and inequality restriction s.' *Econometrica* 54(5), 1243–48
[37] McCarthy, David, and Olivia S. Mitchell (2003) 'International adverse selection in life insurance and annuities.' Technical Report
[38] 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
[39] Rabin, Matthew (2000) 'Risk aversion and expected-utility theory: A calibration theorem.' *Econometrica* 68(5), 1281–1292
[40] Rabin, Matthew, and Richard H. Thaler (2001) 'Anomalies: Risk aversion.' *The Journal of Economic Perspectives* 15(1), 219–232
[41] Sydnor, Justin (2006) 'Abundant Aversion to Moderate Risk: Evidence from Homeowners Insurance'

[42] 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

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

[44] 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

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

Università di Palermo and University of York
E-mail address: plidonni@unipa.it