Alma Mater Studiorum - Università di Bologna

Scuola di Dottorato in Scienze Economiche e Statistiche
Dottorato di ricerca in

Metodologia Statistica per la Ricerca Scientifica
XXV ciclo

Multilevel Analysis in Household Surveys: An Application to Health Condition Data

Sara Piombo

Dipartimento di Scienze Statistiche "Paolo Fortunati"
Febbraio 2013

Alma Mater Studiorum - Università di Bologna

Scuola di Dottorato in Scienze Economiche e Statistiche
Dottorato di ricerca in

Metodologia Statistica per la Ricerca Scientifica
XXIV ciclo

Multilevel Analysis in Household Surveys: An Application to Health Condition Data

Sara Piombo

Coordinatore:
Prof.ssa Angela Montanari

Tutor:
Prof.ssa Rossella Miglio

Co-Tutor:
Prof. Leonardo Grilli

Settore Disciplinare: SECS-S/05
Settore Concorsuale: 13/D3

Dipartimento di Scienze Statistiche "Paolo Fortunati"
Febbraio 2013

to Matteo, Antonio
and Didi

Contents

List of Tables

List of Figures

Introduction

In the last decades multilevel modeling became a very popular method for model estimate where units are nested within groups. Hierarchical structure involves dependence between observations collected within the same group, nevertheless most statistical methods traditionally assume independence among observations.
Multilevel Models, also called Hierarchical Models, handle these hierarchical data structures which are frequently encountered in many research areas.
These models (introduced by H. Goldstain in 1987) have historically been used in educational research where hierarchies occur naturally: students nested within classrooms, classrooms nested within schools and schools nested within districts.
Recent advances in statistical computing capabilities have made these models easily available to researchers in several disciplines and application areas: social science, epidemiology, longitudinal and survival analysis, educational achievement, psychometric, biostatistics and econometrics.
The aim of this study is to apply multilevel regression model in context of household surveys. Hierarchical structure in this type of data is characterized by many small groups (considering individuals at level 1 and households at level 2) which leads to some restrictions on type of multilevel models applicable. Istat survey on "Health Condition and Use of Health Services" is a good candidate for this task, given a very large sample size and the available territorial levels (geographic hierarchy levels). This survey has the objective to analyze the health behaviors and use of health services and place them in relation to the demographic and socio-economic citizens. Through this information, collected directly from households, it is possible to construct indicators of health status and quality of life, presence of disability, risk factors, diseases and prevention. In last years comparative and multilevel analysis in the field of perceived health have grown in size. In particular these studies are focalized on the health condition of the elderly, health territorial differences and the effect of deprived socio-economic context on health (Pirani and Salvini 2012, Olsen and Dahl 2007; Mackenbach et al. 2008).

The purpose of this thesis is to develop a multilevel analysis with tree level of hierarchy for Physical Component Summary outcome to:

- evaluate magnitude of within and between variance at each level (individual, household and municipality);
- explore which covariates affect on perceived physical health at each level;
- compare model-based and design-based approach in order to establish informativeness of sampling design;
- estimate a quantile regression for hierarchical data.

The target population are the Italian residents aged 18 years and older.

Chapter 1

The Istat Survey on Health Conditions

1.1 Survey description

The Italian survey on population health condition conducted by the Italian National Institute of Statistics (Istat) is a part of the ensemble of multi-purpose household surveys, including several thematic surveys repeated every five years. The first survey on health conditions was carried out in 1980, while the two most recent were made in 1999-2000 and in 2004-2005 respectively; the next one is still in progress. This work will examine latest survey available, the 2004-2005 edition. The topics involved in this survey are: perception of individual health status, presence of chronic diseases, disability and invalidity, lifestyle, risk factors such as smoking, body weight and physical activity, prevention, recourse of health services, consumption of drugs, and for women questions about pregnancy and lactation. To avoid seasonal effect, the survey was carried out during a whole year. Four-monthly interviews were made in December 2004, March 2005, June 2005 and September 2005. Subjects included in the survey are Italian population resident in households, excluding residents who live in convents, communities, nursing homes, prisons and barracks.
Sample was formed by 60,730 households distributed in 1,474 Italian municipalities with different population size. In this survey sample size was increased significantly compared to other household surveys (usually of 24,000 households) to allow estimates for small territorial areas: sub-regional areas. A particular sub-regional areas are the so-called "large areas" (AV) formed by aggregation of local health units.
The sampling design is two-stage (municipalities and households) with stratification of the first stage units (municipalities) by size.

Within each of the 68 Large Area, the municipalities were classified according to their population size and divided into two subsets: the municipalities with large population size constitutes a separate stratum and are defined self-representative (AR), remaining areas are defined not-self-representative (NAR) and are divided, by dimension, in stratum of equal size; within each NAR stratum a municipality sample (four within each stratum) is selected with proportional probability to size. Within each municipality (AR and NAR), households are selected systematically from the Registry offices.

Survey unit is the "de facto household" relating to households drawn from municipality population registers. "De facto household" is defined as a group of persons normally resident in the same house who are joined by kinship, affinity, affective or friendship relationships. For every "de facto household" all members were interviewed using two paper questionnaires: one filled in by the interviewer through a direct interview, one completed by respondents and containing sensitive questions.

1.2 The micro-data files

Istat provides to users and researchers a micro-data file containing the individual information collected (128,040 records) and makes it anonymous on about 650 variables. A second micro-data is available, the so called "*simplified*" file. This file contain a restricted number of variable but more territorial detail, therefore it is chosen as work-file. In order to facilitate analysis, synthetic indexes were constructed and inserted in this second file; description will be reported in following.

Main variables and indicators in "simplified" data-set .

Structural variables and local context

- *Demographic variables:* five and ten year age groups and some specific age, gender.
- *Geographic variables:* geographical area, region, Large area and municipality of residence respondent.
- *Status variables:* qualifications, years of study, socio-economic household condition, housing condition.

Health status indicators

- *On perceived health status:* Physical Component Summary, Mental Component Summary, Mental Health and Vitality (Well-being) indices.

- *Chronic diseases diagnosed:* Infarct, Heart disease, Stroke, Arthrosis and Arthrtis, Osteoporosis, Hepatic cirrhosis, Tumor, Ulcer, Tumor in the past, Anxiety and Depression, Alzheimer and senile dementia, kidney or liver or other type of Stones, Thyroid disease, Asthma, Diabetes, Hypertension, Bronchitis or Emphysema, Chronicity index.

Risk factors

- *Smoke:* smoker, former smoker, number of smoking years.

- Body-mass index, diets, physical activity, sport.

Prevention control and prophylaxis

- vaccinations, blood tests, blood pressure control, osteoporosis test, controls for the prevention of female cancers.

Indicators on health services use

- Medications intake, doctor examinations, diagnostic tests, hospitalization, day hospital, rehabilitation therapies, first aid, consumption of health services index, opinions on health services.

Variables on maternity

- Pregnancy, check-up during pregnancy, risk factors and illnesses during pregnancy, type of birth, breastfeeding.

Indexes description:

- *Housing condition index:* calculated on the basis of 6 items of inadequate housing conditions: 1) no bathroom, 2) absence of heating, 3) a too small house, 4) presence of humidity stains, 5) bad housing conditions, 6) less than a room for each household member.
  This index takes values from 1 to 7: 1=very bad condition, that is presence of all six negative situations; 2= presence of 5 out of 6 negative situations, and so on until to 7= none of the negative situation occurs.

- *Physical Component Summary (PCS) and Mental Component Summary (MCS).* To detect health status perception, this survey includes the SF-12 questionnaire (Short Form Health Survey) already used in many European empirical studies. Through twelve questions studying eight different aspects of health: physical activity, limitations due to physical health, emotional state, physical pain, perception of general health, vitality, social and work activity, mental health. The summary of the response scores on each question leads to two indices: one referring to the physical state (PCS) and the other on the psychological state (MCS). Low score indicates the poor health and high score indicates excellent health status. In this survey PCS range from 11,1 to 68,9 and MCS range from 7.5 to 72.3.

- *Vitality (VT) and Mental Health (MH).* To construct these indices the questionnaire SF-36 is used[1]. The VT index investigates the level of energy and tiredness through four questions, while for MH index five questions are used to study four main dimensions of mental health: anxiety, depression, loss of behavior or emotional control, and psychological well-being.

- *Chronicity index.* Made on the basis of self-perceived health status reported by each individual and the presence of various chronic diseases or invalidity. This index incorporates diseases weight on health of each individual by using odds ratios of feeling badly or very badly depending on the presence of each specific chronic disease. This index can take value from 0 to 100. Zero score indicates the lowest level of health, 100 indicates the highest level of health.

- *Disability and Invalidity.* A person is defined "disabled" if in presence of serious difficulty in at least one of these dimensions: confined to bed, a chair or at home; some difficulty moving or in daily functions, communicating (sight, hearing, speech). In detecting disability phenomenon, Istat always referred to questionnaire drawn up in the 80s by a work group from OCSE on the basis of the OMS classification (ICIDH- International Classification of Impairment, Disease, Disability and handicap-1980).
  Although limits of this instrument are known, up to now no ICF operationalization (International Classification of Functioning, Disability and Health), approved by OMS and shared internationally, is available yet. The two sections of population (disabled and invalid) are only partially overlapping, as population showing invalidity can be struck down even partially with correspondent disabilities and vice versa, but not all disabled people have got a disability recognition. Next to disability phenomenon, the survey points out some specific types of "invalidity" as well: mobility handicaps; mental insuffi-

[1]The methodology is reported in http://www.sf-36.org and http://crc.marionegri.it/qdv/

ciency; blindness; deaf-mutism and deafness, invalidity for mental disease. It concerns dimensions that are no perfectly overlapping to the correspondent types of disabilities, and that is for both the different conceptual approach underlying two phenomena and for the differences in adopted instruments. Detecting disability means to assess the level of autonomy reduction in carrying out the main functions, as a consequence of deficit or disablement due to the disease; taking into consideration a possible use of aids. "Invalidity", instead, is referred to disablement that affects an organ and therefore it is independent of the overall assessment of self-sufficiency.
The phenomenon is then detected, in contrast to what happens for disability, for children up to six years as well. Moreover, instruments adopted for the survey are not comparable; in fact disability derives from a synthesis of questions, invalidity is instead measured in terms of presence or absence of the specific disablement declared by respondents.

Because some variables has been detected only for 14 or 18 years old and over, the data-set under analysis will be restricted to sub-sample of individuals being 18 or more years old. This choice leads to a loss of about 22,000 individual records (from 128,040 to 105,844), the number of households remains unchanged, while the number of member per household will be reduced. The following table shows the sample size and the average size of the groups at different hierarchical levels. We note that household average size is small (2.1 members), but average size of the group at higher levels is satisfactory for the purposes of a multilevel analysis (e.g. about 34 households by municipality) and especially number of groups at each level is consistent (except for regional level).

Table 1.1: Number of units and average for each hierarchical level in the sub-sample of subjects being 18 years or over.

	N	average size by group
Regions	20	3.40
Large Areas	68	21.54
Municipalities	1465	34.45
households	50474	2.10
Individuals	105844	-

1.3 Outcome variable

In this survey, large sample size allow analyses on health characteristics at several sub-national levels, but given small level-2 size (household), is advisable to analyse continuous response variable. Continuous variables available in the data-set concern perceived health indices.
Over the past years, many studies have been conducted to describe and compare perceived health status mainly in elderly population (König et al, 2010) different geographical areas (K. Olsen, 2007) and for health inequality (K. Humphries and E. van Doorslaer,2000).
Our purpose is to explore perceived health status in overall Italian population, evaluate homogeneity magnitude of units belonging from the same cluster and select covariates having greatest impact on self-reported health. Differently from more common studies we perform a 3 level model, furthermore 2-nd level of analysis was household rather than a commonly used geographical area (E. Pirani e S. Salvini 2012, G. Costa at al. 2003).
Initially our interest focused on SF-12 questionnaire indexes: Physical Component Summary (PCS) and Mental Component Summary (MCS). Later due the intention to perform a particular multilevel model including quantile regression, we restrict analysis to PCS index only.
In this survey PCS range is from 11.12 to 68.70: width interval 57.58. It is characterized by a strong skewness as shown in table 1.2 and in figure 1.1.

Table 1.2: Physical Component Summary (PCS): several descriptive statistics in sub-sample of 18 aged and over.

Range:	11.12 - 68.70		
Mean	49.89	Median	54.32
Std. Dev.	9.67	Interq. Range	10.12
Variance	93.48	Skewness	-1.40
Mode	$\approx$ 56.00	Kurtosis	4.12

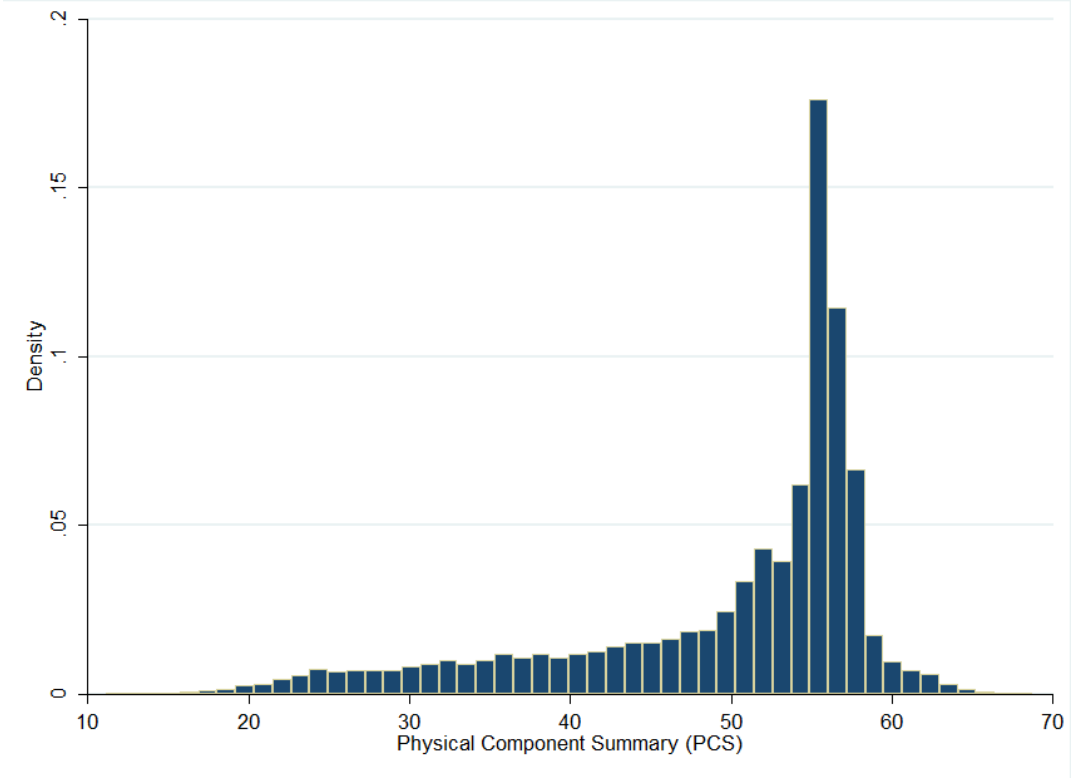


Figure 1.1: Physical Component Summary (PCS) histogram.

The six questions involved in determination of this index are:

- How is your general health?
- Are you currently limited because of health status in performing activities of moderate physical effort (such as moving a table, use a vacuum cleaner, bowling or take a bike ride, etc.)?
- Are you currently limited in going up several flights of stairs?
- During the last 4 weeks, have you been less efficient than you would like in working or other regular daily activities as a consequence of your physical health?
- During the last 4 weeks, have you had to limit some types of work or other activities as a consequence of your physical health?
- During the last 4 weeks, how much has physical pain hampered your normal work (both at home and outside)?

Chapter 2

Multilevel models

2.1 Introduction

In classical linear regression analysis, observations are independently and identically distributed (i.i.d.). This assumption is not realistic in many situations, such as hierarchical data. We have hierarchical data when observations (level 1 of hierarchy) are nested in groups (level 2 of hierarchy).
A classic example are pupils (level 1) nested in classes (level 2) and classes nested in schools (level 3) or, in the case of longitudinal data, units (level 2) observed at several occasions over time (level 1).
Dependencies between individual observations also occurs in survey researches, when the sample is not taken at random, but clustered sampling from geographical areas. In general, respondents from the same group will be more similar to each other (positively correlated), than respondents from different groups and therefore assumption of independent observations is violated.
If this assumption does not hold, the estimates of the standard errors are, in general, too small and lead to an error rate of the first type higher than the nominal level α.
For instance, if we consider achievements of pupils in mathematics as a response variable, dependency of the observations in the same class or school may stem from: sharing the same environment, if in the same class, sharing the same teachers or coming from the same neighborhood. In this study individuals belonging to the same household, the source of similarity on health characteristics may arise from hereditary factors, living in the same environment, having same habits and diet, as well similar social background and culture. In epidemiology we would expect to find more similar disease rates within the same geographical and administrative areas than across geographical and administrative not bordering areas.
The multilevel models are appropriate to address the dependence issue, but also

allow us to understand the relationships within and between clusters and to analyze and explain the variability at each level of hierarchy. In fact, we can view multilevel models as a regression with coefficients (intercept and slopes) that can vary by group. The terms "multilevel models" and "multilevel analysis" are used mainly in the social science (sociology, education, psychometric, etc.), while in other research areas are used terms as: hierarchical linear models, mixed models or random coefficient models.

2.2 Multilevel linear model

Before formalizing multilevel linear models, we consider the classic linear regression model. For example in the simplest case of one response variable and one explanatory (covariate) variable, model equation is

$$y_i = \beta_0 + \beta_1 x_i + e_i \tag{2.1}$$

with following assumptions:

a) $y_i \mid x_i \sim N(\mu_{x_i}, \sigma^2), \quad \mu_{x_i} \equiv E(y_i \mid x_i) = \beta_0 + \beta_1 x_i$

b) $Var(y_i \mid x_i) = Var(e_i \mid x_i) = \sigma^2$, homoskedasticity

c) $e_i \sim N(0, \sigma^2)$ i.i.d.

Proceeding with the above example, this could be a linear regression applied to a sample of school students where i indexes students, the response y is an attainment measure and x is a predictor, such as a prior test score. The residuals e_i are assumed to have a normal distribution with variance σ^2, and independently distributed. As pointed out above, this will not be generally true. Two randomly chosen students in the same school will tend to be more alike in their attainments, in this case adjusted for the predictor x, than two students chosen at random from different schools. As a result of the factors mentioned above.
One way accounting dependence between units is to extend model (2.1) as follows

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + u_j + e_{ij} \tag{2.2}$$

with following assumptions:

a) $u_j \sim N(0, \sigma_u^2)$ i.i.d.

b) $e_i \sim N(0, \sigma_e^2)$ i.i.d.

c) $e_{ij} \perp u_j \ \ \forall i, j$

where i indexes students as before and the additional subscript, j, indexes schools (or classes). We now have an explicit term, the school residual, u_j, which allows each school to contribute an "effect" to the response, i.e. to have a different intercept for each school crossing the y axis at a different point. The lines are parallel to each other and allow us to order the groups according to their average intercept (or random effect); for example we can order the schools according to average score of their students in mathematical tests (conditioning to x).

The model (2.2) is known as a "random intercept" or "variance components" model where σ_u^2 is the "between-school" variance and σ_e^2 is the "between-student" variance. In general, we can also assume that the school and student residuals are independent so that the total variance is given by: $\sigma_u^2 + \sigma_e^2$.
Whereas in equation (2.2) we have chosen to model the school effect as a random variable depending on a single parameter, the variance σ_u^2, an alternative to such a "random effects" model would be to fit school as a "fixed effect", using, for example, a set of m–1 dummy variables where m is the number of schools. In some special circumstances this may be preferred, but more usually we would wish to consider the set of schools (or geographical areas or households) as a randomly chosen sample from a population of schools (or areas or households) about which we wish to make inferences.

It is possible to introduce at school level (level 2) predictors into the model, such as the resources available to the school or the average prior attainment of all the pupils in the school, to ascertain their effects on the response variable. An important advantage of the random effects approach is that it allows us to do this straightforwardly, and also to examine the effects on the between-school variance. With the fixed effects model we cannot introduce further school level effects at all, since the available degrees of freedom have been taken up completely by the dummy variables. In a fixed effects model the coefficients of the dummy variables will provide direct estimates for the school effects.
From model (2.2) we can show that correlation between two students in the same school is

$$\rho = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}, \tag{2.3}$$

which is known as the *intraclass correlation coefficient* (ICC), and for this particular model (2-level random intercept), but not more generally, is also the proportion of the total variance due to schools, and called the *variance partition coefficient* (VPC). If we suppose that model (2.2) represents the "true" structure for the data but instead we fit model (1.2), as already noted, we shall obtain biased inferences. However, the estimates of the regression coefficients are generally consistent, but the standard errors are too small so that confidence intervals will be too small

and significance tests too optimistic, especially for level 2 predictors. Confidence intervals based on the simple regression estimate will be too short and significance test will reject the null hypothesis too often.

For many data structures, a single random effect for a higher level unit will not be adequate. Thus, in the schools example it is commonly found that the relationship with x will vary from school to school: the lines for each school are no longer parallel. This gives us a "random coefficient" model that can be written as

$$\begin{aligned} y_{ij} &= \beta_{0j} + \beta_{1j}x_{ij} + e_{ij} \ ; \\ \beta_{0j} &= \gamma_{00} + u_{0j} \\ \beta_{1j} &= \gamma_{10} + u_{1j} \end{aligned} \tag{2.4}$$

and the "combined model" becomes

$$y_{ij} = \gamma_{00} + \gamma_{10}x_{ij} + u_{1j}x_{ij} + u_{0j} + e_{ij} \ ; \tag{2.5}$$

in which the first 2 addends are called "fixed effects" or "fixed part" of the model and the last three terms are called the "random efects" or "random part" of the model because it contains the random variables. We assume that $\mathbf{u}$ is bivariate normal with variance-covariance matrix

$$\begin{pmatrix} u_{0j} \\ u_{1j} \end{pmatrix} \sim N \begin{pmatrix} \sigma^2_{u0} & \sigma_{u01} \\ \sigma_{u01} & \sigma^2_{u1} \end{pmatrix}, \ \mathbf{e} \perp \mathbf{u}$$

The only difference from the model (2.2) is that the coefficient of x is assumed to be random across schools, with mean γ_{10}, variance σ^2_{u1} and a covariance σ_{u01} with the random intercept term. Having a random coefficient also implies that the between-school variance, and hence the VPC, are (quadratic) functions of x (not unique ICC). In fact the variance of y given x is

$$Var(y_{ij} \mid x_{ij}) = \sigma^2_{u0} + 2\sigma_{u01} + \sigma^2_{u1}x^2_{ij} + \sigma^2_e \ ,$$

namely we have heteroskedasticity in the conditional distributions.

We can also insert a level 2 covariate, for example years of teacher experience w, in this way

$$\begin{aligned} y_{ij} &= \beta_{0j} + \beta_{1j}x_{ij} + e_{ij} \ ; \\ \beta_{0j} &= \gamma_{00} + \gamma_{01}w_j + u_{0j} \\ \beta_{1j} &= \gamma_{10} + \gamma_{11}w_j + u_{1j} \end{aligned} \tag{2.6}$$

Substantially at level 2, the level 1 coefficients become outcomes. The "combined model" becomes

$$y_{ij} = \gamma_{00} + \gamma_{01}w_j + \gamma_{10}x_{ij} + \gamma_{11}w_jx_{ij} + u_{1j}x_{ij} + u_{0j} + e_{ij} \ ; \tag{2.7}$$

in which the term $\gamma_{11} w_j x_{ij}$ in the fixed part is a "cross-level interaction".

We can insert several others covariates at level 1 (e.g. gender, age, ethnicity) and at level 2 (e.g. students per class) and expressing the model in matrix notation it becomes

$$Y_j = X_j \Gamma + Z_j U_j + E_j \tag{2.8}$$

where Y_j is a $n_j \times 1$ response vector for cluster j, X_j in a $n_j \times p$ design matrix for the fixed effects, Γ is a $p \times 1$ vector of unknown fixed parameters, Z_j is an $n_j \times r$ design matrix for the random effects, U_j is the $r \times 1$ vector of unknown normal random effects and E_j is the $n_j \times 1$ normal residual vector.
Clearly the number of parameters in a multilevel regression can easily become very large as we can inferred from (2.7 and 2.8). Even with a modest number of explanatory variables, multilevel regression analysis implies a complicated model. Generally, we do not estimate the complete model with all cross-level iterations, first because this is likely to get us into computational problems, but also because it is very difficult to interpret such a complex model. In general it is preferable to estimate more limited models that include only those parameters that have proven their worth in previous research, or are of special interest for the topic question.
Multilevel models can also be defined for three or more levels and the number of parameters increases further as well as their difficulty of interpretation.

Back to the model (2.4), if a categorical explanatory variable has a random coefficient then this can be interpreted as each category having a different level 2 variance (a particular instance of where the level 2 variance is a function of a level 1 predictor). This leads on to the idea of more general ways of modeling variation. Thus, for example, if we have a predictor variable "gender" then not only could we model different between-school variances for boys and girls, we could also consider modeling different between-student variances for boys and girls. The between-boy variance is allowed to be different from the between-girl variance. It turns out that we can formulate directly quite complex linear and non-linear models for the variance between units at any level of data hierarchy.

Another extension is to generalized linear models such as a logistic model for a categorical response. Supposing that we have a binary response, e.g. whether or not a student passes the math exam. A basic two-level model would have two components: the first expresses the probability of a positive response as a function of student and school characteristics in the "fixed part" of the model together with a level 2 residual:

$$logit(\pi_{ij}) = \beta_0 + \beta_1 x_{1ij} + \beta_j x_{2j} + u_j;\ u_j \sim N(0, \sigma_u^2) \tag{2.9}$$

where x_1 could be a student characteristic (number of hours of study, gender, ect.) and x_2, say, a level 2 school characteristic (public or not). Our actual response, y, is (0,1) and we specify the second part of the model where this has a Bernoulli distribution or binomial distribution with denominator 1 and probability π_{ij}.
The model (2.5) can be extended to ordered or unordered multi-category responses and also specified for other link functions.

2.3 Variance explained

In ordinary multiple regression analysis the squared multiple correlation R^2 is an important statistic interpreted as the proportion of variance modelled by the explanatory variables. In multilevel regression analysis, the issue of modelled or explained variance is a complex one. First, there are unexplained variances at several levels. This itself makes the proportion of explained variance a more complicated concept than in single-level regression analysis. Second, if there are random slopes, the model is inherently more complex, and the concept of explained variance has no unique definition anymore. Among the different approaches proposed to indicate how well we are predicting the outcomes in a multilevel model, a straightforward approach is the one that examine the proportion of explained variance in a sequence of models.
The "intercept-only model" or "null model"

$$y_{ij} = \beta_0 + u_j + e_{ij} \tag{2.10}$$

is a baseline model, because it does not introduce any explanatory variables (except the constant intercept term) and decomposes the total variance of the outcome variable into two levels. Obviously, if the level 2 variance in this model is zero, we can perform a classical linear regression model.

In the case of random intercept model, to calculate a statistic analogous to the multiple R^2, we must express the difference between null model variance at level 1 or 2 and the correspondent variance of the estimated model as a proportion of the total variance in that level (Raudenbush and Bryk, 2002).
Working separately level by level, for the proportion of variance explained at the first level we can use

$$R_1^2 = \frac{\sigma_{e|b}^2 - \sigma_{e|m}^2}{\sigma_{e|b}^2} \tag{2.11}$$

where $\sigma_{e|b}^2$ is the lowest-level residual variance for the null model and $\sigma_{e|m}^2$ is the lowest-level residual variance for the comparison model.

Similarly for the proportion of variance explained at the second level we use

$$R_2^2 = \frac{\sigma_{u|b}^2 - \sigma_{u|m}^2}{\sigma_{u|b}^2} \tag{2.12}$$

where $\sigma_{u|b}^2$ is the second-level residual variance for the null model and $\sigma_{u|m}^2$ is the second-level residual variance for the comparison model.
We proceed in the same way for the case of a random intercept model with more than 2 level of hierarchy.
A problem using the formulas above is that it is possible to arrive at the conclusion that a specific explanatory variable has a negative contribution to the explained variance, leading to a negative R^2, which is an impossible value[1].

In the random slope models the residual variances depend on the scale of the explanatory variables; there is no current solution for computation of a comparable R^2.

2.4 Estimation and hypothesis testing

The most frequent method to estimate the values of the regression coefficients and the intercept and slope variances is the maximum likelihood method. The maximum likelihood (ML) method is a general estimation procedure, which produces estimates for the population parameters that maximize the probability of observing the data that are actually observed, given the model.
An advantage of the maximum likelihood estimation method is that it is generally robust, and produces estimates that are asymptotically efficient, normal and consistent. The asymptotically properties in the case of multilevel models do not refers to the increasing of the total number of the observations, but is realized increasing the number of groups at the highest hierarchical level. With large samples, ML estimates are usually robust against mild violations of the assumptions, such as having not normal errors. The maximum likelihood estimation proceeds by maximizing a function called the likelihood function. Two different likelihood functions are used in multilevel regression modeling. One is "the full maximum likelihood" (FML); in this method, both the regression coefficients and the variance components are included in the likelihood function. The other estimation method is "restricted maximum likelihood" (RML); here only the variance components are included in the likelihood function, and the regression coefficients are estimated in

[1]Snijders and Bosker, 2012, cap. 7 explain way this result may happen and in which circumstances.

a second estimation step. Both methods produce parameter estimates with associated standard errors and an overall model deviance, which is a function of the likelihood. FML treats the regression coefficients as fixed but unknown quantities when the variance components are estimated, but does not take into account the degrees of freedom lost by estimating the fixed effects.
The degrees of freedom (df) in a multilevel model are calculated with reference to a number of units for the level 1, and to a number of groups for the level 2[2] . RML estimates the variance components after removing the fixed effects from the model. As a result, FML estimates of the variance components are underestimates (ignoring degree of freedom) if the number of groups is small. RML should lead to better estimates, when the number of groups is small. However FML is preferable, because it has two advantages over RML: the computations are generally easier and, since the regression coefficients are included in the likelihood function, a likelihood ratio test (LR) can be used to compare two models that differ both in the fixed part (the regression coefficients) than in the random part (the variances and covariances of random part). With RML, only differences in the random part can be compared with this test.
Computing the maximum likelihood estimates requires an iterative procedure. At the start, the computer program generates reasonable starting values for the various parameters (in multilevel regression analysis these are usually based on single-level regression estimates). In the next step, the procedure tries to improve on the starting values, to produce better estimates. This second step is repeated (iterated) many times. After each iteration, the program inspects how much the estimates have actually changed compared to the previous step. If the changes are very small, the program concludes that the estimation procedure has converged and that it is finished. There are several iterative algorithms mainly used: Fisher Scoring (proposed by N. Longford), Expectation-Maximum (EM) algorithm (implemented in Stata), Newton-Raphson algorithm (implemented in SAS).

Reports below, as an example, hierarchical construction in steps of the likelihood for the random intercept model like in (2.2). Denoting with $\theta = (\Psi, \sigma_u^2)$ the parameters vector, where σ_u^2 is the parameter of random effects (variance) and $\Psi = (\beta_0, \beta_1, \sigma_e^2)$ are all other parameters (fixed effects and the residual variance parameter). Than the likelihood function for an observed y_{ij} and fixed u_j are $L_{ij}(\Psi \mid u_j)$ (density of a normal $N(\beta_0 + \beta_1 x_{ij} + u_j, \sigma_e^2)$ for fixed u_j evaluated at the observed y_{ij}). Now the observations in the same group are conditional independent

[2]If we are testing the coefficient of a level 1 variable in which the total number of units is N and the total number of explanatory variables is r, we have $df = N - r - 1$. To test a coefficient of a level 2 variable in which the total number of groups is M and the total number of level 2 explanatory variables is q, we have $df = M - q - 1$. To test a coefficient of a cross-level interaction between level 1 variable x and level 2 variable w, when the model contains a total of s other level 2 variables also including with this variable x, we have $df = M - s - 1$.

given u_j and than the j-th cluster conditional likelihood is

$$L_j(\Psi \mid u_j) = \prod_{i=1}^{n_j} L_{ij}(\Psi \mid u_j) \tag{2.13}$$

Note that the equation defining a multilevel model likelihood includes the random effects, but they are unobservable, so the random effect must be integrated out. We obtain the marginal likelihood function of cluster j-th in this way

$$L_j(\theta) = \int L_j(\Psi \mid u_j)\ p(u_j|\sigma_u^2)\ du_j \tag{2.14}$$

Now for the independence between clusters the marginal likelihood is obtained by the product of the marginal cluster likelihoods

$$L(\theta) = \prod_{j=1}^{J} L_j(\theta)$$

and then we proceed to the maximization with respect to θ

$$\hat{\theta}_{ML} = \underset{\theta}{argmax}\ L(\theta)$$

obtaining maximum likelihood estimates.

Bayesian methods are often used even in the cases where ML estimation is difficult. An empirical Bayes parameter estimate for the j-th group is an optimally weighted average of the parameter estimates using the j-th group data and the entire sample data. For example, in the 2 level random intercept model as $y_{ij} = \gamma_{00} + u_{0j} + e_{ij}$, the "Empirical Bayes" (EB) estimate (prediction in reality) of the intercept $\beta_{0j} = \gamma_{00} + u_{0j}$ and denote with β_{0j}^{EB}, can be expressed as a weighted linear combination of $\hat{\beta}_{0j}$ and $\hat{\gamma}_{00}$:

$$\beta_{0j}^{EB} = \lambda_{0j}\hat{\beta}_{0j} + (1 - \lambda_{0j})\hat{\gamma}_{00} \tag{2.15}$$

where $\hat{\beta}_{0j}$ is the level-1 OLS (ordinary least squares) estimate of group mean ($\bar{y}_{.j}$) while $\hat{\gamma}_{00}$ is the OLS estimate of the grand (overall) mean ($\bar{y}_{..}$), of the outcome variable; λ_{0j} is the "reliability" of $\hat{\beta}_{0j}$. If all groups had λ_{0j}=1, the multilevel parameter estimates would be equivalent to those of OLS regression. With a high reliability of $\hat{\beta}_{0j}$ the EB estimate β_{0j}^{EB} would be weighted more heavily by the group mean $\hat{\beta}_{0j}$; otherwise, β_{0j}^{EB} would be weighted more by the overall mean $\hat{\gamma}_{00}$. When the group size n_j is small, the weight of reliability λ_{0j} in (2.15) would be small,

and then the contribution of $\hat{\beta}_{0j}$ to β_{0j}^{EB} would *"shrink"* and the multilevel model parameter estimate would be pulled toward the overall mean. The EB estimator is, also, called a *shrinkage estimator*. The degree of shrinkage depend on the precision of the level-1 OLS estimate or its reliability. The shrinkage estimator approach is even called a *borrowing strength* approach; that is, it borrows information from all the groups to support statistical estimation for the groups with insufficient observations. In the "Null model" the reliability can be computed as

$$\lambda_{0j} = \frac{\sigma_{u0}^2}{\sigma_{u0}^2 + (\frac{\sigma_e^2}{n_j})} \tag{2.16}$$

which can also be written as a function of the ICC, ρ:

$$\lambda_{0j} = \frac{n_j \rho}{1 + (n_j - 1)\rho} \tag{2.17}$$

Note that the reliability is a function of n_j and of ICC: for a given n_j, a larger ICC leads to a larger reliability λ_{0j}, and for a given ICC a larger n_j lead to a larger reliability.
The "posterior means" (so called in Bayesian theory), β_{0j}^{EB}, can be used to see which groups have an extraordinarily high or low values on the response variable, given their values on the explanatory variables. They can also be used in a residual analysis, for checking the assumption of normality for the random effects, and for detecting outliers. For this purpose we compute the EB residuals as

$$\hat{u}_{0j}^{EB} = \hat{\beta}_{0j}^{EB} - \hat{\gamma}_{00} = \lambda_{0j}(\bar{y}_{.j} - \bar{y}_{..}) \tag{2.18}$$

where $(\bar{y}_{.j} - \bar{y}_{..})$ is the OLS residual of the level 2 regression multiplied by the reliability.

Several alternative methods for estimation exist as Iterative Generalized Least Squares (IGLS) proposed by H. Goldstein, Generalized Estimating Equations (GEE), Bayesian methods and Bootstrap methods. The Bootstrap methods can be used to improve the parameter estimates and the standard errors. When group size is small or the distribution of the outcome variable is not normal, bootstrap methods can be used for robust multilevel modeling.

Test for random effects: this means to test if the variance and covariance components (e.g. σ_{u0}^2, σ_{u1}^2, σ_{u01}) are equal to zero. The LR-test used in general for comparisons between model with a different number of parameters and therefore for this purpose inherent a different number of random part parameters. Nevertheless, in the case of variance (which cannot be negative - we have a one tail

test) the p-value of the Chi-square test must be divided by 2. In alternative a one-tailed Wald Z test (computed as the ratio of the parameter estimate to its standard error) can be used. These two test are asymptotically equivalent if the number of group is huge, so the LR test is to prefer.

Test for fixed-effects: to test a significance of a single coefficient can be used a Wald Z test; for simultaneously test on several coefficients (as for the dummy variable coefficients relating to one categorical variable) LR test can be used.

Chapter 3

Models specification for Physical Component Summary

3.1 Introduction

In household surveys, hierarchical structure is characterized by a small number of units n_j (subjects) per groups j (households), but in general by a very large number of groups J. In multilevel models having a large number of groups is better than having few groups of large size. Table 3.1 shows size and number of level-2 groups in the sub-sample (without subjects under 18 age, consequently level-2 group sizes do not coincide with household sizes).

Table 3.1: Level 2 sizes (households) in the sub-sample of 18 aged and over.

	Number of subjects		Number of groups	
Level 2 size	Freq.	Percent	Freq.	Percent
1	14,453	13.66	14,453	28.63
2	45,386	42.88	22,693	44.96
3	25,290	23.89	8,430	16.7
4	15,768	14.9	3,942	7.81
5	4,075	3.85	815	1.61
6	696	0.66	116	0.23
7	168	0.16	24	0.05
8	8	0.01	1	0
Total	105,844	100	50,474	100

For multilevel linear models with normal random effects at each level, having "small n and large J" scenario generally creates no problems in statistical inference of estimating fixed regression coefficient. This result does not extend to the case of non-linear link functions (for example logistic regression) and non normal random effects.
While, when we wish to study random coefficients (cluster-specific intercepts or slopes) or to estimate variance and covariance components at the second level of the hierarchy, we need that the fit of the model at level 1 is very good in order to obtain good estimate. In essence, the lack of data at level 1 requires more restriction, namely the slopes do not vary (Roudenbush, 2008).
With small sample sizes at the group level, variance components are more susceptible to bias; the estimate tend to be estimated too small with standard errors that may be biased downward.
One method for obtaining better tests and confidence intervals is to correct the asymptotic standard errors, using the so-called Huber/White or *sandwich estimators* which will be used here because the outcome variable is not-normal and highly skewed. In effect, sandwich estimators are recommended if the normality and homoscedasticity assumptions of the hierarchical linear model are not satisfied. Also sandwich estimators require an adequate number of units at the highest level of hierarchy. The use of these estimators in multilevel models may involve larger standard errors especially for the parameters of the random part of the model (Maas and Hox, 2004).

Taking into account the previous methodological remarks, we will proceed to model a multilevel linear model with only random intercepts. The parameters will be estimated using the FLM (full maximum likelihood) method and the standard errors will be calculated using cluster-robust method (i.e. with sandwich estimator). The term *cluster* here refers to a variable indicating the level for which the observations are considered independent, that are the highest-level units in a multilevel data structure. In our case the cluster variable for robust estimator, due to computational problems, will be the 103 provinces at the time of the survey.
The software used to carry out the analysis was mainly Stata version 12.0; it allows the inclusion of sampling weights in multilevel model estimate.

3.2 Null models

Null models are progressively constructed for each possible hierarchical level in order to check if there is correlation between observations and groups belonging to the same cluster. The null models are constructed according to formulas given

below in which u identifies household (level 2) random effects, v is the municipality (level 3) random effects and so on until (level 6) random effects referred to geographical area γ; e are level-1 residuals.

$$y_{ij} = \beta_0 + u_j + e_{ij}$$
$$y_{ijk} = \beta_0 + u_{jk} + v_k + e_{ijk}$$
$$\dots$$
$$y_{ijksrg} = \beta_0 + u_{jksrg} + v_{ksrg} + \xi_{srg} + \zeta_{rg} + \gamma_g + e_{ijksrg}$$

Table 3.2: Null models with 2, 3 and 6 levels of hierarchy

MODEL	Null-2L		Null-3L		Null-6L	
Fixed effects	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Intercept	49.62	0.03	49.57	0.06	49.52	0.22
Random part	Parameter	S.E.	Parameter	S.E.	Parameter	S.E.
6-Geographical area: var(γ)					0.18	0.16
5-Region: var(ζ)					0.07	0.09
4-Large area: var(ξ)					0.25	0.09
3-Municipality: var(v)			2.32	0.17	1.88	0.15
2-Household: var(u)	25.72	0.48	23.69	0.47	23.69	0.47
1-Residuals: var(e)	69.11	0.43	69.11	0.43	69.10	0.43

Table 3.3: Variance Partition Coefficients (VPC)

MODEL	Null-2L	Null-3L	Null-6L
6-Geographical area: var(γ)			0.19
5-Region: var(ζ)			0.07
4-Large area: var(ξ)			0.26
3-Municipality: var(v)		2.44	1.97
2-Household: var(u)	*27.12*	24.90	24.89
1-Residuals: var(e)	72.88	72.66	72.61

In the 3 level model we note that: 1) the 3th level variance is quite low; 2) the correlation between two persons of the same municipality and household is equal to

$corr(y_{ijk}, y_{i'jk}) = \frac{\sigma_v^2+\sigma_u^2}{\sigma_v^2+\sigma_u^2+\sigma_e^2}$=27.34%

and it is not more greater than the ICC of model with 2 levels ($\rho = 27.12\%$). Therefore, even if the LRtests are conservative for all these null models with respect to the next smaller in number of level, hereinafter we will use the model with only 3 levels in consideration of the poor variability at the highest hierarchical levels. Moreover, the number of groups at levels 5 and 6 of hierarchy is very small to carry out a multilevel analysis.

3.3 Models selection

Given explanatory variable asymmetry, various attempts were made to normalize it, also applying following transformations on data: logarithmic, square root, cube root, reciprocal and Box-Cox. None of the various transformation held satisfactory results. Therefore we proceeded using the variable as it is, but with robust estimates for coefficients and parameters standard errors (sandwich estimator) as previously stated.
Looking at PCS histograms by age groups (Fig. 3.1) evident differences appear: the shape distribution changes from negative skewed in young people ages to positively skewed in older ages and there is presence of heteroscedasticity.

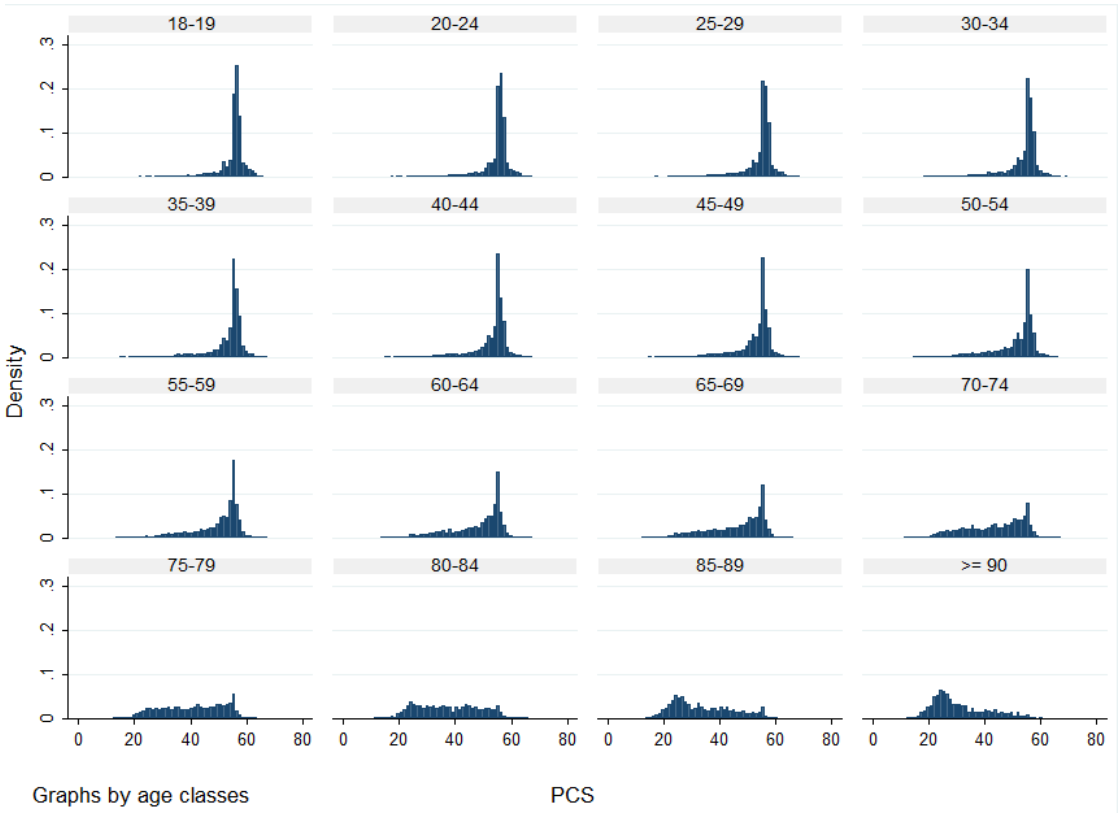


Figure 3.1: Physical Component Summary (PCS) histogram by age groups.

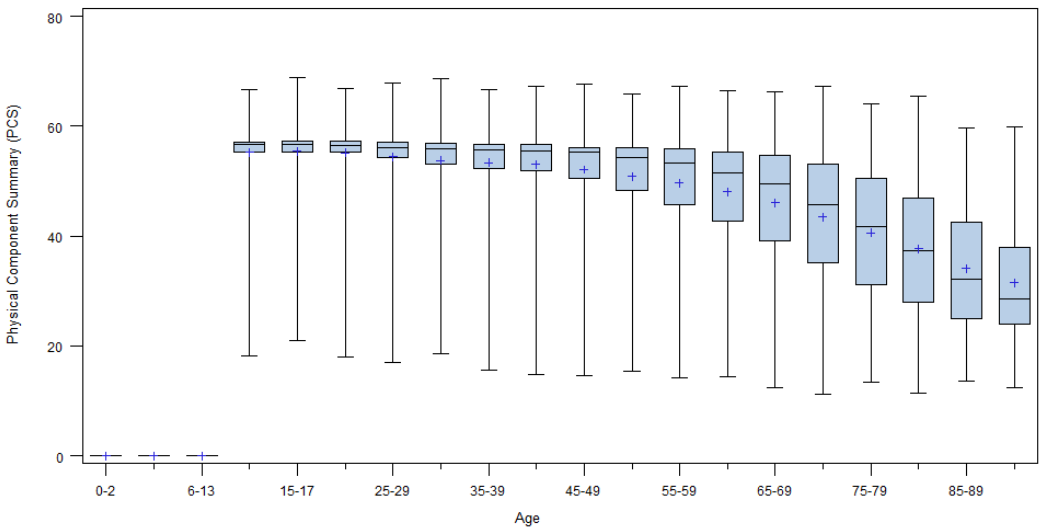


Figure 3.2: Box-plot of Physical Component Summary (PCS) by age groups: horizontal bar in middle of each box represents the median, + represents the mean.

In order to model heteroschedasticity at level 1 we add a dummy variable, which multiplies level-1 residuals. This variable indexes appropriate combinations of age groups and gender as homogeneous as possible within, in terms of average and variance on PCS score.

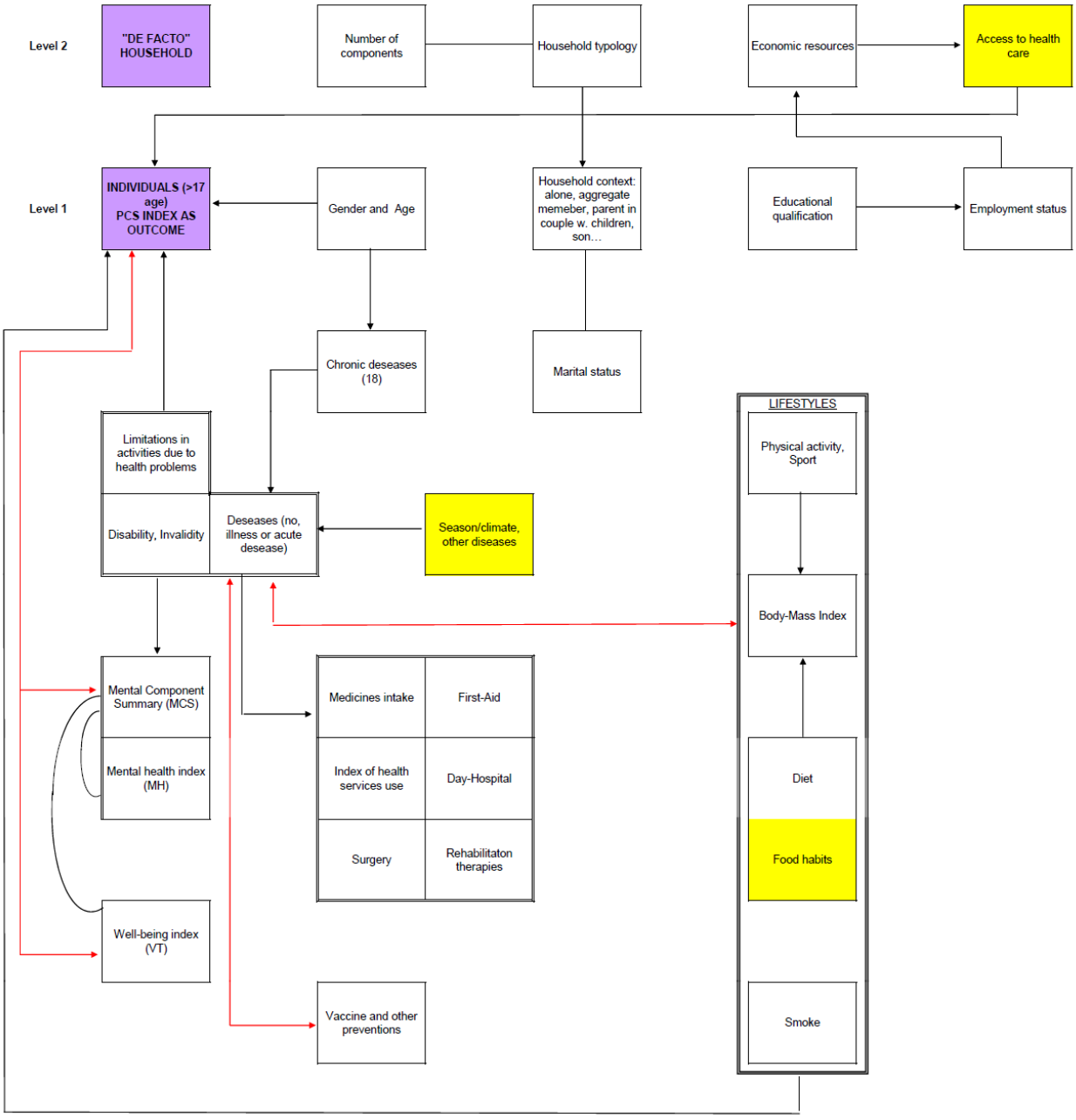
Other covariates at level 1

Model building strategies can be either top-down or bottom-up. The top-down approach starts with a model that includes the maximum number of fixed and random effects that are considered for the model.This procedure is described by West et al.(2007). The opposite strategy is mostly used, which is bottom-up. This strategy is described by Hox (J. Hos, 2010).

Given the large number of possible covariates to test we have decided to adopt the bottom-up approach. We have followed this procedure because initially the very large sample size involves significance on almost every available variables in the data set. First socio-demographic variables were considered.

In addition to age and gender, the following explanatory variables were tested: nationality, marital status, qualification as number of years of study, qualification as categories, professional status, work experience, employment sector of economy, occupational position and household context of the respondents. Among these

two were found to be more explanatory: educational qualification (as categorical) and employment status. Explanatory power of employed variable is mainly due to "unable to work" category.
For remaining covariates and indices on health conditions, prevention, chronic disease etc. it was necessary to reason on diagram showing relationship and connections between PCS and other variables in the data set (Fig.3.3). All these variables are significant and also associated each other.
Among lifestyle variables, only *Body-mass index* and *sport* practice result statistically significant. Question linked to the second of these variables distinguishes among sedentary, active (persons who make movement such as walking or have a job that involves physical effort) and people who make sport; only who perform sportly or agonistic activity has a significant advantage on PCS score.
In the same way, variables of the group regarding *medicine consumption, rehabilitation therapies etc.* are significant, but have been excluded because can be view like a consequences of affirmative response on other variables detected, such as chronic diseases, disability and illnesses or accidents.
Variable*Limitations in activities* is largely overlapping disability and invalidity and refers to a time period greater than the PCS index (last 6 months instead of the last 4 weeks), so it was excluded from the analysis. Between disability and invalidity we decide for *disability* variable since more accurate in picking the real physical limitations.
The variable *disease* is included in the model because it is both a possible expression of symptoms resulting from having chronic diseases and expression of disorders and illnesses which derive from other causes (e.g. seasonal illness, accidents, injuries, other diseases not considered among chronic ones). It provides three modality that distinguish among having no health problem, illness or acute disease in the past 4 weeks.



LEGEND:

Yellow: variable not available

Arrow: causal relation

Bidirectional red arrow: reciprocal interdipendence

Line: connection

Curved line: presence of common questions

Figure 3.3: Diagram showing (not exhaustive) connections between groups of covariates and the PCS index.

With regard to chronic diseases synthetic *chronicity index* has been initially used, in order to achieve a parsimonious model. This index has high correlation with PCS (ρ =-0.6,) but it is a consequence of construction methodology, so considering all the collected *chronic diseases* was the only way to proceed and it also allows us to sort each of them by impact on physical health of the respondents. With the intent of a parsimonious model, *number of chronicity* has being included once as continuous and once as categorical divided into classes. Best fitting in term of likelihood remains the one previously described.
The three indices Mental Component Summary (MCS), Vitality (VT) and Mental Health (MH) do not have a poor correlation with PCS (ρ =0.26, ρ =0.56, ρ =0.43 respectively), and are very similar to each other in part due to common questions involved in the indexes. Therefore it will not be possible to enter all of these, but at most one alone. Clearly being sick influences the psychological/mental status and vice versa, so each of them were initially excluded from the model performing the **Model 1**.

However, by testing their inclusion one at a time and comparing the obtained models, VT seems to have less noise than MH and MCS index on other coefficients and in capability to improve likelihood; also the standard errors are reduced compared with others. Beside a study conducted on SF-36 questionnaire accuracy found that VT is the sub-domain (with the exception of Social Functioning sud-domain) that giving minor contribution to PCS summarizing (C. Taft at al.,2001). Observing questions involved in VT summary, it seems to be an expression of physical well-being related to effort ability. So, for a more intuitive interpretation of the results we renamed it "well-being". Well-being (VT) have strong impact on Depression and Alzheimer coefficients. After entering VT in the model, the depression coefficient changes from negative to positive, then we removed it; Alzheimer lose significance, but we decided to retain it.
So we obtain a second model: **Model 2**.

In order to render VT coefficient easier interpretable, it was centered on sample mean (62.8) and divided by 10.
A similar transformation is applied to Body-mass index (BMI). According to World Health Organization (WHO) definition, BMI is a simple index of weight-for-height that is commonly used to classify underweight, overweight and obesity in adults. It is defined as the weight in kilograms divided by the square of the height in metres (kg/m^2). Cut-off points for these category are: BMI ¡18.5 = underweight, BMI in 18.5-24.9 =normal-weight, BMI in 25-29.9=overweight and BMI¿30=obese. In order to reduce the number of model parameters, BMI will be used as continuous variable. An other motivation related with this choice stems from worsening of

likelihood in the model having BMI divided by classes.
For the same reasons described for VT covariate, BMI will be mean-centered (24.8) and divided by 5. In fact the interval scores between BMI average (normal-weight) and overweight, and between overweight and obese is about 5, so increasing the transformed BMI of one unit we obtain the effect of transition, for example, from normal-weight to overweight and so on.
Several interactions could be added, but the only one between Disabled and Diseases was tested.

Covariates at level 2 and over .
The available variables at level 2 are not so many. We tested the significance of the following: number of household components (both as continuous and as categorical), dichotomous variable distinguishing one-person household and multi-component household, household typology, housing condition index, economic resources. Even if the one-person household variable was significant,the more informative variable "household composition" (number of household member as categorical) and "household typology" ware been inserted. "Household typology" distinguishes among the following categories: living alone, couple without children, couple with children and other typology (single parent with children, household with added member, two or more unrelated adults, ect).
"Household composition" was inserted both to explore the impact of different number of household membership on outcome variable and to model heteroscedasticity at level 2 in conditional distribution of PCS. After a careful analysis of level 2 residuals, the benefit in level 2 heteroscedasticity modelling appear not appreciable, moreover computational effort increases, so the intention was discarded.
"Household typology" replace "Household composition" variable in model 2. This change was made to allow any comparison with other studies in which it is most commonly used. As a result an additional model was estimate: **Model 3**.

The explanatory variable on household economic resources was significant and it distinguishes between household with good or adequate resources and those with poor or insufficient resources according to householder statement. Remaining ones are not significant.
At the 3rd level (municipality) we tested several external variables without success: latitude, longitude, altitude, self-representative municipality (AR) as defined by sampling design, coastal municipality, mountain and urban degree. Only the latter is significant.

Urban degree is an index provided by Istat and consists of 3 categories:

- *High*: densely populated areas, built by aggregation of local units having contiguous territory, higher density to 500 inhabitants per square kilometre and total amount of population at least 50,000 inhabitants.

- *Intermediate*: zones obtained by aggregation of local units territorial not belonging to the previous group, with a density more than 100 people per square kilometre and which have one of this added characteristics: total population over 50,000 inhabitants or neighbouring to areas belonging from previous group.

- *Low*: remaining areas that have not been classified in previous groups.

Searching for a possible contextual effect, we have test significance for several external variables:

- variables having regional detail: relative poverty incidence, household expenditure for health care consumption, government expenditure on health;
- variables having provincial detail: average amount of cash at banks by households, value added per capita;
- variables having municipality detail: average per capita of taxable IRPEF[1] income (only for Veneto region);

all these were found not significant.

3.4 Final models

Covariates selection process led to estimate 3 final models. All these models are carried out using robust estimators.
Final Models are quite large and contain estimation on 58 (model 1 and 2) and 57 (model 3) parameters relative to 29 explanatory variables (including 18 chronic diseases) plus 14 random part parameters (standard deviation).

Parameter estimates and their associated standard error estimates are presented in Tables 3.4 (fixed part) and 3.5 (random part) for null model and final models 1, 2 and 3.

[1]Personal Income Tax.

Table 3.4: Final models: Fixed part

MODEL	Model 1		Model 2		Model 3	
Fixed effects	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Intercept	54.96	0.16	53.59	0.18	53.58	0.17
Age	0.003***	0.005	0.028	0.005	0.028	0.005
Age^ 2	−0.001	0.000	−0.001	0.000	−0.001	0.000
Female	−0.53	0.04	−0.10**	0.04	−0.10**	0.04
Qualification: ref. No Educ. Qualif.						
University	1.53	0.13	1.32	0.13	1.32	0.13
Upper Secondary	1.41	0.11	1.18	0.11	1.18	0.11
Lower Secondary	1.07	0.10	0.86	0.10	0.86	0.10
Primary	0.65	0.11	0.52	0.11	0.52	0.11
Employment: ref. Employed						
Unemployed	−0.01***	0.10	−0.24*	0.09	−0.25*	0.09
Not working	−0.11**	0.05	−0.33	0.05	−0.32	0.05
Unable to work	−5.85	0.29	−5.16	0.27	−5.17	0.27
Body-mass Index	−0.38	0.03	−0.38	0.03	−0.37	0.03
Sport	0.79	0.04	0.50	0.04	0.50	0.04
Disabled	−10.18	0.31	−8.93	0.28	−8.94	0.28
Disease: ref. No health problem						
Ilness	−1.85	0.09	−1.41	0.09	−1.41	0.09
Acute disease	−6.81	0.11	−6.01	0.10	−6.01	0.10
Interaction: Disabled x Disease						
Disabled-Ilness	0.55***	0.36	0.47***	0.32	0.47***	0.32
Disabled-Acute disease	3.14	0.35	3.23	0.31	3.23	0.31
Well-being (VT)			0.98	0.02	0.98	0.02
Infarct	−3.14	0.22	−2.75	0.20	−2.74	0.20
Heart disease	−2.89	0.14	−2.45	0.13	−2.45	0.13
Stroke	−2.75	0.21	−2.18	0.20	−2.18	0.20
Arthrosis,Arthrit.	−4.61	0.10	−4.00	0.09	−3.99	0.09
Osteoporosis	−2.27	0.13	−2.04	0.12	−2.04	0.12
Hepatic cirrhosis	−2.60	0.52	−1.96	0.49	−1.96	0.49
Cancer	−4.51	0.32	−3.73	0.29	−3.73	0.29
Parkinson	−2.30	0.38	−1.63	0.35	−1.63	0.35

Continued on next page

– *continued from previous page*

MODEL	Model 1		Model 2		Model 3	
Fixed effects	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Ulcer	−1.11	0.18	−0.76	0.17	−0.76	0.17
Cancer in the past	−1.60	0.19	−1.30	0.18	−1.30	0.18
Depression	−0.51	0.12				
Alzheimer	−1.70	0.36	−0.76**	0.34	−0.77**	0.34
Stones	−1.43	0.16	−1.21	0.14	−1.21	0.14
Thyroid disease	−0.92	0.13	−0.68	0.12	−0.68	0.12
Asthma	−1.31	0.12	−1.11	0.12	−1.11	0.12
Diabetes	−1.62	0.13	−1.35	0.13	−1.35	0.13
Hypertension	−0.61	0.08	−0.45	0.07	−0.45	0.07
Bronchitis,Emph.	−1.26	0.13	−0.88	0.11	−0.88	0.11
Household composition: ref. 1 component						
2 components	−0.28	0.07	−0.28	0.07		
3 components	−0.38	0.07	−0.35	0.06		
4 components	−0.30	0.07	−0.26	0.07		
5 or more comp.	−0.26**	0.11	−0.28*	0.10		
Household typology: ref. Living alone						
Childless couple					−0.35	0.07
Couple with child.					−0.29	0.06
Other typology					−0.22**	0.09
Insuf. econ. resources	−0.54	0.06	−0.21	0.05	−0.21	0.05
Urban degree: ref. Low						
Intermediate	0.20*	0.07	0.24	0.07	0.24	0.07
High	0.39	0.08	0.37	0.08	0.37	0.08

*Coefficient not significant: *** "p¿0.05"; ** "p¿0.01"; * "p¿0.001"*

In the three final models estimated standard errors are small and slightly higher only for coefficients related to low sample frequencies (such as Hepatic cirrhosis, Parkinsonism, Alzheimer, Unable to work) and for interaction coefficients (Disabled-Illness and Disabled-Acute disease), denoting high precision of the estimates due to large sample size.
Specifically we note that in model 2 and 3 standard errors are smaller or equal than in model 1 due to entering of well-being (VT) index. Beside insert VT in

model 2 leads to a reduction on many coefficients related to chronic diseases and female. Loss of significance on sex covariate after VT entering seems to explain differences between men and women on PCS score. In our sample both PCS average (M=51, F=48.9) and well-being (VT) average (M=66.3, F=59.6) show lower score for women than for men. Gender differences in health are well know and widely studied. Many researches highlight a lower level of women health measures after adjusting for several socio-demographic and socio-economic variables (D. Cherepanov et al., 2010). The different result of model 2 estimates reveals that a possible gender difference on perceived physical health may be due to women feeling less full of energy and lively but more tired and worn out than men (see Appendix B - Well-being (VT) questions).
Such explanation is plausible for the reduction of the chronic disease coefficients: be affected by chronic disease will generally be feel less lively, full of energy and more frequently tired than others.
Coefficient estimates on qualification indicate positive association between subjects PCS score and individual level education. Comparable association between educational qualification and level of health has been shown in many preview studies (G. Costa et al.,2003). Between Employed and Unemployed status there is not a significant difference on response variable in each models, while Not Working becomes significant in models 2 and 3, moreover interpretation of this effect is non easy because within this category converge different type of subjects: students, retirees, housewives, unoccupied. Strong negative effect on PCS score due to being unable to work emerge in all three models. The health risk factor role of BMI and Sport activity is commonly known. In recent years research has been conducted to estimate the impact of obesity on Quality of Life (QoL): obesity-associated decrements on QoL tend to be most pronounced on physical domains of functioning. It is interesting to note that these variables have an effect even even in our models. Transition from normal-weight to overweight and from overweight to obese leads an average score lower of more than a third of point. We expected a greater impact, that is probably partly caught by chronic diseases covariates.
Naturally, disability is a factor of great negative influence on physical health: the highest in the models among explanatory variables.
Another strong negative effect on PCS is attributable to been affected by a Acute disease (appendicitis, pneumonia,etc.) in the last 4 weeks.

If Disability and Acute disease are both present, the average impact on outcome variable is not the sum of related coefficient, but it is mitigated for about 3 score.

Chronic diseases all have naturally negative (from mild to strong) contribute

on physical health perceived. The most penalizing are: Arthrosis and Arthritis, Cancer, Infarct, Heart disease and Stroke.

Level-2 independent variables, "Household composition" (model 1 and 2) and "Household typology" (model 3) have as referent category *one person household*. These covariates bring out an average PCS score lower against who do not live alone. Who lives in households with more than one component seems to have a bit worse physical health than someone who lives alone. The negative magnitude is similar in all category, with the exception for who lives in *other household typology* or in *large household* (5 or more components). Similar evidence was found by G. Costa (Costa et al., 2003).

Households declaring scarce economic resources have a negative effect especially in model 1. The negative relationship between socio-economic status (SES) and health has been documented in many contest (G. Smith et al. 1990).
Finally is of interest to observe "urban degree" effect. People living in high urbanization areas have a slightly better PCS score. Probably this variable is able to capture the easier access to medical cure in larger city. Comparable effects have been highlighted in other studies with similar covariates (G. Costa et al., 2003; R. Edwing et al., 2003).

Any differences emerge between random part of the three final models.
Municipality standard deviation of random effects is small but still significant compare with model without level 3 and tested using an LR- test. At level 2 we note that household intercept variance does not decrease so much compared to null model; this is probably due to small number level-2 covariates included in the models.
Residuals have a standard deviation increasing with age and higher for females than for males.
Estimate parameters of model 2 and 3 are smaller than model 1. We conclude that PCS variability is explained mainly by level-1 variables.

Table 3.5: Final models: Random part

MODEL	Model 1		Model 2		Model 3		Null model	
Random effects	Paramet.	S.E.	Paramet.	S.E.	Paramet.	S.E.	Paramet.	S.E.
3-Municipality: sd(v)	0.71	0.03	0.70	0.04	0.70	0.04	0.97	0.05
2-Household: sd(u)	1.82	0.03	1.79	0.03	1.79	0.03	2.42	0.07
1-Residuals: Independent by Sex and Age								
F ≤29: sd(e)	4.58	0.09	4.61	0.09	4.61	0.09	5.30	0.09
F 30-49: sd(e)	5.53	0.06	5.39	0.06	5.39	0.06	6.39	0.08
F 50-64: sd(e)	6.87	0.06	6.55	0.06	6.55	0.06	9.65	0.14
F 65-74: sd(e)	7.63	0.08	7.18	0.08	7.18	0.08	13.39	0.21
F 75-84: sd(e)	8.13	0.10	7.58	0.10	7.58	0.10	17.78	0.22
F ≥85: sd(e)	7.87	0.13	7.36	0.12	7.36	0.12	22.38	0.28
M ≤29: sd(e)	4.11	0.10	4.08	0.10	4.08	0.10	5.04	0.10
M 30-49: sd(e)	5.00	0.06	4.87	0.06	4.87	0.06	5.95	0.07
M 50-64: sd(e)	6.02	0.08	5.70	0.07	5.70	0.07	8.00	0.13
M 65-74: sd(e)	7.19	0.09	6.69	0.09	6.69	0.09	11.26	0.21
M 75-84: sd(e)	7.87	0.10	7.26	0.09	7.26	0.09	15.32	0.24
M ≥85: sd(e)	8.09	0.25	7.41	0.26	7.42	0.26	20.19	0.34

Explained variances are reported in Table 3.5 and performed by the analogous R^2 of classical linear regression as

$$R_l^2 = \frac{\sigma^2_{null} - \sigma^2_{estimate}}{\sigma^2_{null}}$$

Model 2 and 3 perform better than model 1. The percentage of variance explained is consistent at all levels. We note that variances explained at level 1 increases with age for both genders.

Table 3.6: Variance and percentage of explained variance of final models with respect to null model: (R_l^2).

	Variance				**R_l^2**		
MODEL	Null	Mod.1	Mod.2	Mod.3	Mod.1 (%)	Mod.2 (%)	Mod.3 (%)
3-Municipality: sd(v)	0.93	0.50	0.49	0.49	46.82	47.95	47.96
2-Household: sd(u)	5.86	3.30	3.19	3.19	43.64	45.56	45.57
1-Residuals: Independent by Sex and Age							
F ≤29: sd(e)	28.07	20.95	21.26	21.25	25.35	24.27	24.28
F 30-49: sd(e)	40.84	30.60	29.04	29.04	25.06	28.88	28.89
F 50-64: sd(e)	93.08	47.20	42.88	42.87	49.29	53.94	53.94
F 65-74: sd(e)	179.27	58.27	51.59	51.59	67.50	71.22	71.22
F 75-84: sd(e)	316.01	66.08	57.41	57.42	79.09	81.83	81.83
F ≥85: sd(e)	500.70	61.92	54.10	54.14	87.63	89.19	89.19
M ≤29: sd(e)	25.40	16.85	16.67	16.66	33.64	34.38	34.39
M 30-49: sd(e)	35.40	25.02	23.73	23.74	29.32	32.95	32.94
M 50-64: sd(e)	64.04	36.26	32.49	32.49	43.38	49.28	49.27
M 65-74: sd(e)	126.85	51.76	44.72	44.73	59.19	64.74	64.73
M 75-84: sd(e)	234.72	61.98	52.71	52.74	73.60	77.54	77.53
M ≥85: sd(e)	407.44	65.40	54.97	54.99	83.95	86.51	86.50

A residual analysis at each level can help to check for assumption of normal distribution. Following quantile-quantile plot graphics report standardized residuals at level 1 and standardized Empirical Bayes (EB) random effects at level 2 and 3 for final models 1 and 3. Gaussian distribution at level 3 is roughly satisfied, while at levels 2 and 1 not hold the hypothesis does not hold (bad fit on the tails). Scatter-plot of standardized residuals and provided random effects reveal some heteroscedasticity particularly at level 1.

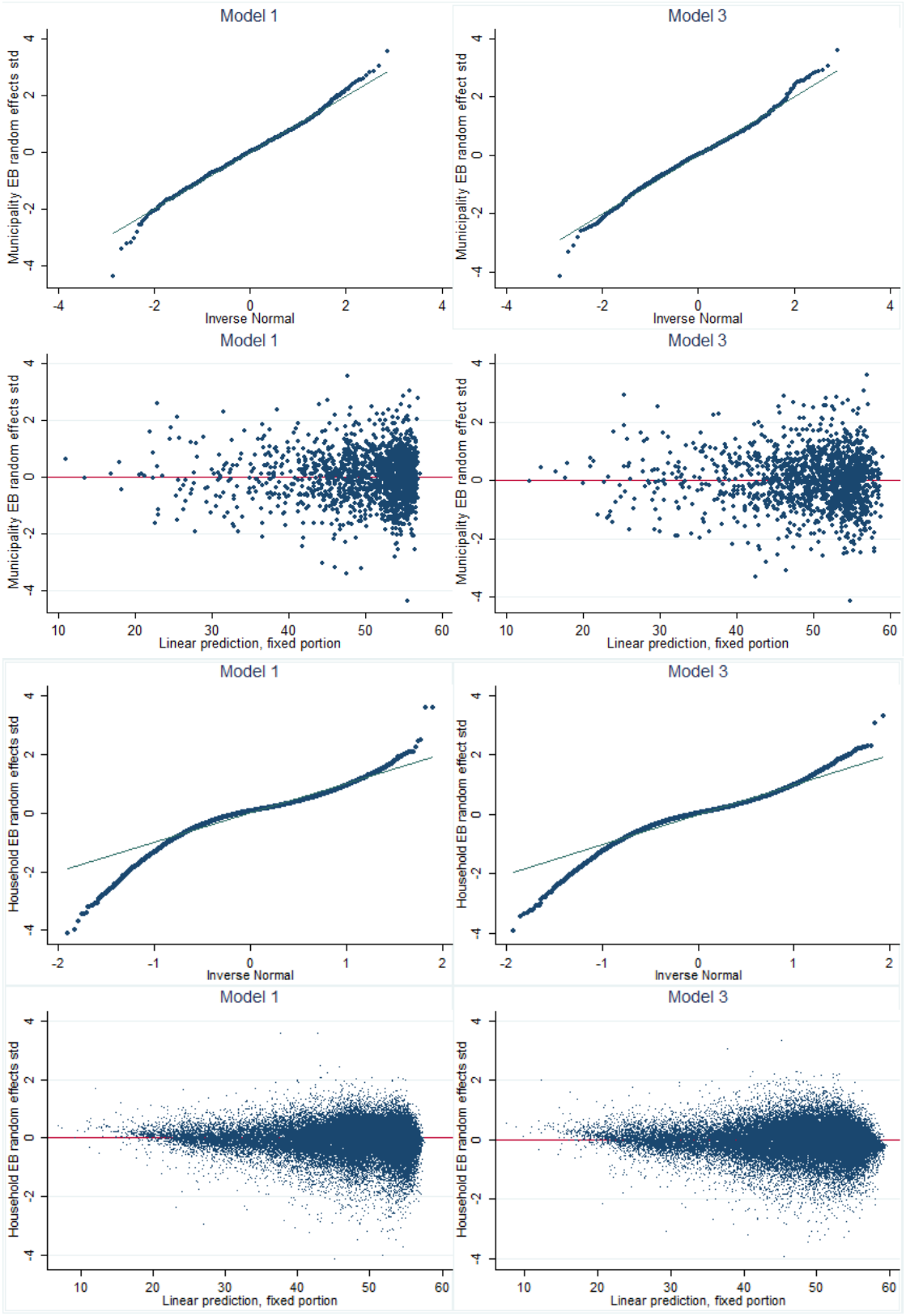


Figure 3.4: Quantile-quantile plot of standardized random effects and scatter-plot of standardized random effects versus fixed part predictions for Model 1 and Model 3.

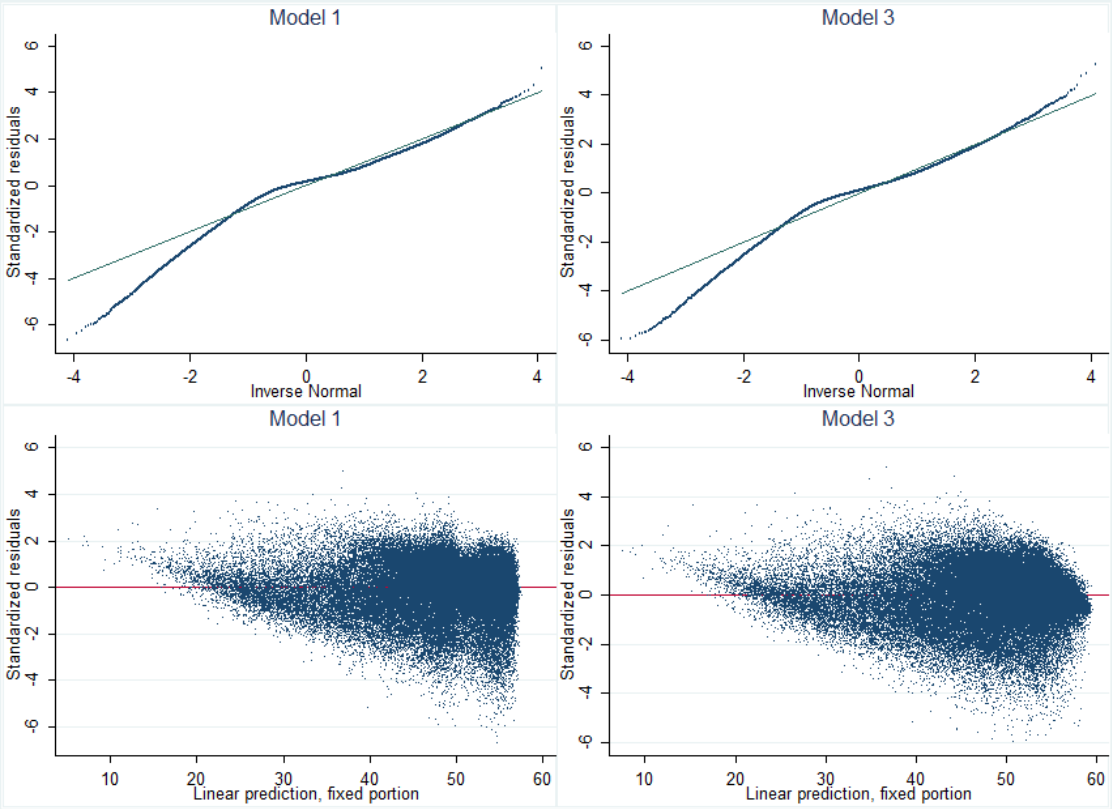


Figure 3.5: Quantile-quantile plot of standardized level 1 residuals (top panel) and scatter-plot versus fixed part prediction (bottom panel) for model 1 and 3.

An analysis of provided municipality random effects will show presence of municipality with intercept significantly different from zero. In this aim we display the so-called Caterpillar graphic for model 3. These graphics show random effects ranked from smaller to largest and related 95% confidence interval. The high number of units at level 3 (1465) makes impossible to show all them, therefore we show only the 35 lowest and the 35 highest random effects. These graphics can not bring out all municipalities having non-zero random effect. Municipalities with intercept different from zero are listed in table 3.7.
Municipality having negative intercept are 29, municipalities having negative random effects are 30. Note that non-zero municipality random effects is not due to little group-size and are equally distributed across regions and geographical areas. That is, estimated model 3 explains territorial differences. It is also evident in the map (Figure 3.7).

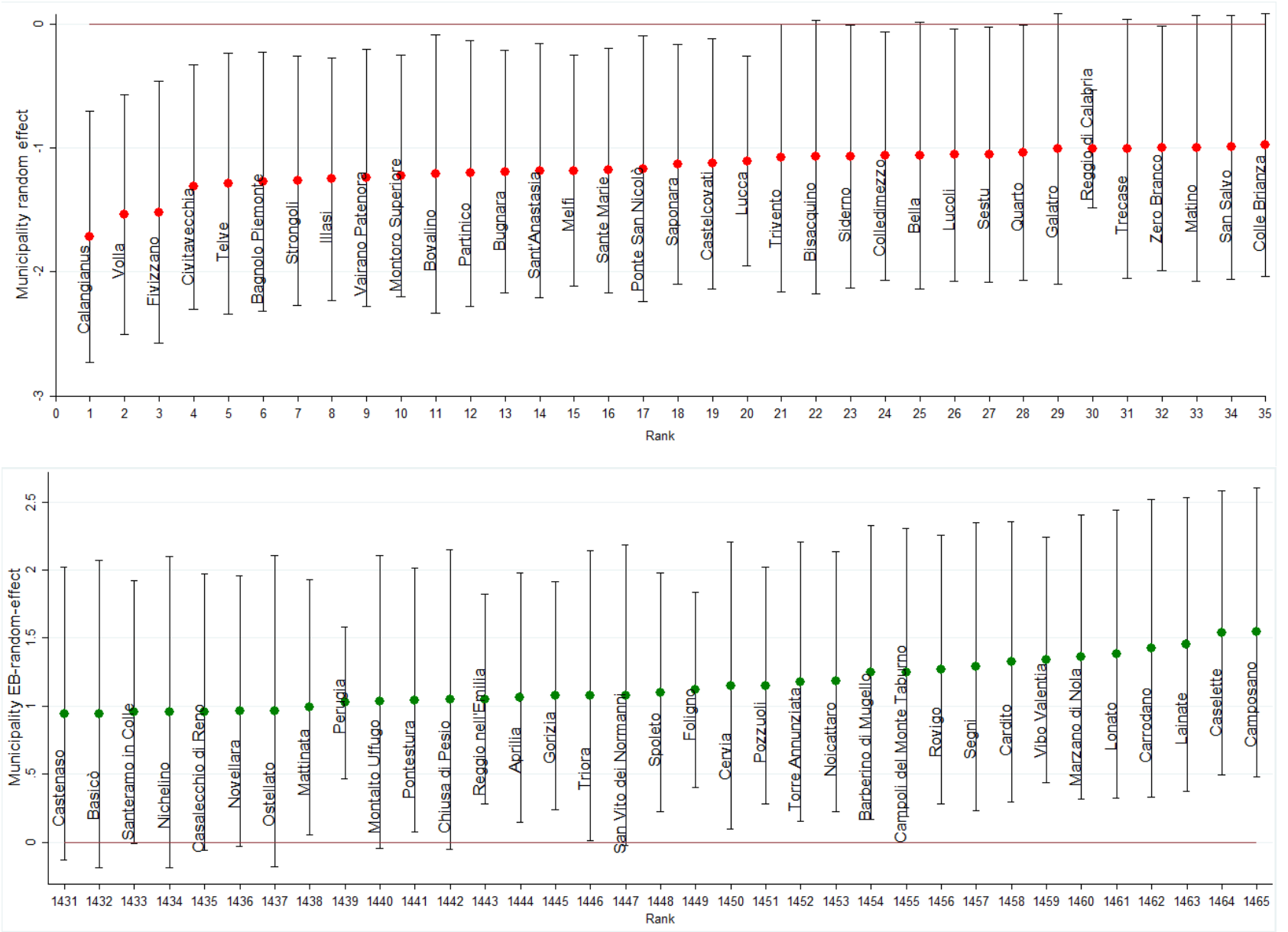


Figure 3.6: Caterpillar graphics: ranking of smaller (top panel) and greatest (bottom panel) 35 municipality unstanardized EB random effects v_k of model 3 and their approximated 95% confidence interval.

Table 3.7: Municipalities with random effects significantly different from zero in model 3.

Geogr.area	Region	Province	Municipality	Urban degree	Rank	Ran.Eff.	Households	Subjects.
Islands	Sardegna	Sassari	Calangianus	Low	1	-	27	79
South	Campania	Napoli	Volla	High	2	-	29	103
Center	Toscana	Massa	Fivizzano	Low	3	-	27	60
Center	Lazio	Roma	Civitavecchia	High	4	-	38	88
North-East	Trentino-Alto Adige	Trento	Telve	Low	5	-	27	67
North-West	Piemonte	Cuneo	Bagnolo Piemonte	Low	6	-	25	66
South	Calabria	Crotone	Strongoli	Low	7	-	32	74
North-East	Veneto	Verona	Illasi	Interm.	8	-	29	84
South	Campania	Caserta	Vairano Patenora	Interm.	9	-	27	76
South	Campania	Avellino	Montoro Superiore	Interm.	10	-	30	107
South	Calabria	Reggio di Calabria	Bovalino	Interm.	11	-	18	52
Islands	Sicilia	Palermo	Partinico	Interm.	12	-	20	54
South	Abruzzo	L'Aquila	Bugnara	Low	13	-	43	97
South	Campania	Napoli	Sant'Anastasia	High	14	-	25	79
South	Basilicata	Potenza	Melfi	Low	15	-	36	111
South	Abruzzo	L'Aquila	Sante Marie	Low	16	-	42	76
North-East	Veneto	Padova	Ponte San Nicolò	High	17	-	25	60
Islands	Sicilia	Messina	Saponara	Interm.	18	-	36	101
North-West	Lombardia	Brescia	Castelcovati	Interm.	19	-	28	70
Center	Toscana	Lucca	Lucca	Interm.	20	-	59	156
South	Calabria	Reggio di Calabria	Siderno	Interm.	23	-	27	65
South	Abruzzo	Chieti	Colledimezzo	Low	24	-	40	86
South	Abruzzo	L'Aquila	Lucoli	Low	26	-	35	78
Islands	Sardegna	Cagliari	Sestu	Interm.	27	-	23	72
South	Campania	Napoli	Quarto	High	28	-	22	73
South	Calabria	Reggio di Calabria	Reggio di Calabria	High	30	-	258	716
North-East	Veneto	Treviso	Zero Branco	Interm.	32	-	33	88
Center	Marche	Pesaro - Urbino	Fano	Interm.	42	-	64	153
North-West	Valle d'Aosta	Aosta	Aosta	Low	131	-	198	424

Continued on next page

– *continued from previous page*

Geogr. area	Region	Province	Municipality	Urban degree	Rank	Ran.Eff.	Households	Subjects
Center	Toscana	Firenze	Firenze	High	1270	+	736	1523
North-East	Emilia-Romagna	Bologna	Bologna	High	1305	+	565	1181
South	Puglia	Bari	Bari	High	1353	+	237	630
North-West	Piemonte	Alessandria	Alessandria	Interm.	1413	+	99	199
South	Molise	Isernia	Isernia	Low	1425	+	59	176
South	Puglia	Foggia	Mattinata	Low	1438	+	39	109
Center	Umbria	Perugia	Perugia	Interm.	1439	+	181	456
North-West	Piemonte	Alessandria	Pontestura	Low	1441	+	34	86
North-East	Emilia-Romagna	Reggio nell'Emilia	Reggio nell'Emilia	High	1443	+	82	195
Center	Lazio	Latina	Aprilia	Interm.	1444	+	47	111
North-East	Friuli-Venezia Giulia	Gorizia	Gorizia	Interm.	1445	+	74	149
North-West	Liguria	Imperia	Triora	Low	1446	+	33	60
Center	Umbria	Perugia	Spoleto	Interm.	1448	+	57	140
Center	Umbria	Perugia	Foligno	Interm.	1449	+	91	254
North-East	Emilia-Romagna	Ravenna	Cervia	Interm.	1450	+	26	71
South	Campania	Napoli	Pozzuoli	High	1451	+	48	141
South	Campania	Napoli	Torre Annunziata	High	1452	+	25	86
South	Puglia	Bari	Noicattaro	High	1453	+	33	99
Center	Toscana	Firenze	Barberino di Mugello	Low	1454	+	21	59
South	Campania	Benevento	Campoli del Monte Taburno	Interm.	1455	+	23	69
North-East	Veneto	Rovigo	Rovigo	Interm.	1456	+	37	92
Center	Lazio	Roma	Segni	Interm.	1457	+	23	62
South	Campania	Napoli	Cardito	High	1458	+	28	71
South	Calabria	Vibo Valentia	Vibo Valentia	Interm.	1459	+	41	131
South	Campania	Avellino	Marzano di Nola	Interm.	1460	+	22	67
North-West	Lombardia	Brescia	Lonato	Interm.	1461	+	25	71
North-West	Liguria	La Spezia	Carrodano	Low	1462	+	35	53
North-West	Lombardia	Milano	Lainate	High	1463	+	24	62
North-West	Piemonte	Torino	Caselette	Interm.	1464	+	26	67
South	Campania	Napoli	Camposano	High	1465	+	22	65

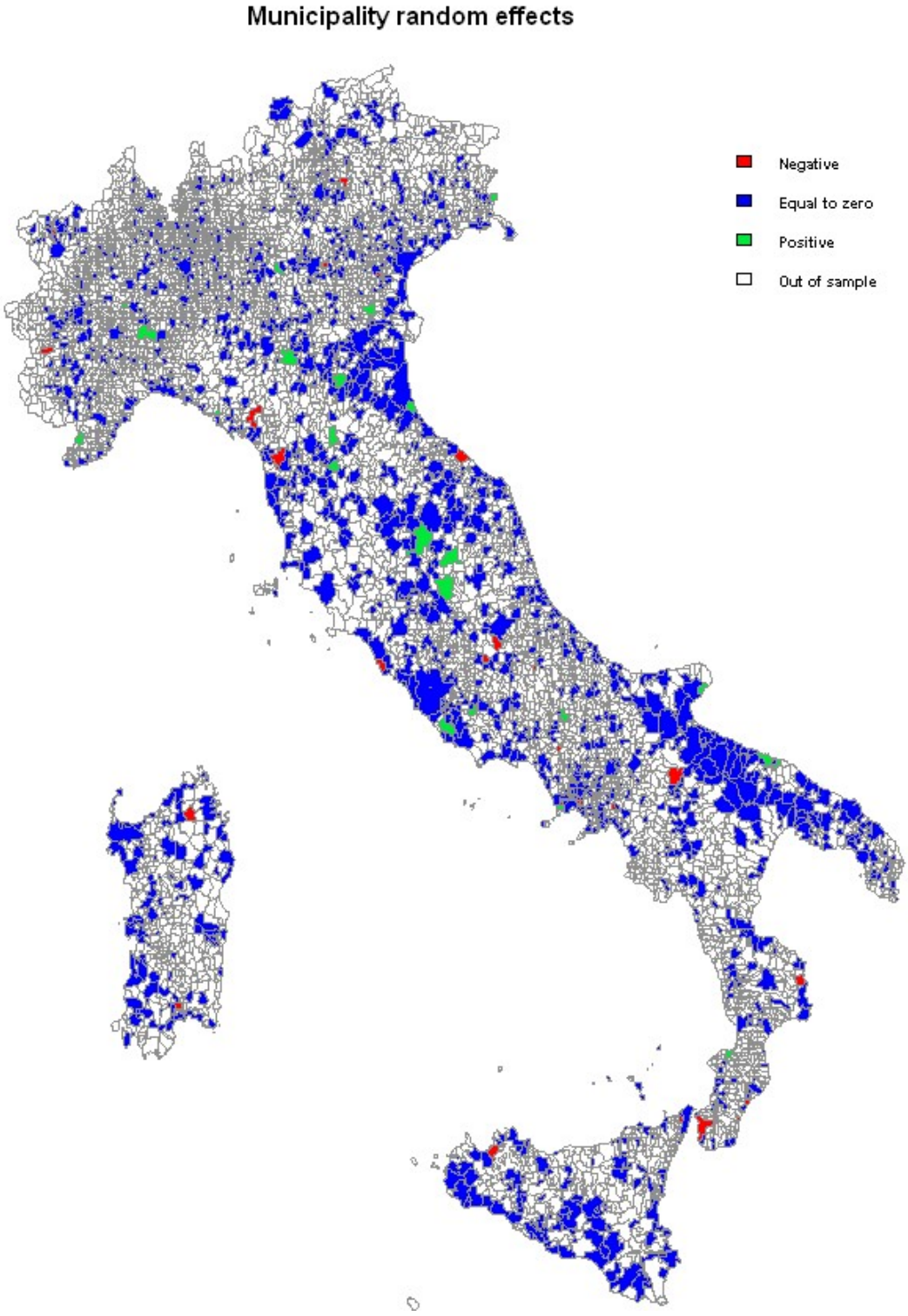


Figure 3.7: Municipality random effects map (model 3)

3.5 Are sampling weights informative?

Large surveys generally use stratified sampling designs. As a result, the individuals in the population are included in the sample with different probabilities. These probabilities are known as "sampling probabilities" or " inclusion probabilities". Such diversity of inclusion probabilities must be taken into account in the data analysis. The reason for using sampling weights in the estimation is to avoid bias. A sampling design with heterogeneous inclusion probabilities will provide a sample that is a biased representation of the population. This fact, however, does not necessarily lead to bias in the estimators for the parameters. If the predictor variables are represented differently in the sample than their distribution in the population (which is one type of bias), not necessarily the estimators of the parameters of the hierarchical model will result to be biased. The parameters are biased only if the distribution of the residuals is affected by the sampling design. In this case the survey design is said to be "informative". If the residuals in the model are independent from sampling design and from sampling weights, then the use of weights is superfluous. In multilevel analysis of survey data, if we can be confident that the model is well specified and the sample design is unrelated to the residuals, it is better not to take into account of the weights when doing the analysis and proceed as usual. In fact, a simulation study carried out by Pfefferman (1998) shows that the scaled weighted estimators perform well, but there is a little disadvantage in terms of bias and precision in using these estimators when sampling is non-informative. So an analysis design-based is to follow only if there are reasons against an analysis model-based.

We describe in section 3.3.2, same useful methods too choose between a model-based and a design-based analysis for a given multilevel data set.

3.5.1 Inclusion probabilities and scaling methods

In a 2-level model, assuming a two-stage sample design, where clusters are selected independently with some probability and, given that a cluster j has been extract, a sample of level-one units within this cluster is selected, in general the inclusion probabilities are defined as follows

π_j= inclusion probability for level-2 unit j
$\pi_{i|j}$= inclusion probability for level-1 unit i, given the inclusion of cluster j.

The marginal probability of level-1 unit i in cluster j is given by the product of these 2 probability

$\pi_j \pi_{i|j}$ = inclusion probability of level-1 unit i of cluster j.

The *design weights* or *base weights* of the sample are the inverse of the inclusion probabilities:

$$w_j = \frac{1}{\pi_j} \tag{3.1}$$

$$w_{i|j} = \frac{1}{\pi_{i|j}} \tag{3.2}$$

To use weights in 2-level models, the separate sets of the weights at level 1 and level 2 are needed, corresponding to the separate inclusion probabilities at each level. Using these weights the *effective sample sizes* can be calculated as follows:

$$N^{eff} = \frac{(\Sigma_j w_j)^2}{\Sigma_j w_j^2} \tag{3.3}$$

$$n_j^{eff} = \frac{(\Sigma_i w_{i|j})^2}{\Sigma_i w_{i|j}^2} \tag{3.4}$$

The effective sample size is defined such that a weighted sample gives the amount of information as a simple random sample with sample size equal to the effective sample size [2].
The ratio of effective sample sizes to actual (real) sample sizes are called the *design effects* [3]

at level 2 and 1 we have respectively:

[2] A general definition of the *effective sample size*, n_{eff}, in a complex survey design of size n (if $n/N \approx 0$) is given by the ratio between n and the design effect: $n_{eff} = \frac{n}{deff}$.
If the sampling design is itself cause of estimate fluctuations, the effective sample size obtained is lower than the actual sample size. Various sampling techniques, being more efficient than sample random sampling, have an effective sample size greater than the the actual sample size

[3] *Design effect* definition: Given an unbiased estimator $\hat{\theta}$ of a parameter θ and a sampling design, is called "design effect" the ratio between variance of $\hat{\theta}$, $V(\hat{\theta})$ and variance $V_o(\hat{\theta})$ of the same estimator in a simple random sampling, both with the same sample size.

$$Deff = \frac{V(\hat{\theta})}{V_o(\hat{\theta})}$$

.

$$deff_2 = \frac{N^{eff}}{N}, \quad deff_{1j} = \frac{n_j^{eff}}{n_j}. \tag{3.5}$$

The design effects give an indication of the potential loss of statistical efficiency incurred by following a design-based rather than model-based analysis. In a level-1 analysis the scales of these weights are irrelevant, while in multilevel designs scaling the level-2 weights still is irrelevant, but the scale of the level-1 weights is important.
At present clear guidelines about the best way of scaling do not exist in the literature. Among the most known methods of scaling we describe two of the most known and used: Pfeffermann's methods. Pfeffermann et al. (1998) propose two methods of scaling to be used in design-based estimation for multilevel models. The first is the "method 1", also called "method B" by Asparouhov (2006), and provides scaled weights summing to effective sample size n_j^{eff},

$$method\,1: \quad w_{i|j}^{*} = \frac{n_j^{eff}}{\Sigma_i w_{i|j}} w_{i|j}. \tag{3.6}$$

The second is the "method 2" ", also called "method A by Asparouhov (2006), and provides scaled weights summing to actual sample size n_j,

$$method\,2: \quad w_{i|j}^{\circ} = \frac{n_j}{\Sigma_i w_{i|j}} w_{i|j}. \tag{3.7}$$

The estimation of these models is performed by a weighted maximum likelihood: "pseudo-maximum-likelihood".

3.5.2 Some procedures to explore the informativeness of survey weights

As before mentioned, if there is association between sampling design and residuals in the hierarchical linear model, then the sampling weights are informative and a model-based approach risks being biased and inconsistent, that is, the parameter estimates may be a misrepresentation of the true parameters in the hierarchical linear model even in large samples. T.Snijders and R. Bosker (2012) propose a variety of methods that can be followed to assess the association of sampling design with the model of interest. Nevertheless, these methods do not have the aim to assert, by an affirmative or negative response, about the informativeness of sampling weights. The proposed methods have rather the purpose to get ideas

about whether and how the design variables and the survey weights may be associated with the model of interest and to guide the researcher in deciding between a model-based and a design-based analysis. Some of these methods are the following:

1. Take a careful look on variability of the weights in each level and on effective sample size. When the variability of the weights is high and/or the design effects is low, it may be a signal that the design is far from a random sample, which can lead to biases of model-based estimators and to low efficiency of design-based estimators.

2. Add design variables to model of interest. This is feasible if the variables that determine the probability of inclusion (the design variables) are known and are part of the data set to be analyzed. However, especially for large surveys, it is common that the information about the sampling design and how the sampling weights is obtained, is only partial. If the design variables are available, then they may be included in the model; it can be tested whether they affect the dependent variable and whether the other results are sensitive to their inclusion. This makes sense especially as additions to a model that is already rich in explanatory variables. The only objective in this exploration is to assess whether the design variables have effects in addition to the variables already in the model of interest.

3. Apply the hierarchical linear model to parts of the data differing with respect to inclusion probabilities or design variables. The procedure consists to split the data into parts according to the weights at each level and analyze each part separately by multilevel model according to model of interest, and assess the magnitude of the differences in results between the various parts.

4. Add weight variables to the model. Non-informativeness of the design means that the design variables are independent of the residuals. This can be tested by using the weights as covariate added to the hierarchical linear model.

5. Compare model-based and design-based estimators.

3.6 Scaled weights computation

In order to compare design-based and model-based estimates we consider two models that include scaled sample weights respectively by the Pfeffermann method 1 and method 2 and compare them with the unweighted model. For this purpose will be necessary to derive the conditional weights at each hierarchical level.
This type of analysis, was made indeed in some intermediate steps during the specification of the models, and gave us a first idea about weighted informativeness.

In the Istat household surveys inclusion probability of the j-th selected household in the k-th municipality of the h-th stratum is defined as (PD Falorsi and S. Falorsi, 1995):

$$\pi_j = \frac{n_h P_{hk}}{P_h} \frac{m_{hk}}{M_{hk}} \tag{3.8}$$

where as previously:
i denote level-1 unit (subject)
j denore level-2 unit (household)
k denote level-3 unit (municipality)
h denote stratum

and

P_h is the number of residents in stratum h
P_{hk} is the number of residents in the municipality k of stratum h
M_{hk} is the number of households in the municipality k of stratum h
m_{hk} is the number of drown households in the municipality k of stratum h
n_h is the number of drown municipalities in stratum h.

Reciprocal of π_j is the so-called *base weight*.

The available *final weights*, are not the simple reciprocal of the inclusion probability and therefore does not correspond directly to the weights [4] .
They are multiplied by factors correcting for non-response and for domains of interest (i.e. corrected by a factor that allows to satisfy the equality condition between known population totals and the corresponding sample estimates for each domain of interest (territorial areas) (see Istat, year)). We write synthetically the data-set final weights as

$$w_{ij}^{'} = w_{ij} \cdot f_{ijD}$$

in which f_{ijD} indicates all adjustment factors applied to base weights (D denote the domain).

Adjustment factors are not available and therefore cannot be removed. For coherence with level 1 weights, we will derive level-2 conditional probabilities in

[4]E. L. Korn and B. I. Graubard (2003) suggest that the base weights should be used, because the adjustments introduced to compensate for non-response and for domain of interest (post-stratification), lead to a lower efficiency estimators (variance of the estimators increases). Moreover they proposed another weighting method, requiring knowledge of higher-order inclusion probabilities which often is not available; therefore it is less widely applied. However, in their simulations their method does appear to perform well.

way to contain the correction factors at each level. Now, the inclusion probabilities of the individual are equal to the inclusion probability of the corresponding household because all the household members were interviewed (cluster sampling of households within municipalities). Therefore we have

$$\pi_{ij} = \pi_j = \pi_{i|j}$$

And so it is not necessary to scale level-1 weights, but only the level-2 ones. For the reasons mentioned before we will obtain the conditional inclusion probability at level 2 as follows: rewriting the inclusion probability of household j as

$$\pi_j^{'} = \pi_j \cdot 1/f_{ijD}$$

being the inclusion probability of municipality k belonging to the h-th stratum given by:

$$\pi_{k|h} = \frac{n_h P_{hk}}{P_h}$$

we will obtain the conditional probability of household j-th given the municipality k in this way:

$$\pi_{j|kh}^{'} = \frac{\pi_j^{'}}{\pi_{k|h}}$$

instead as m_{hk}/M_{hk}.

The computational intervention is restricted only to NAR (not self-representing) municipalities because the AR (self-representing) municipalities form stratum itself and enter in the sample with certainty.
By distinguishing the two cases we have:

- if $k \in AR$ is $\pi_{h|k} = 1$
  and than the data-set weights $w_j^{'}$ remain as they are;

- while if $k \in NAR$ the conditional inclusion probability is:

$$w_{j|kh}^{'} = \frac{n_h P_{hk}}{P_h} w_j^{'}$$

At this point it would be sufficient multiply survey weights by the probability to draw the municipality in the stratum, but in the absence of information on stratum population (missing information on the list of municipalities by stratum) for the NAR municipalities we have replaced stratum population, P_h, with Large-Area population, P_A, and number of municipalities extracted in the stratum, n_h, with number of municipalities extracted in Large-Area, n_A. We retain that this

approximation does not involve big changes [5]. So replacing:
if $k \in NAR$ the conditioned weight is calculated as follows:

$$w_{j|kA}^{'} = \frac{n_A P_{Ak}}{P_A} w_j^{'}$$

obtaining the conditioned weights to be scaled.

For these computations the resident population at January 1st 2005 has been used.
To $w_{j|kA}^{'}$ obtained we applied Pfefferman scaling method 1 and 2, obtaining the second level scaled weights to be used for the estimation of corresponding weighed models.

$$method\,1: \quad w_{j|kA}^{*} = \frac{n_{kA}^{eff}}{\Sigma_j w_{j|kA}^{'}} w_{j|kA}^{'}. \tag{3.9}$$

$$method\,2: \quad w_{j|kA}^{\circ} = \frac{n_{kA}}{\Sigma_j w_{j|kA}^{'}} w_{j|kA}^{'}. \tag{3.10}$$

.

3.7 Design-based and Model-based analysis

In order to explore the informativeness of survey weights, the procedures 1, 2, 4 and 5 of section 3.3.2 have been applied. In more detail:

1. Computation of sampling weights variance in the survey. It is very high and equal respectively to 77907.97 for the individual weights and 78008.49 for the household weights. The design effect is equal to $deff_2$=0.7251 and the effective sample size is equal to N_{eff}=92837.18 [6]. These results suggest the application of at least some of the described methods in order to understand whether and how much they can have an impact on estimates;

2. Estimation of model having variable that identifies primary sample units (PSUs) as covariate.

3. Estimation of model containing sampling weights as a covariate;

[5] NAR municipalities are been extracted with probability proportional to population size within each stratum; 4 municipalities per stratum were extracted.

[6] These values refer to the entire sample.

4. Estimation of two weighed models using level-2 scaled weights according to Pfeffermann methods.

The comparison of these models will be done with the model 2 which has been re-performed excluding the heteroscedasticity modeling, without loss of significance in comparison.

The results of our estimated models are illustrated in Table 3.7 and 3.8.

Table 3.8: Compare among model 2, model including design variable (D.V.) and model including sample weights (S.W.) as covariates (PSUs stands for Primary Sample Units).

MODEL	Unweighted		with D.V.		with S.W.	
Fixed effects	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Intercept	53.44	0.19	53.55	0.21	53.47	0.19
Age	0.040	0.005	0.039	0.005	0.039	0.005
Age^ 2	−0.002	0.000	−0.002	0.000	−0.002	0.000
Female	−0.11**	0.05	−0.11**	0.05	−0.11**	0.05
Qualification: ref. No Educ. Qualif.						
University	1.29	0.14	1.28	0.14	1.28	0.14
Graduate	1.16	0.12	1.16	0.12	1.16	0.12
Secondary	0.86	0.11	0.86	0.11	0.86	0.11
Elementary	0.52	0.12	0.52	0.12	0.52	0.12
Employment: ref. Employed						
Unemployed	−0.27*	0.09	−0.27*	0.09	−0.27*	0.09
Not working	−0.44	0.05	−0.44	0.05	−0.44	0.05
Unable to work	−4.37	0.24	−4.37	0.24	−4.37	0.24
Body-mass index	−0.43	0.03	−0.43	0.03	−0.43	0.03
Sport	0.63	0.04	0.63	0.04	0.63	0.04
Disabled	−8.88	0.27	−8.88	0.27	−8.88	0.27
Disease: ref. No health problem						
Illness	−1.47	0.09	−1.47	0.09	−1.47	0.09
Acute disease	−6.05	0.11	−6.06	0.11	−6.05	0.11
Interaction: Disabled x Disease						
Disabled-Illness	0.77*	0.29	0.77*	0.29	0.77*	0.29
Disabled-Acute disease	3.64	0.29	3.64	0.29	3.64	0.29
Well-being (VT)	1.12	0.02	1.12	0.02	1.12	0.02
Infarct	−2.57	0.19	−2.57	0.19	−2.56	0.19
Heart disease	−2.31	0.12	−2.31	0.12	−2.31	0.12
Stroke	−2.03	0.20	−2.03	0.20	−2.03	0.20
Arthrosis,Arthrit.	−3.86	0.08	−3.86	0.08	−3.86	0.08
Osteoporosis	−1.89	0.11	−1.89	0.11	−1.88	0.11
Hepatic cirrhosis	−1.84	0.46	−1.83	0.46	−1.84	0.46

Continued on next page

– *continued from previous page*

MODEL	Unweighted		with D.V.		with S.W.	
Fixed effects	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Cancer	−3.45	0.27	−3.45	0.27	−3.45	0.27
Parkinsonism	−1.59	0.34	−1.59	0.34	−1.59	0.34
Ulcer	−0.63	0.17	−0.63	0.17	−0.63	0.17
Cancer in the past	−1.22	0.19	−1.22	0.19	−1.22	0.19
Alzheimer	−0.71**	0.32	−0.71**	0.32	−0.71**	0.32
Stones	−1.06	0.13	−1.06	0.13	−1.06	0.13
Thyroid disease	−0.55	0.12	−0.55	0.12	−0.55	0.12
Asthma	−1.20	0.12	−1.20	0.12	−1.20	0.12
Diabetes	−1.28	0.12	−1.28	0.12	−1.28	0.12
Hypertension	−0.40	0.07	−0.40	0.07	−0.40	0.07
Bronchitis,Emph.	−0.74	0.10	−0.74	0.10	−0.74	0.10
Household composition: ref. 1 component						
2 components	−0.36	0.07	−0.36	0.07	−0.36	0.07
3 components	−0.46	0.07	−0.46	0.07	−0.45	0.07
4 components	−0.36	0.07	−0.36	0.07	−0.36	0.07
5 or more comp.	−0.41	0.11	−0.41	0.11	−0.41	0.11
Insuf. econ. resour.	−0.21	0.06	−0.21	0.06	−0.21	0.06
Urban degree: ref. Low						
Intermediate	0.27	0.07	0.27	0.07	0.29	0.07
High	0.39	0.09	0.37	0.10	0.42	0.09
PSU			0.00***	0.00		
Sample weight					0.00***	0.00
Random effects	Param.	S.E.	Param.	S.E.	Param.	S.E.
3-Municipality: sd(v)	0.83	0.04	0.82	0.04	0.83	0.04
2-Household: sd(u)	2.12	0.04	2.12	0.04	2.12	0.04
1-Residual: sd(e)	5.70	0.05	5.70	0.05	5.70	0.05

*Coefficient not significant: *** "p¿0.05"; ** "p¿0.01"; * "p¿0.001"*

Table 3.9: Compare between unweighted model 2 and weighted models including weights scaled by Pfeffermann's method 1 and method 2.

MODEL	Unweighted		Method 1		Method 2	
Fixed effects	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Intercept	53.44	0.19	53.45	0.22	53.48	0.23
Age	0.040	0.005	0.035	0.005	0.035	0.006
Age^ 2	−0.002	0.000	−0.002	0.000	−0.002	0.000
Female	−0.11**	0.05	−0.09***	0.06	−0.08***	0.05
Qualification: ref. No Educ. Qualif.						
University	1.29	0.14	1.47	0.15	1.49	0.15
Graduate	1.16	0.12	1.36	0.13	1.37	0.13
Secondary	0.86	0.11	0.98	0.12	1.01	0.13
Elementary	0.52	0.12	0.68	0.14	0.71	0.14
Employment: ref. Employed						
Unemployed	−0.27*	0.09	−0.28**	0.12	−0.27**	0.12
Not working	−0.44	0.05	−0.46	0.07	−0.47	0.07
Unable to work	−4.37	0.24	−4.46	0.27	−4.35	0.29
Body-mass index	−0.43	0.03	−0.43	0.03	−0.43	0.04
Sport	0.63	0.04	0.55	0.05	0.53	0.05
Disabled	−8.88	0.27	−8.74	0.30	−8.74	0.30
Disease: ref. No health problem						
Illness	−1.47	0.09	−1.47	0.12	−1.46	0.11
Acute disease	−6.05	0.11	−5.95	0.11	−6.00	0.12
Interaction: Disabled x Disease						
Disabled-Illness	0.77*	0.29	0.85**	0.34	0.84**	0.37
Disabled-Acute disease	3.64	0.29	3.73	0.34	3.76	0.34
Well-being (VT)	1.12	0.02	1.12	0.02	1.13	0.02
Infarct	−2.57	0.19	−2.51	0.21	−2.52	0.21
Heart disease	−2.31	0.12	−2.27	0.15	−2.30	0.15
Stroke	−2.03	0.20	−2.09	0.26	−2.14	0.27
Arthrosis,Arthrit.	−3.86	0.08	−3.93	0.10	−3.92	0.10
Osteoporosis	−1.89	0.11	−1.88	0.14	−1.86	0.14
Hepatic cirrhosis	−1.84	0.46	−2.17	0.50	−2.12	0.49
Cancer	−3.45	0.27	−3.22	0.32	−3.12	0.32

Continued on next page

– *continued from previous page*

MODEL	Unweighted		Method 1		Method 2	
Fixed effects	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Parkinsonism	−1.59	0.34	−1.49	0.41	−1.47	0.40
Ulcer	−0.63	0.17	−0.41**	0.19	−0.45**	0.19
Cancer in the past	−1.22	0.19	−1.24	0.22	−1.21	0.25
Alzheimer	−0.71**	0.32	−0.77***	0.40	−0.74**	0.36
Stones	−1.06	0.13	−1.06	0.16	−1.11	0.16
Thyroid disease	−0.55	0.12	−0.61	0.16	−0.63	0.16
Asthma	−1.20	0.12	−1.10	0.12	−1.12	0.12
Diabetes	−1.28	0.12	−1.40	0.14	−1.38	0.14
Hypertension	−0.40	0.07	−0.39	0.09	−0.40	0.09
Bronchitis,Emph.	−0.74	0.10	−0.83	0.14	−0.83	0.14
Household composition: ref. 1 component						
2 components	−0.36	0.07	−0.46	0.09	−0.49	0.09
3 components	−0.46	0.07	−0.52	0.10	−0.56	0.10
4 components	−0.36	0.07	−0.41	0.09	−0.45	0.09
5 or more comp.	−0.41	0.11	−0.47	0.13	−0.50	0.13
Insuf. econ. resour.	−0.21	0.06	−0.21*	0.07	−0.20*	0.07
Urban degree: ref. Low						
Intermediate	0.27	0.07	0.22*	0.08	0.22*	0.08
High	0.39	0.09	0.32*	0.12	0.33*	0.12
Random effects	Param.	S.E.	Param.	S.E.	Param.	S.E.
3-Municipality: sd(v)	0.83	0.04	0.82	0.04	0.99	0.04
2-Household: sd(u)	2.12	0.04	2.12	0.05	2.04	0.05
1-Residual: sd(e)	5.70	0.05	5.72	0.05	5.73	0.05

*Coefficient not significant: *** "p¿0.05"; ** "p¿0.01"; * "p¿0.001"*

Models including PSUs and the one including sample weights as covariates, do not show any effect either on parameter estimates (in both fixed part and random part) neither on standard errors.

The use of weighted estimators involve small enlargement in some coefficient standard errors and some changes on coefficients, mainly in coefficient related to diseases or other category having low frequencies in the sample. However, these changes do not are generally capable of affecting the validity of the results achieved with model based estimate.

We conclude that model-based analysis can be performed; sample weights are not informative for model under analysis.

Chapter 4

Multilevel linear quantile regression

4.1 Introduction

Given the pronounced asymmetry of response variable and not normal distribution of residuals we search for more robust estimation methods. Quantile regression has appeared a good choice for both robustness properties and possibility of more extensive analysis of the response variable. The main features of this type of regression are:

- complete mean regression with a complete picture of distributional effects;
- robustness against outliers and asymmetry;
- be distribution-free;
- naturally handles of heteroscedasticity;
- be equivariant to monotone transformations.

We are induce to search for a quantile regression method applicable in the context of hierarchical data for the following reasons:

1. the errors distribution is not well approximated by a normal;
2. the response variable is skewed on the left and the mean is not a good representation of the PCS central tendency;
3. the presence of heteroscedasticity in the covariate conditional distributions.

In the latest years, the need for extend the capabilities of quantile regression for independent data to deal with hierarchical data structure has led to distinct approaches. We focused on the one proposed by Geraci and Bottai (M. Geraci and M. Bottai 2007, 2011) based on asymmetric Laplace density (ALD).
The aim of this part of the work will be to apply their method using the recent procedure developed by M. Geraci in R: the `lqmm` package (version 1.02). Accordingly the results showed in this chapter were performed using software R.
In the following paragraphs we will describe "Quantile Regression" (QR) synthetically, and then we will illustrate the "Linear Quantile Mixed Model" (LQMM) approach developed by Geraci and Bottai.

4.2 Quantile regression

Quantile regression is a way to estimate the conditional quantiles of a response variable distribution in the linear model that provides a more complete view of possible relationships between variables. Quantile regression, as introduced by Koenker and Basset (1978), may be viewed as an extension of least squared estimation of conditional mean models to the estimation of an ensemble of models for several conditional quantile functions. Indeed, in analogy with classical linear regression methods, based on minimizing sums of squared residuals and intending to estimate models for conditional mean functions, quantile regression methods are based on minimizing asymmetrically weighted absolute residuals and intending to estimate conditional median functions (a special case) and conditional quantile functions for several others quantiles.
If we consider data in the form $(\mathbf{x}_i^T, y_i)$, for i=1,...,N, where y_i are independent observations of a continuous random variable having distribution function F not known, and $\mathbf{x}_i^T$ are row p-vector of a design matrix X, we can define the τ-th quantile regression coefficients as any solution $\hat{\beta}_\tau$, $\hat{\beta}_\tau \in \Re^p$, of the following minimization problem

$$\hat{\beta}_\tau = min_{\beta \in \Re^p} \Big[\sum_{i \in (i: y_i \geq \mathbf{x}_i^T \beta)} \tau |y_i - \mathbf{x}_i^T \beta| \quad + \sum_{i \in (i: y_i < \mathbf{x}_i^T \beta)} (1-\tau) |y_i - \mathbf{x}_i^T \beta| \quad \Big] \qquad (4.1)$$

where $0 < \tau < 1$ indicates the quantile.
For $\tau = 1/2$ we obtain the median regression and the above equation simplifies to

$$\hat{\beta}_{0.5} = min_{\beta \in \Re^p} \sum_i |y_i - \mathbf{x}_i^T \beta|$$

All observations greater than the absolute differences between the observations and the unknown solution are weighted with τ , all observations below the solution are

weighted with $(1-\tau)$. QR requires that the τ-th quantile of the error term be zero. Note that each quantile estimate is τ dependent. The resulting minimization problem is a linear function of the parameters and can be solved by linear programming methods (simplex algorithm). In general, to obtain the parameter standard errors the bootstrap method is used. Another common way to write the above minimization problem is to use the "check function" ρ_τ defined as

$$\rho_\tau(v) = \tau v I_{[0,\infty)}(v) - (1-\tau) v I_{(-\infty,0)}(v) \tag{4.2}$$

where $I(.)$ is the indicator function. The check function allows us to reformulate (4.1) in the following way

$$\hat{\beta}_\tau = min_{\beta \in \Re^p} \sum_i \rho_\tau(y_i - \mathbf{x}_i^T \beta) \tag{4.3}$$

For each covariate the estimated coefficient may be interpreted as the impact of a one unit change of the covariate on response variable living unchanged the others. In their paper Koenker and Basset (1978) also established consistency and asymptotic normality of those estimators for fixed x_i.

4.3 Linear Quantile Mixed Models

The multilevel linear models, likewise to linear regression, estimate the conditional expectation of a response variable taking into account the hierarchical data structure, but they are not used to characterize the entire conditional distribution of a dependent variable. Quantile regression models allow it but cannot deal hierarchical data. Geraci and Bottai (M. Geraci and M. Bottai, 2007, 2011) have introduced a new method for quantile regression with mixed effects, the "Linear Quantile Mixed Model" (LQMM) that, in this context, we call "Multilevel Linear Quantile Regression" (MLQR). They propose a conditional quantile regression model for continuous responses where random effects are inserted along with fixed coefficient predictors to take into account the dependence between units in the context of multilevel data analysis. The approach is based on the link existing between the minimization of weighted least absolute deviations of quantile regression (4.1) and the maximization of a Laplace likelihood. Yu, Lu, and Stander in a work of 2001 presented alternative definition of quantile regression carried out by using the Asymmetric Laplace Distribution (ALD) function. This distribution also appeared in paper by Koenker and Machado (1999) on goodness of fit for quantile regression.
We say that a random variable Y is distributed as an ALD with parameters μ,

σ and τ and we write it as $Y \sim ALD(\mu, \sigma, \tau)$, if the corresponding probability density is given by

$$f(y|\mu, \sigma, \tau) = \frac{\tau(1-\tau)}{\sigma} exp\Big\{ -\rho_\tau\Big(\frac{y-\mu}{\sigma}\Big)\Big\}, \tag{4.4}$$

where $\rho_\tau(v)$ is the check function (also called loss function) defined in (4.2), $0 < \tau < 1$ is the skewness parameter, $\sigma > 0$ is the scale parameter, and $-\infty < \mu < +\infty$ is the location parameter. The support of the random variable Y is the real line. Note that the loss function ρ assigns weight τ or $1-\tau$ to the observations greater or, respectively, less than μ and that $P(y \leq \mu) = \tau$. Therefore, the distribution splits along the scale parameter into two parts, one with probability τ to the left, and one with probability $1-\tau$ to the right. Denoting $\mu_i = \mathbf{x}_i^T\beta$ and $\mathbf{y} = (y_1, \ldots, y_N)$, and assuming that $y_i \sim ALD(\mu_i, \sigma, \tau)$, the likelihood for N independent observations is, bar a proportionality constant,

$$L(\beta, \sigma; \mathbf{y}, \tau) \propto \sigma^{-1} exp\Big\{ -\sum_{i=1}^{N} \rho_\tau\Big(\frac{y_i - \mu_i}{\sigma}\Big)\Big\}. \tag{4.5}$$

If we consider σ a nuisance parameter, then the maximization of the likelihood in (4.5) with respect to the parameter β is equivalent to the minimization of Koenker function in (4.3). Therefore, ALD proves a useful link between the likelihood and the inference for QR estimation. Now the extension to hierarchical case data involves the inclusion of the random effects.
Currently `lqmm` package has been developed only for 2-level case, and requires:

- y continuous response variable;
- to fix quantile of interest τ (e.g., τ=0.25);
- $y = X\beta + Zu + \varepsilon$ linearity relation between dependent and independent variables ; u denote random effects;
- $\varepsilon \sim ALD(0, \sigma, \tau)$ so we target the quantile τ;
- $u \perp \varepsilon$ independence between level-1 and level-2 residuals ;

Random effects covariance matrix Ψ can be diagonal or having a symmetric structure. We assume that y_{ij}, conditionally on u_j, for $i = 1, \ldots, n_j$ and $j = 1, \ldots, J$ are independently distributed according to an ALD

$$f(y_{ij}|\beta, u_j, \sigma) = \frac{\tau(1-\tau)}{\sigma} exp\Big\{ -\rho_\tau\Big(\frac{y_{ij} - \mu_{ij}}{\sigma}\Big)\Big\} \tag{4.6}$$

where $\mu_{ij} = \mathbf{x}_{ij}^T\beta + u_j$ is the linear predictor of the τ-th quantile. The conditional quantile to be estimated, τ, is fixed and known. The random effects induce a correlation structure among observations on the same group. Assume that u_j are identically distributed according to the same density f_u characterized by a τ-dependent parameter ψ (i.e. $\psi(\tau)$), and they are mutually independent. Assume that e_{ij} are independent, and u_j and e_{ij} are independent of one another.

Let $\mathbf{y}_j = (y_{j1}, \ldots, y_{jn_j})$ and $f(\mathbf{y}_j|\beta, u_j, \sigma) = \prod_{i=1}^{n_j} f(y_{ij}|\beta, u_j, \sigma)$ be the density for the j-th group conditional on the random intercept u_j.The complete-data density of $(\mathbf{y}_j, u_j)$, per $j = 1, \ldots, J$ is given by

$$f(\mathbf{y}_j, \mathbf{u}_j|\beta, \sigma, \psi) = f(\mathbf{y}_j|\beta, u_j, \sigma)f(u_j|\psi) \tag{4.7}$$

where $f(u_j|\psi)$ is the density of u_j and β, σ, ψ are the parameters of interest. If we let $\mathbf{y} = (\mathbf{y}_1, \ldots, \mathbf{y}_J)$ and $\mathbf{u} = (u_1, \ldots, u_J)$ the joint density of $(\mathbf{y}, \mathbf{u})$ based on J groups is given by:

$$f(\mathbf{y}, \mathbf{u}|\beta, \sigma, \psi) = \prod_{j=1}^{J} f(\mathbf{y}_j|\beta, u_j, \sigma)f(u_j|\psi) \tag{4.8}$$

We obtain the marginal density of $\mathbf{y}$ by integrating out the random effects, leading to $f(\mathbf{y}|\beta, \sigma, \psi) = \int_{R^J} f(\mathbf{y}, \mathbf{u}|\beta, \sigma, \psi)d\mathbf{u}$. Thus, the inference about the parameters β, σ, ψ should be based on the marginal likelihood, $L(\beta, \sigma, \psi; \mathbf{y}) = \sum_j^J L(\beta, \sigma, \psi; \mathbf{y}_j)$. The latter involves an integral that, in general, does not have a closed form solution. For estimations `lqmm` procedure provides numerical integration of the likelihood $L_j(\beta, \sigma, \psi \mid y_j) = \int f(y_j, u_j \mid \beta, \sigma, \psi)du_j$ via:

- Gauss-Hermite quadrature if u is assumed normally distributed

- Gauss-Laguerre quadrature if u is assumed ALD distributed

We can denote the marginal log-likelihood with $l_j(\beta, \sigma, \psi \mid y_j) = logL_j(\beta, \sigma, \psi \mid y_j)$. For fixed effects β and covariance matrix Ψ estimation is provided by one of the two following optimization methods:

- derivative-free optimization (with Nelder-Mead algorithm)

- subgradient optimization (gradient search procedure)

Parameters interpretation is the same as in QR: $\hat{\beta}_\tau$ is the estimated coefficient vector at the τ-th quantile and each element of this vector (coefficient) expresses the marginal change in the τ-th quantile on the response variable due to a 1 unit change in the associated covariate, living the others unchanged. If we are in the

contest of a random intercept model $\hat{\psi}_u^2$ will be a measure of the dispersion of the cluster-specific random intercept at the τ-th quantile. The estimators are asymptotically normal, therefore a Wald-test can be used for testing hypothesis. In the same way that QR, coefficients standard errors are calculated via bootstrap. Variance components can be tested by LR-tests.
The scale parameter σ of the ALD does not have a straightforward interpretation since the use of the Laplace distribution for the conditional response responds to the need for a likelihood approach to quantile regression rather than for the observation to be effectively Laplacian. Hence for assessing fit of the model, AIC and log-likelihood values can be examined (provided in the `lqmm` output). This application does not provide standard errors for random effects standard deviations.

4.4 Multilevel linear quantile regression estimations

For an experimental application of the procedure `lqmm` we will use the final models 2 (having *household composition* as covariate) and 3 (having *household typology as covariate*), but recalculated excluding third hierarchical level and residuals heteroscedasticity at level 1. These changes on the final models are necessary because this recent application is currently developed for models with only two levels, the second change because it does not involve large differences in the estimated coefficients, therefore, priority was given to exposure and computational motivations. The choice of estimating a multilevel linear quantile regression for two very similar models will allow us to assess for stability of estimate procedure even for little variations.
The application is carried out by assuming random effects normally distributed, thus the estimates will be calculated by numerical integration using the method of Gauss-Hermite quadrature, and as optimization algorithm the Nelder-Mead has chosen (which is set by default in the procedure). The sample size does not allow high number of bootstrap replications, so the calculation of standard errors have been performed with 100 bootstrap replications that are considered nevertheless as sufficient.
Seven quantile estimation are fixed for each model: 0.10, 0.25, 0.33, 0.50, 0.67, 0.75 and 0.90. Quantile 0.33 was added because it corresponds approximately to the sample average of PCS ($\tau_{0.324} = 49.89$); quantile 0.67 be chosen for symmetry. Table 4.1 reports unconditional quantile values of response variable.

Table 4.1: Sample quantiles of PCS index for 18 years and older.

Quantile	0.05	0.10	0.25	0.33	0.50	0.67	0.75	0.90	0.95
PCS score	27.92	33.89	46.03	50.10	54.32	55.91	56.15	57.34	58.43

The Tables below report the obtained estimates for the two final models. For model 3 the standard errors are calculated only for central quantiles. The number of degrees of freedom are: 45 in model 2 and 44 in model 3.

Table 4.2: Multilevel linear quantile regression: estimated parameters and standard errors for model 2 and model 3: quantile 0.10 (standard errors for model 3 not performed).

MODEL	Model 2		Model 3
Quantile	Quantile 0.10		Quantile 0.10
Fixed effects	Coef.	S.E.	Coef.
Intercept	47.54	0.94	46.37
Age	0.003***	0.018	0.020
Age^ 2	−0.001*	0.000	−0.001
Female	−0.01***	0.08	0.07
Qualification: ref. No Educ. Qualif.			
University	2.51	0.56	2.88
Upper Secondary	2.19	0.52	2.48
Lower Secondary	1.96	0.52	2.20
Primary	1.51*	0.45	1.68
Employment: ref. Employed			
Unemployed	−0.54*	0.16	−0.48
Not working	−0.71	0.13	−0.72
Unable to work	−5.45	0.65	−4.90
Body-mass Index	−0.29	0.07	−0.33
Sport	0.47	0.09	0.42
Disabled	−9.65	0.66	−8.89
Disease: ref. No health problem			
Illness	−1.95	0.13	−1.76
Acute disease	−7.49	0.22	−7.46
Interaction: Disabled x Disease			
Disabled-Illness	2.93	0.78	1.77
Disabled-Acute disease	7.10	0.77	6.49
Well-being (VT)	1.47	0.03	1.53
Infarct	−2.24	0.41	−1.75
Heart disease	−2.26	0.31	−2.01
Stroke	−1.12***	0.60	−0.07
Arthrosis,Arthrit.	−4.85	0.20	−4.93
Osteoporosis	−1.54	0.25	−1.41
Hepatic cirrhosis	−2.41***	1.36	−2.72

Continued on next page

– *continued from previous page*

MODEL	Model 2		Model 3
Quantile	Quantile 0.10		Quantile 0.10
Fixed effects	Coef.	S.E.	Coef.
Cancer	−3.83	0.66	−3.57
Parkinsonism	−2.22**	0.86	−1.44
Ulcer	−0.15***	0.30	−0.14
Cancer in the past	−0.97**	0.45	−0.65
Alzheimer	−0.66***	0.83	1.00
Stones	−0.73**	0.34	−0.69
Thyroid disease	−0.62*	0.23	−0.64
Asthma	−0.96	0.25	−0.38
Diabetes	−1.31	0.31	−0.75
Hypertension	−0.57	0.14	−0.63
Bronchitis,Emph.	−0.29***	0.26	−0.54
Household composition: ref. 1 component			
2 components	0.30***	0.32	
3 components	−0.21***	0.36	
4 components	−0.03***	0.39	
5 or more comp.	−0.05***	0.41	
Household typology: ref. Living alone			
Childless couple			0.09
Couple with child.			−0.35
Other typology			−0.25
Insuf. econ. resources	−0.31***	0.23	0.16
Urban degree: ref. Low			
Intermediate	0.29***	0.20	0.59
High	0.57**	0.21	0.91
Random effects	Paramet.	S.E.	Paramet.
2-Household: sd(u)	5.18		5.29
ALD scale (sd)	0.65(6.50)	0.00	0.65(6.50)
Log-likelihood	-355704.8		-355936.9
AIC	711499.5		711961.7

*Coefficient not significant: *** "p¿0.05"; ** "p¿0.01"; * "p¿0.001"*

Table 4.3: Multilevel linear quantile regression: estimated parameters and standard errors for model 2 and model 3: quantile 0.25.

MODEL	Model 2		Model 3	
Quantile	Quantile 0.25		Quantile 0.25	
Fixed effects	Coef.	S.E.	Coef.	S.E.
Intercept	50.66	0.45	51.44	0.56
Age	0.030**	0.013	0.025***	0.016
Age^ 2	−0.002	0.000	−0.002	0.000
Female	−0.09***	0.05	−0.13*	0.05
Qualification: ref. No Educ. Qualif.				
University	1.68	0.30	1.54	0.33
Upper Secondary	1.49	0.29	1.35	0.33
Lower Secondary	1.24	0.28	1.14	0.31
Primary	0.84*	0.25	0.83*	0.28
Employment: ref. Employed				
Unemployed	−0.61	0.11	−0.61	0.12
Not workinging	−0.61	0.08	−0.61	0.08
Unable to work	−4.84	0.34	−4.87	0.42
Body-mass Index	−0.36	0.04	−0.35	0.05
Sport	0.54	0.06	0.47	0.06
Disabled	−10.24	0.34	−10.06	0.41
Disease: ref. No health problem				
Illness	−1.82	0.08	−1.85	0.08
Acute disease	−7.70	0.10	−7.70	0.13
Interaction: Disabled x Disease				
Disabled-Illness	1.95	0.44	2.02	0.42
Disabled-Acute disease	6.72	0.36	6.56	0.43
Well-being (VT)	1.31	0.02	1.30	0.02
Infarct	−2.80	0.19	−2.72	0.22
Heart disease	−2.42	0.18	−2.39	0.18
Stroke	−1.95	0.34	−2.17	0.32
Arthrosis,Arthrit.	−4.71	0.12	−4.71	0.13
Osteoporosis	−2.01	0.16	−2.04	0.16
Hepatic cirrhosis	−1.79*	0.68	−2.00	0.58
Cancer	−3.82	0.36	−3.75	0.40

Continued on next page

– *continued from previous page*

MODEL Quantile	Model 2 Quantile 0.25		Model 3 Quantile 0.25	
Fixed effects	Coef.	S.E.	Coef.	S.E.
Parkinsonism	−1.51	0.42	−1.35*	0.47
Ulcer	−0.57*	0.18	−0.55**	0.22
Cancer in the past	−1.52	0.28	−1.44	0.30
Alzheimer	−0.41***	0.45	−0.43***	0.51
Stones	−0.97	0.18	−0.99	0.21
Thyroid disease	−0.60	0.16	−0.59*	0.18
Asthma	−1.22	0.16	−1.19	0.17
Diabetes	−1.54	0.14	−1.54	0.16
Hypertension	−0.50	0.09	−0.44	0.09
Bronchitis,Emph.	−0.63	0.14	−0.66	0.14
Household composition: ref. 1 component				
2 components	−0.33**	0.13		
3 components	−0.46*	0.14		
4 components	−0.24***	0.16		
5 or more comp.	−0.28***	0.18		
Household typology: ref. Living alone				
Childless couple			−0.52	0.17
Couple with child.			−0.57	0.17
Other typology			−0.56*	0.18
Insuf. econ. resources	−0.19***	0.12	−0.36**	0.14
Urban degree: ref. Low				
Intermediate	0.36	0.09	0.30*	0.09
High	0.51	0.09	0.46	0.11
Random effects	Paramet.	S.E.	Paramet.	S.E.
2-Household: sd(u)	0.00		0.00	
ALD scale (sd)	1.54(6.49)	0.01	1.54(6.49)	0.01
Log-likelihood	-347730.8		-347749.9	
AIC	695551.7		695587.7	

*Coefficient not significant: *** "p¿0.05"; ** "p¿0.01"; * "p¿0.001"*

Table 4.4: Multilevel linear quantile regression: estimated parameters and standard errors for model 2 and model 3: quantile 0.33

MODEL	Model 2		Model 3	
Quantile	Quantile 0.33		Quantile 0.33	
Fixed effects	Coef.	S.E.	Coef.	S.E.
Intercept	52.44	0.32	52.39	0.37
Age	0.040	0.008	0.045	0.011
Age^ 2	−0.002	0.000	−0.002	0.000
Female	−0.14*	0.04	−0.17	0.04
Qualification: ref. No Educ. Qualif.				
University	1.19	0.23	1.16	0.23
Upper Secondary	1.01	0.23	0.96	0.23
Lower Secondary	0.80	0.23	0.74*	0.22
Primary	0.40***	0.21	0.32***	0.19
Employment: ref. Employed				
Unemployed	−0.42	0.08	−0.40	0.08
Not workinging	−0.48	0.06	−0.45	0.06
Unable to work	−4.67	0.32	−4.62	0.29
Body-mass Index	−0.36	0.03	−0.36	0.03
Sport	0.43	0.04	0.44	0.04
Disabled	−10.65	0.47	−10.53	0.38
Disease: ref. No health problem				
Illness	−1.58	0.08	−1.58	0.07
Acute disease	−7.76	0.13	−7.78	0.14
Interaction: Disabled x Disease				
Disabled-Illness	1.41*	0.42	1.33*	0.40
Disabled-Acute disease	6.59	0.46	6.52	0.43
Well-being (VT)	1.12	0.02	1.12	0.02
Infarct	−2.80	0.27	−2.84	0.24
Heart disease	−2.61	0.15	−2.59	0.14
Stroke	−2.18	0.29	−2.13	0.24
Arthrosis,Arthrit.	−4.72	0.12	−4.74	0.10
Osteoporosis	−2.16	0.13	−2.14	0.13
Hepatic cirrhosis	−1.58**	0.61	−1.54*	0.56
Cancer	−3.87	0.31	−3.83	0.29

Continued on next page

– *continued from previous page*

MODEL Quantile	Model 2 Quantile 0.33		Model 3 Quantile 0.33	
Fixed effects	Coef.	S.E.	Coef.	S.E.
Parkinsonism	−1.88	0.42	−1.84	0.41
Ulcer	−0.61	0.18	−0.66	0.16
Cancer in the past	−1.39	0.28	−1.33	0.29
Alzheimer	−0.38***	0.34	−0.40***	0.31
Stones	−0.98	0.16	−0.99	0.19
Thyroid disease	−0.66	0.14	−0.64	0.13
Asthma	−1.32	0.16	−1.27	0.14
Diabetes	−1.61	0.15	−1.61	0.13
Hypertension	−0.47	0.08	−0.48	0.08
Bronchitis,Emph.	−0.93	0.14	−0.95	0.12
Household composition: ref. 1 component				
2 components	−0.19***	0.10		
3 components	−0.24**	0.11		
4 components	−0.13***	0.11		
5 or more comp.	−0.13***	0.13		
Household typology: ref. Living alone				
Childless couple			−0.22**	0.10
Couple with child.			−0.21**	0.10
Other typology			−0.20**	0.10
Insuf. econ. resources	−0.20**	0.10	−0.16**	0.07
Urban degree: ref. Low				
Intermediate	0.14***	0.10	0.20*	0.07
High	0.29*	0.10	0.33	0.07
Random effects	Paramet.	S.E.	Paramet.	S.E.
2-Household: sd(u)	2.21		2.21	
ALD scale (sd)	1.88(6.35)	0.01	1.88(6.35)	0.01
Log-likelihood	-342094.6		-342090.9	
AIC	684279.1		684269.7	

*Coefficient not significant: *** "p¿0.05"; ** "p¿0.01"; * "p¿0.001"*

Table 4.5: Multilevel linear regression: model 2 and model 3 without level 3.

MODEL	Model 2		Model 3	
Fixed effects	Coef.	S.E.	Coef.	S.E.
Intercept	53.38	0.20	53.37	0.19
Age	0.039	0.005	0.040	0.005
Age^ 2	−0.002	0.000	−0.002	0.000
Female	−0.12*	0.04	−0.12*	0.04
Qualification: ref. No Educ. Qualif.				
University	1.33	0.15	1.33	0.15
Upper Secondary	1.18	0.13	1.18	0.13
Lower Secondary	0.88	0.11	0.88	0.11
Primary	0.52	0.12	0.52	0.12
Employment: ref. Employed				
Unemployed	−0.30*	0.09	−0.31*	0.09
Not workinging	−0.44	0.05	−0.44	0.05
Unable to work	−4.42	0.24	−4.42	0.24
Body-mass Index	−0.44	0.03	−0.44	0.03
Sport	0.64	0.04	0.64	0.04
Disabled	−8.89	0.27	−8.90	0.27
Disease: ref. No health problem				
Illness	−1.42	0.09	−1.42	0.09
Acute disease	−6.01	0.11	−6.00	0.11
Interaction: Disabled x Disease				
Disabled-Illness	0.76**	0.29	0.76**	0.29
Disabled-Acute disease	3.62	0.29	3.62	0.29
Well-being (VT)	1.11	0.02	1.11	0.02
Infarct	−2.57	0.19	−2.56	0.19
Heart disease	−2.32	0.12	−2.32	0.12
Stroke	−2.03	0.20	−2.03	0.20
Arthrosis,Arthrit.	−3.86	0.08	−3.86	0.08
Osteoporosis	−1.89	0.11	−1.89	0.11
Hepatic cirrhosis	−1.82	0.46	−1.82	0.46
Cancer	−3.45	0.27	−3.45	0.27
Parkinsonism	−1.61	0.34	−1.61	0.34

Continued on next page

– *continued from previous page*

MODEL	Model 2		Model 3	
Fixed effects	Coef.	S.E.	Coef.	S.E.
Ulcer	−0.65	0.17	−0.65	0.17
Cancer in the past	−1.18	0.19	−1.18	0.19
Alzheimer	−0.68**	0.32	−0.68**	0.32
Stones	−1.07	0.13	−1.06	0.13
Thyroid disease	−0.56	0.12	−0.56	0.12
Asthma	−1.19	0.12	−1.19	0.12
Diabetes	−1.30	0.12	−1.30	0.12
Hypertension	−0.40	0.07	−0.40	0.07
Bronchitis,Emph.	−0.76	0.10	−0.75	0.10
Household composition: ref. 1 component				
2 components	−0.36	0.07		
3 components	−0.47	0.07		
4 components	−0.39	0.08		
5 or more comp.	−0.45	0.10		
Household typology: ref. Living alone				
Childless couple			−0.42	0.08
Couple with child.			−0.42	0.07
Other typology			−0.34	0.09
Insuf. econ. resources	−0.19*	0.06	−0.20*	0.06
Urban degree: ref. Low				
Intermediate	0.32	0.07	0.32	0.07
High	0.46	0.08	0.46	0.08
Random effects	Paramet.	S.E.	Paramet.	S.E.
2-Household: sd(u)	2.25	0.04	2.25	0.04
1-Residuals: sd(e)	5.70	0.05	5.70	0.05
Log-likelihood	-341341.3		-341342.6	
AIC	682772.7		682773.2	

*Coefficient not significant: *** "p¿0.05"; ** "p¿0.01"; * "p¿0.001"*

Table 4.6: Multilevel linear quantile regression: estimated parameters and standard errors for model 2 and model 3: quantile 0.50.

MODEL	Model 2		Model 3	
Quantile	Quantile 0.50		Quantile 0.50	
Fixed effects	Coef.	S.E.	Coef.	S.E.
Intercept	54.43	0.49	54.23	0.58
Age	0.059	0.009	0.062	0.010
Age^ 2	−0.002	0.000	−0.002	0.000
Female	−0.20	0.04	−0.20	0.04
Qualification: ref. No Educ. Qualif.				
University	0.83**	0.34	0.94**	0.36
Upper Secondary	0.70**	0.30	0.81**	0.33
Lower Secondary	0.56***	0.30	0.66**	0.32
Primary	0.25***	0.28	0.36***	0.28
Employment: ref. Employed				
Unemployed	−0.07***	0.07	−0.08***	0.07
Not working	−0.07***	0.05	−0.04***	0.07
Unable to work	−4.36	0.37	−4.38	0.37
Body-mass Index	−0.23	0.03	−0.23	0.03
Sport	0.32	0.04	0.32	0.05
Disabled	−11.03	0.64	−11.17	0.68
Disease: ref. No health problem				
Ilness	−0.84	0.07	−0.82	0.08
Acute disease	−6.64	0.18	−6.64	0.17
Interaction: Disabled x Disease				
Disabled-Ilness	−0.24***	0.76	−0.12***	0.72
Disabled-Acute disease	4.15	0.79	4.29	0.83
Well-being (VT)	0.63	0.07	0.63	0.08
Infarct	−2.64	0.32	−2.72	0.28
Heart disease	−2.86	0.19	−2.89	0.22
Stroke	−2.32	0.38	−2.36	0.44
Arthrosis,Arthrit.	−4.56	0.16	−4.56	0.17
Osteoporosis	−2.42	0.14	−2.39	0.18
Hepatic cirrhosis	−2.14*	0.66	−2.18*	0.68
Cancer	−3.89	0.40	−3.94	0.49

Continued on next page

– *continued from previous page*

MODEL	Model 2		Model 3	
Quantile	Quantile 0.50		Quantile 0.50	
Fixed effects	Coef.	S.E.	Coef.	S.E.
Parkinson	−2.18	0.44	−2.12	0.46
Ulcer	−0.62*	0.21	−0.64*	0.23
Cancer in the past	−1.09	0.31	−1.16*	0.35
Alzheimer	−1.33*	0.43	−1.37*	0.52
Stones	−1.09	0.24	−1.11	0.22
Thyroid disease	−0.54	0.16	−0.53*	0.17
Asthma	−1.08	0.14	−1.08	0.15
Diabetes	−1.71	0.14	−1.70	0.16
Hypertension	−0.49	0.07	−0.48	0.08
Bronchitis,Emph.	−1.24	0.18	−1.23	0.14
Household composition: ref. 1 component				
2 components	−0.17***	0.12		
3 components	−0.20***	0.12		
4 components	−0.15***	0.13		
5 or more comp.	−0.15***	0.13		
Household typology: ref. Living alone				
Childless couple			−0.15***	0.13
Couple with child.			−0.14***	0.14
Other typology			−0.12***	0.13
Insuf. econ. resources	−0.14***	0.10	−0.11***	0.12
Urban degree: ref. Low				
Intermediate	0.11**	0.05	0.11***	0.07
High	0.19	0.05	0.19*	0.07
Random effects	Paramet.	S.E.	Paramet.	S.E.
2-Household: sd(u)	0.22		0.19	
ALD scale (sd)	2.13(6.01)	0.01	2.13(6.01)	0.01
Log-likelihood	-332549.2		-332543.7	
AIC	665188.3		665175.4	

*Coefficient not significant: *** "p¿0.05"; ** "p¿0.01"; * "p¿0.001"*

Table 4.7: Multilevel linear quantile regression: estimated parameters and standard errors for model 2 and model 3: quantile 0.67.

MODEL	Model 2		Model 3	
Quantile	Quantile 0.67		Quantile 0.67	
Fixed effects	Coef.	S.E.	Coef.	S.E.
Intercept	55.71	0.43	55.48	0.45
Age	0.039	0.011	0.048	0.013
Age^ 2	−0.001	0.000	−0.001	0.000
Female	−0.24	0.04	−0.20	0.04
Qualification: ref. No Educ. Qualif.				
University	0.73**	0.32	0.88*	0.32
Upper Secondary	0.56***	0.32	0.75**	0.32
Lower Secondary	0.44***	0.30	0.66**	0.30
Primary	0.28***	0.29	0.47***	0.28
Employment: ref. Employed				
Unemployed	0.14***	0.08	0.17***	0.09
Not workinging	0.08***	0.06	0.09***	0.07
Unable to work	−3.75	0.35	−3.83	0.32
Body-mass Index	−0.19	0.03	−0.18	0.03
Sport	0.23	0.04	0.25	0.03
Disabled	−10.03	0.77	−10.06	0.66
Disease: ref. No health problem				
Illness	−0.44	0.04	−0.42	0.04
Acute disease	−4.02	0.15	−4.04	0.15
Interaction: Disabled x Disease				
Disabled-Illness	−2.09*	0.67	−1.97*	0.68
Disabled-Acute disease	−1.02***	0.67	−0.82***	0.58
Well-being (VT)	0.39	0.02	0.38	0.02
Infarct	−2.53	0.27	−2.49	0.28
Heart disease	−2.77	0.16	−2.77	0.18
Stroke	−2.34	0.33	−2.29	0.30
Arthrosis,Arthrit.	−3.31	0.10	−3.29	0.09
Osteoporosis	−2.33	0.16	−2.37	0.15
Hepatic cirrhosis	−1.38**	0.54	−1.31**	0.54

Continued on next page

– *continued from previous page*

MODEL Quantile	Model 2 Quantile 0.67		Model 3 Quantile 0.67	
Fixed effects	Coef.	S.E.	Coef.	S.E.
Cancer	−3.38	0.36	−3.42	0.40
Parkinsonism	−2.65	0.53	−2.57	0.60
Ulcer	−0.63	0.16	−0.62	0.15
Cancer in the past	−0.75	0.20	−0.72*	0.22
Alzheimer	−1.94	0.53	−1.97	0.53
Stones	−0.73	0.15	−0.74	0.14
Thyroid disease	−0.35	0.10	−0.38	0.09
Asthma	−0.84	0.11	−0.84	0.11
Diabetes	−1.25	0.12	−1.27	0.14
Hypertension	−0.37	0.06	−0.40	0.06
Bronchitis,Emph.	−1.36	0.12	−1.30	0.12
Household composition: ref. 1 component				
2 components	−0.09***	0.12		
3 components	−0.14***	0.13		
4 components	−0.13***	0.13		
5 or more comp.	−0.09***	0.12		
Household typology: ref. Living alone				
Childless couple			−0.20*	0.08
Couple with child.			−0.24*	0.09
Other typology			−0.15***	0.08
Insuf. econ. resources	−0.10***	0.06	−0.12**	0.05
Urban degree: ref. Low				
Intermediate	0.10**	0.05	0.15*	0.05
High	0.17	0.05	0.22	0.05
Random effects	Paramet.	S.E.	Paramet.	S.E.
2-Household: sd(u)	0.23		0.24	
ALD scale (sd)	1.82(6.15)	0.01	1.82(6.15)	0.01
Log-likelihood	-329120.4		-329107	
AIC	658330.7		658301.9	

*Coefficient not significant: *** "p¿0.05"; ** "p¿0.01"; * "p¿0.001"*

Table 4.8: Multilevel linear quantile regression: estimated parameters and standard errors for model 2 and model 3: quantile 0.75.

MODEL	Model 2		Model 3	
Quantile	Quantile 0.75		Quantile 0.75	
Fixed effects	Coef.	S.E.	Coef.	S.E.
Intercept	56.22	0.31	55.88	0.32
Age	0.037	0.010	0.044	0.008
Age^ 2	−0.001	0.000	−0.001	0.000
Female	−0.16	0.04	−0.16	0.04
Qualification: ref. No Educ. Qualif.				
University	0.56*	0.18	0.73	0.19
Upper Secondary	0.40**	0.18	0.63	0.18
Lower Secondary	0.30***	0.17	0.53*	0.17
Primary	0.17***	0.16	0.37**	0.17
Employment: ref. Employed				
Unemployed	0.14***	0.07	0.18**	0.07
Not working	0.06***	0.06	0.09***	0.05
Unable to work	−3.80	0.37	−3.75	0.36
Body-mass Index	−0.19	0.02	−0.20	0.02
Sport	0.27	0.03	0.28	0.03
Disabled	−9.13	0.66	−9.28	0.73
Disease: ref. No health problem				
Ilness	−0.36	0.04	−0.35	0.04
Acute disease	−2.69	0.13	−2.65	0.15
Interaction: Disabled x Disease				
Disabled-Ilness	−2.96	0.70	−2.77*	0.86
Disabled-Acute disease	−3.98	0.59	−3.99	0.72
Well-being (VT)	0.31	0.02	0.30	0.02
Infarct	−2.34	0.25	−2.38	0.26
Heart disease	−2.47	0.13	−2.49	0.16
Stroke	−2.17	0.28	−2.19	0.30
Arthrosis,Arthrit.	−2.52	0.07	−2.50	0.09
Osteoporosis	−2.09	0.15	−2.11	0.13

Continued on next page

– *continued from previous page*

MODEL	Model 2		Model 3	
Quantile	Quantile 0.75		Quantile 0.75	
Fixed effects	Coef.	S.E.	Coef.	S.E.
Hepatic cirrhosis	-1.38^{*}	0.45	-1.22^{**}	0.48
Cancer	-3.09	0.30	-3.18	0.31
Parkinson	-2.73	0.52	-2.80	0.44
Ulcer	-0.58	0.15	-0.62	0.18
Cancer in the past	-0.74	0.15	-0.76	0.20
Alzheimer	-2.24	0.49	-2.17	0.48
Stones	-0.76	0.14	-0.78	0.14
Thyroid disease	-0.30^{*}	0.11	-0.34^{*}	0.11
Asthma	-0.81	0.10	-0.86	0.11
Diabetes	-1.11	0.09	-1.11	0.12
Hypertension	-0.36	0.05	-0.36	0.06
Bronchitis,Emph.	-1.17	0.12	-1.16	0.10
Household composition: ref. 1 component				
2 components	-0.05^{***}	0.12		
3 components	-0.10^{***}	0.13		
4 components	-0.10^{***}	0.14		
5 or more comp.	-0.09^{***}	0.14		
Household typology: ref. Living alone				
Childless couple			-0.12^{***}	0.10
Couple with child.			-0.16^{***}	0.13
Other typology			-0.07^{***}	0.12
Insuf. econ. resources	-0.05^{***}	0.06	-0.01^{***}	0.06
Urban degree: ref. Low				
Intermediate	0.12^{**}	0.06	0.12^{**}	0.05
High	0.17^{*}	0.05	0.18	0.05
Random effects	Paramet.	S.E.	Paramet.	S.E.
2-Household: sd(u)	0.00		0.00	
ALD scale (sd)	1.54(6.49)	0.01	1.54(6.49)	0.01
Log-likelihood	-329787.8		-329764.7	
AIC	659665.6		659617.3	

*Coefficient not significant: *** "p¿0.05"; ** "p¿0.01"; * "p¿0.001"*

Table 4.9: Multilevel linear quantile regression: estimated parameters and standard errors for model 2 and model 3: quantile 0.90.

MODEL	Model 2		Model 3
Quantile	Quantile 0.90		Quantile 0.90
Fixed effects	Coef.	S.E.	Coef.
Intercept	57.33	0.47	56.96
Age	0.041**	0.016	0.051
Age^ 2	−0.001	0.000	−0.002
Female	−0.17*	0.05	−0.20
Qualification: ref. No Educ. Qualif.			
University	0.81**	0.32	0.99
Upper Secondary	0.79**	0.31	0.95
Lower Secondary	0.57***	0.31	0.73
Primary	0.36***	0.29	0.48
Employment: ref. Employed			
Unemployed	0.13***	0.14	0.27
Not workinging	0.06***	0.10	0.17
Unable to work	−3.71	0.46	−3.27
Body-mass Index	−0.28	0.04	−0.29
Sport	0.39	0.06	0.42
Disabled	−6.68	0.57	−5.92
Disease: ref. No health problem			
Illness	−0.48	0.09	−0.44
Acute disease	−2.75	0.17	−2.69
Interaction: Disabled x Disease			
Disabled-Illness	−2.83	0.70	−2.78
Disabled-Acute disease	−4.85	0.75	−5.62
Well-being (VT)	0.37	0.03	0.34
Infarct	−1.94	0.31	−1.83
Heart disease	−2.28	0.20	−2.13
Stroke	−2.06	0.44	−1.77
Arthrosis,Arthrit.	−2.31	0.10	−2.28
Osteoporosis	−1.91	0.17	−1.75

Continued on next page

– *continued from previous page*

MODEL	Model 2		Model 3
Quantile	Quantile 0.90		Quantile 0.90
Fixed effects	Coef.	S.E.	Coef.
Hepatic cirrhosis	−0.73***	0.57	−0.83
Cancer	−2.52	0.53	−2.60
Parkinsonism	−2.74	0.64	−3.18
Ulcer	−0.76*	0.24	−0.48
Cancer in the past	−0.95*	0.36	−0.63
Alzheimer	−1.92	0.56	−2.65
Stones	−0.72*	0.24	−0.56
Thyroid disease	−0.32***	0.17	−0.36
Asthma	−0.94	0.22	−0.75
Diabetes	−0.95	0.17	−0.95
Hypertension	−0.38	0.09	−0.37
Bronchitis,Emph.	−0.91	0.16	−0.87
Household composition: ref. 1 component			
2 components	−0.13***	0.13	
3 components	−0.03***	0.14	
4 components	−0.04***	0.16	
5 or more comp.	−0.07***	0.18	
Household typology: ref. Living alone			
Childless couple			−0.07
Couple with child.			0.08
Other typology			0.21
Insuf. econ. resources	−0.08***	0.09	−0.11
Urban degree: ref. Low			
Intermediate	0.31*	0.11	0.33
High	0.39	0.11	0.36
Random effects	Paramet.	S.E.	Paramet.
2-Household: sd(u)	0.00		2.56
ALD scale (sd)	0.63(6.35)	0.01	0.64(6.41)
Log-likelihood	-334647.5		-334698.5
AIC	669385.1		669485.1

*Coefficient not significant: *** "p¿0.05"; ** "p¿0.01"; * "p¿0.001"*

The MLQR analysis reveals more information than the MLM. Describe this additional information can be cumbersome, in particular when a long sequence of quantile values are carried-out. A graphical view of MLQR estimates becomes a necessary step in interpreting of the results. Consequently for each covariate, arrays of coefficients for a range of quantiles can be used to determine how a one-unit increase in the covariate affects on the response variable. A typical way to highlight these effects is to plot covariate coefficients and confidence interval in a graphic where a predictor variable effect $\hat{\beta}_\tau$ is on the y-axis and the quantile value τ is on the x-axis.

Following figures report for model 2 these graphs to which is superimposed MLM estimates and confidence intervals. Tow additional graphs on Disabled-Illness and Disabled-Acute disease interaction coefficients are show summing effects of related coefficients. Others two similar graphs are added: the one on standard deviation of household random effects and the other on AIC values.

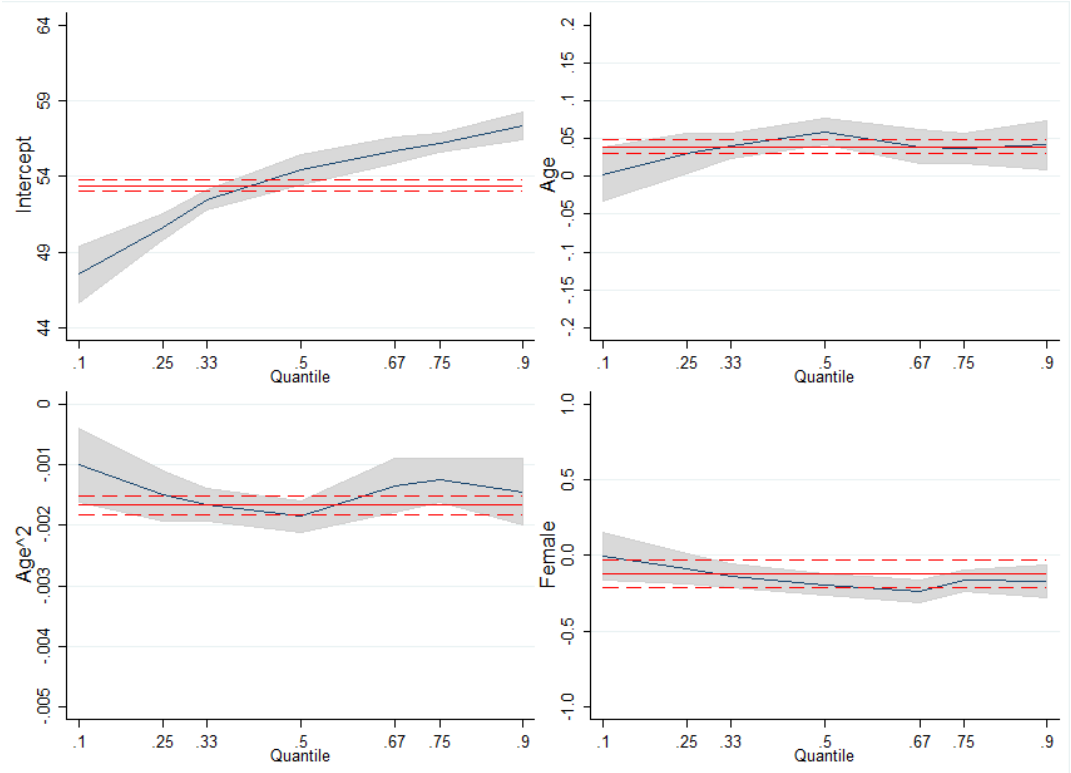


Figure 4.1: MLQR (blue) and MLM (red) estimates and 95% confidence intervals for PCS model 2.

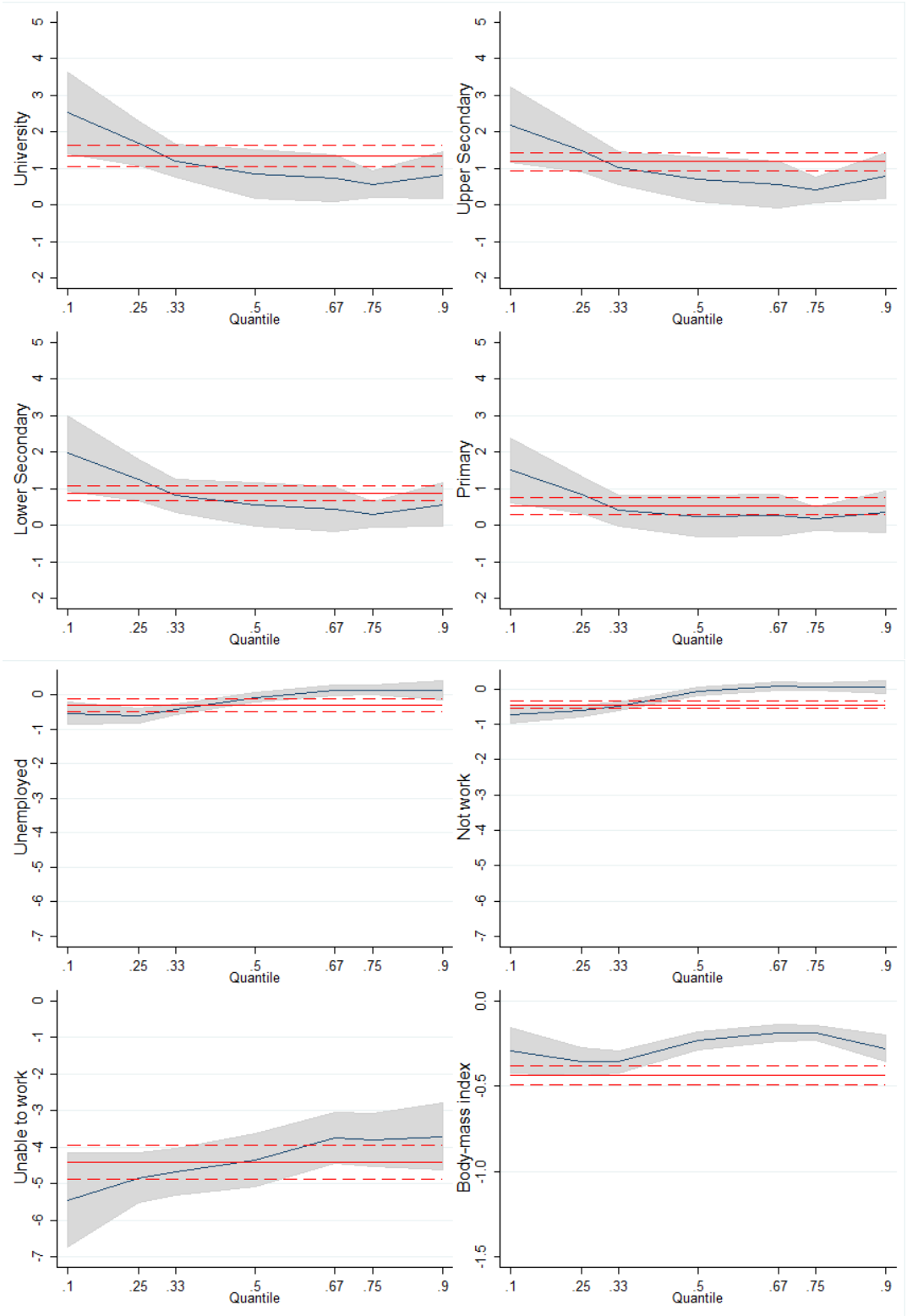


Figure 4.2: MLQR (blue) and MLM (red) estimates and 95% confidence intervals for PCS model 2.

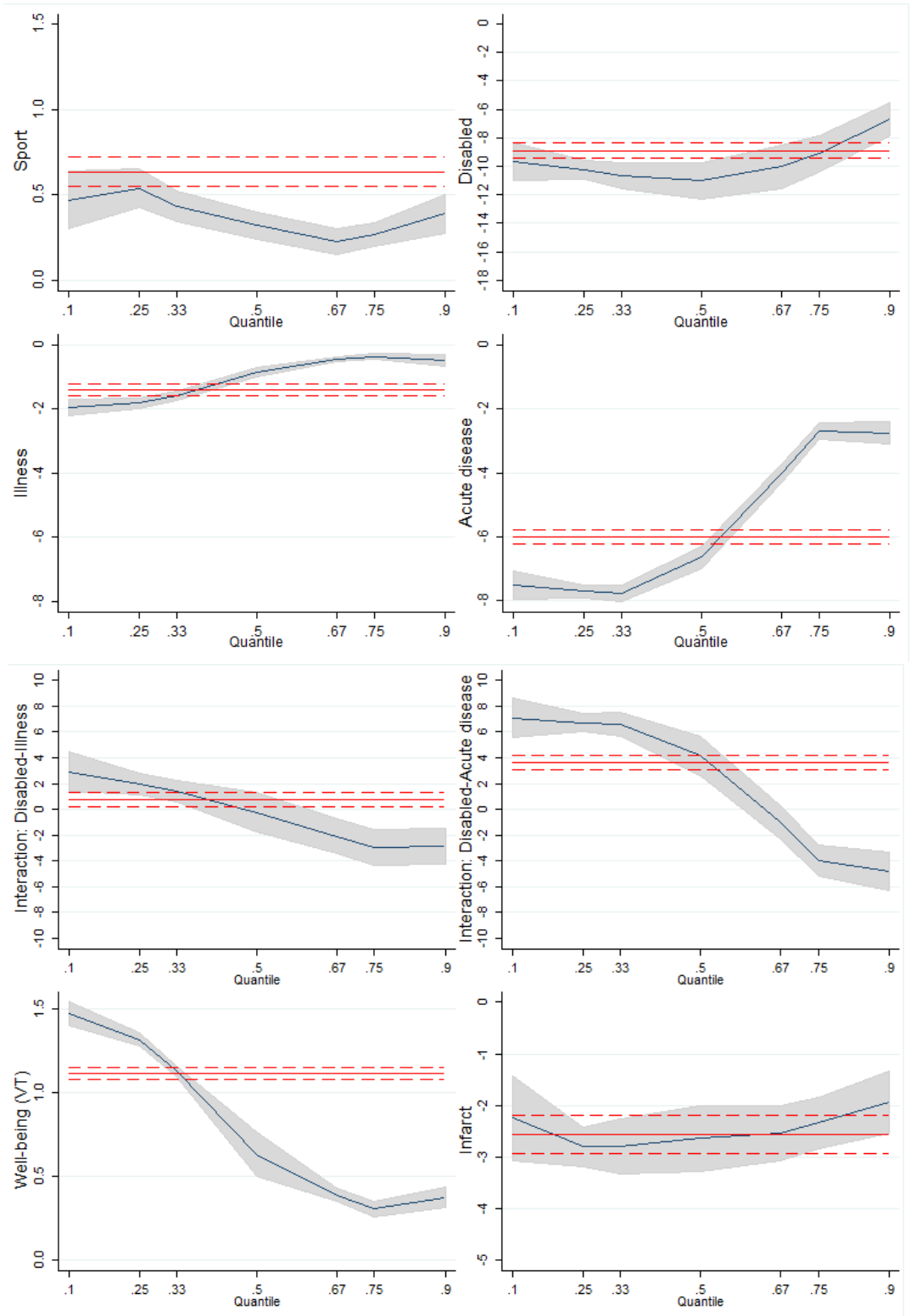


Figure 4.3: MLQR (blue) and MLM (red) estimates and 95% confidence intervals for PCS model 2.

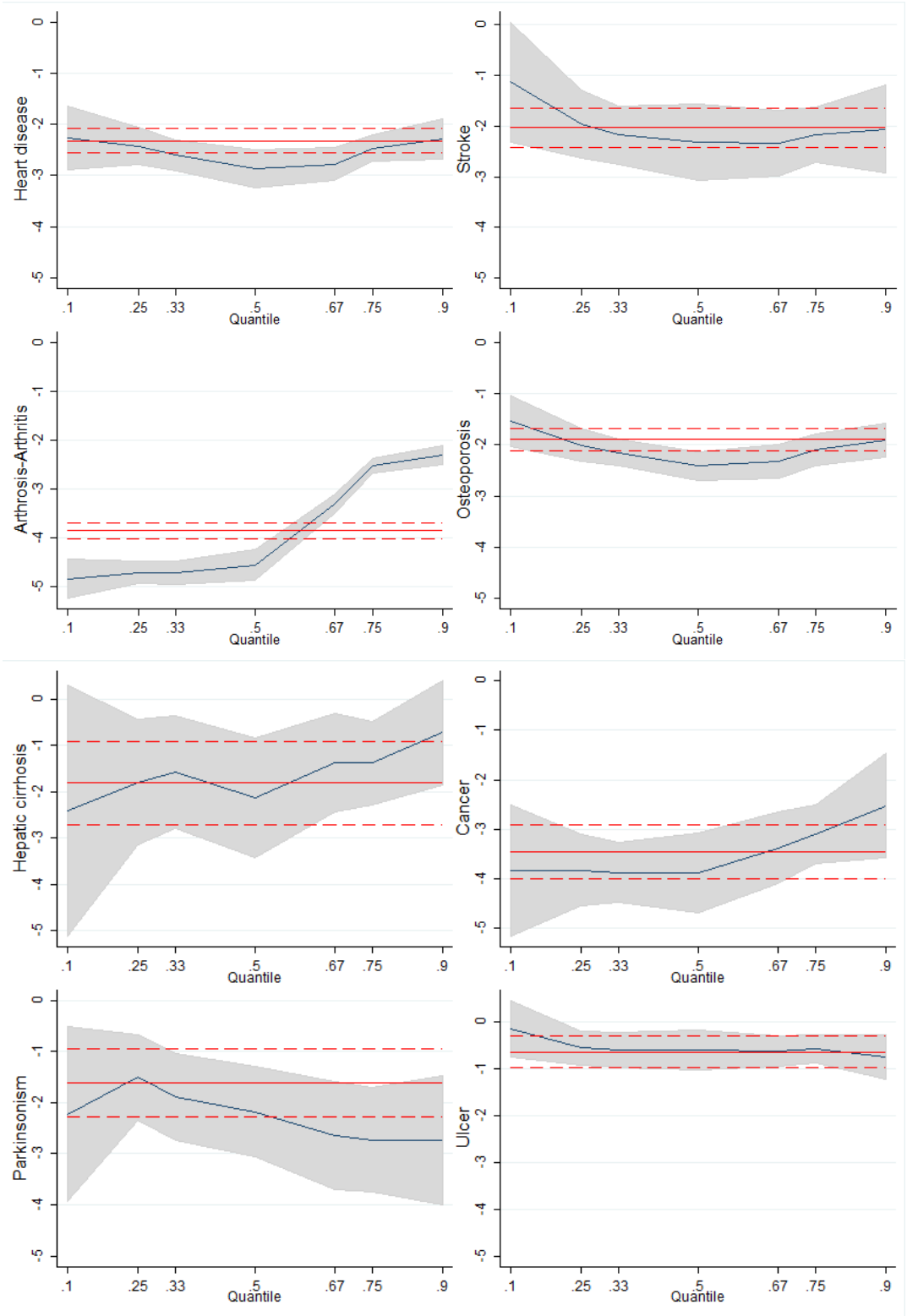


Figure 4.4: MLQR (blue) and MLM (red) estimates and 95% confidence intervals for PCS model 2.

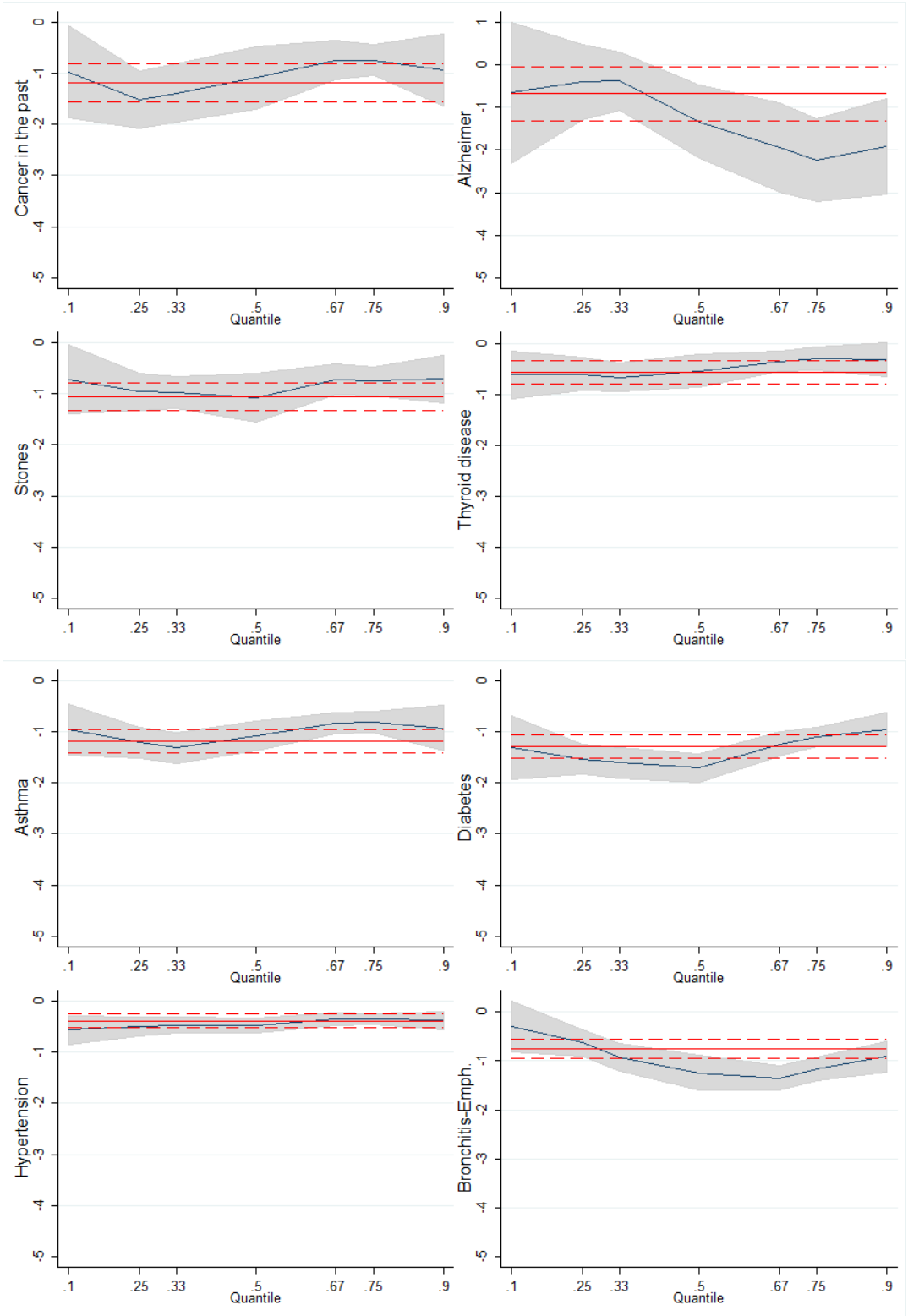


Figure 4.5: MLQR (blue) and MLM (red) estimates and 95% confidence intervals for PCS model 2.

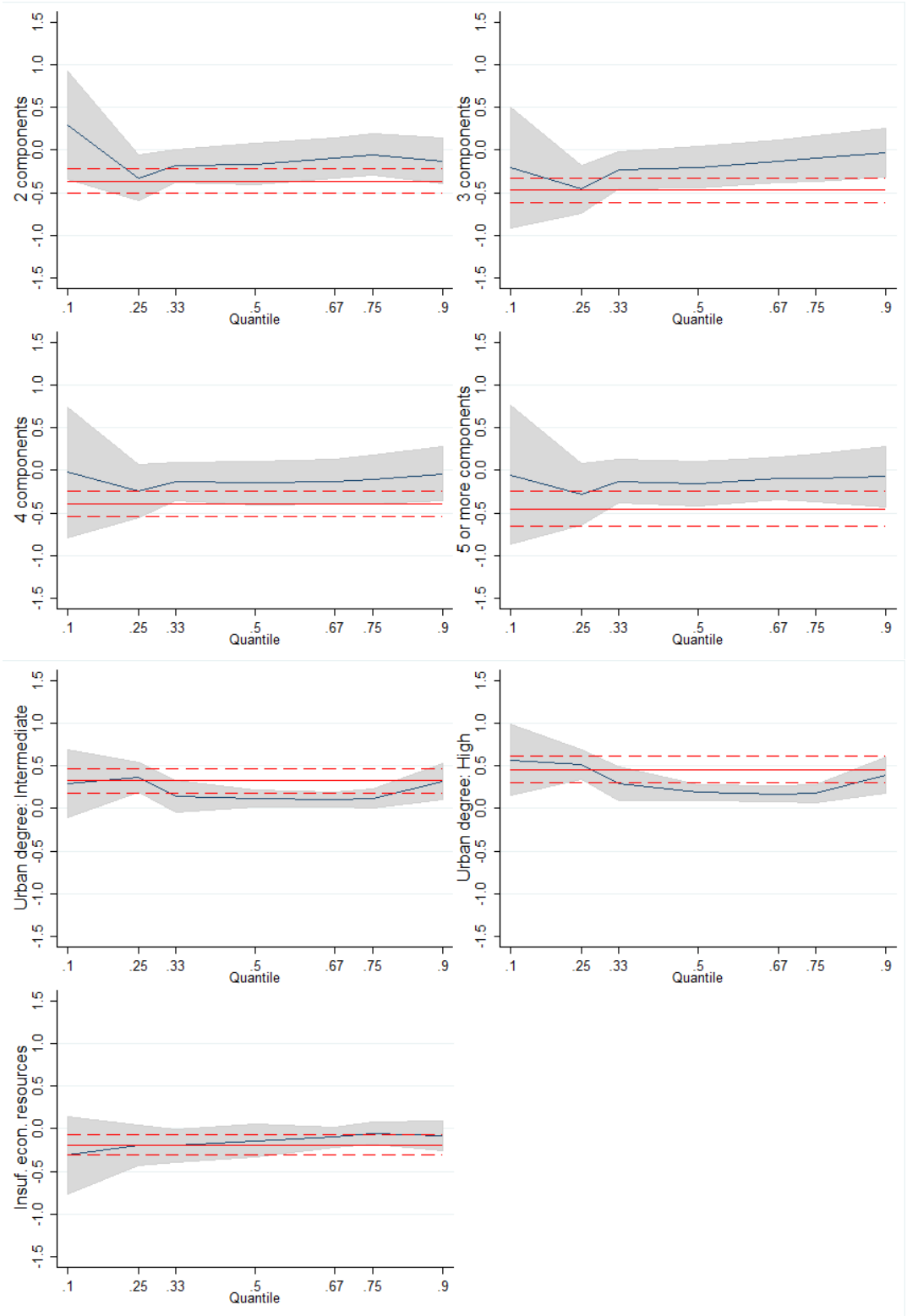


Figure 4.6: MLQR (blue) and MLM (red) estimates and 95% confidence intervals for PCS model 2.

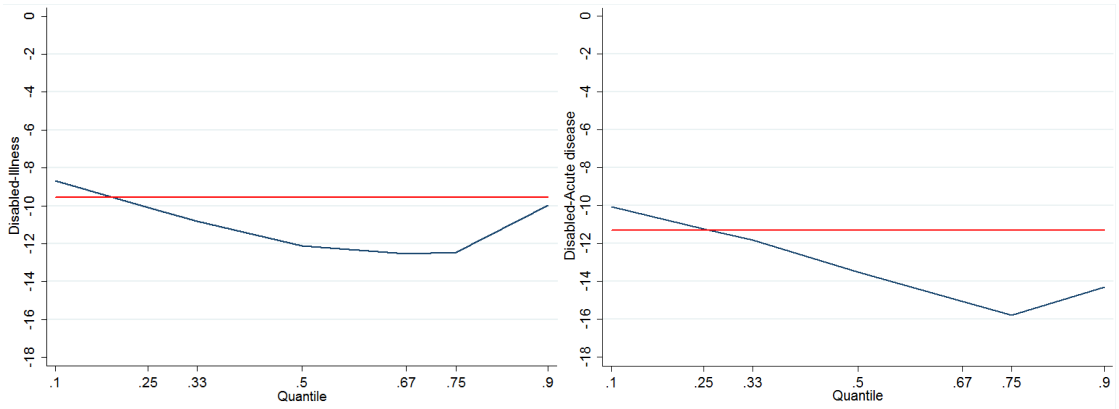


Figure 4.7: MLQR (blue) and MLM (red): combined effect for Interaction Disabled-Illness (on the left) and for Interaction Disabled-Acute disease (on the right) for PCS model 2.

Comparing model 2 and model 3 MLQR estimates appear a substantial overlap of results for fixed effects; only in quantile 0.10 and quantile 0.90 there are some differences (Table 4.2 and 4.9). This is not a surprise because, in the QR, study of the tails of distribution requires a suitable number of observations to ensure stability of the estimates, that is rare to found. The equivalence of point estimates between models 2 and 3 appear evident even comparing results of MLM estimates (Table 4.5) and then we can have confidence on fixed part estimate stability of this novel procedure.
The same thing can not be said for random part parameters estimate due to large difference on level 2 variance estimated for quantile 0.90: sd(u)=0 in model 2 and sd(u)=2.46 in model 3. On the other hand, also the values of intercept standard deviation by quantile seems a bit dubious. Looking on the left of Figure 4.8 we see a fluctuating trend of the random effects standard deviation along the quantiles; particularly surprise us the null standard deviation of household intercept at 0.25 quantile neighbouring to quantile 0.33 value of 2.21, and perfectly comparable with the MLM estimate, 2.25. Nevertheless, the anomalous fluctuations of random effects parameter can be due to very limited size of the level 2 units.
In Figure 4.9 the quantile-quantile plots of random effects of model 2 reveal persistence of some not normal distribution even in MLQR estimate (maybe assume random effects ALD distributed could be improve the estimations of random part of the model).

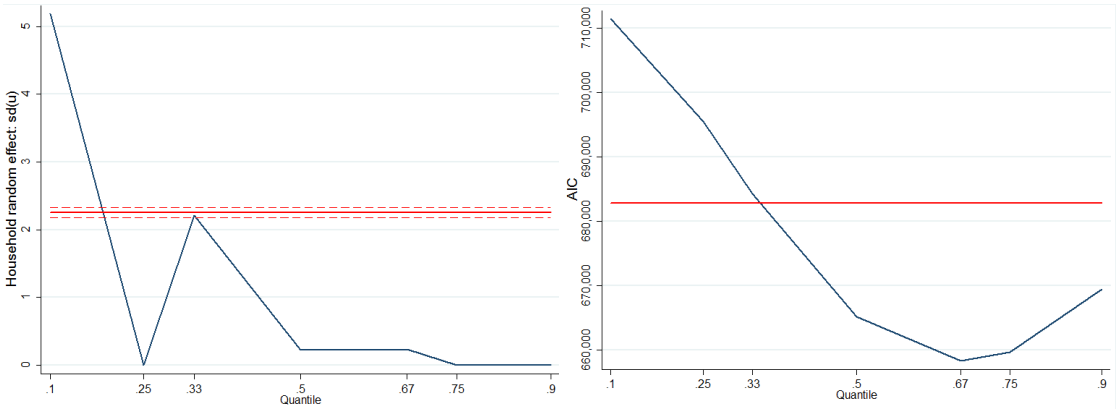


Figure 4.8: MLQR (blue) and MLM (red) estimated standard deviation of household random effects (on the left) and AIC (Akaike's Information Criterion) values (on the right) for PCS model 2.

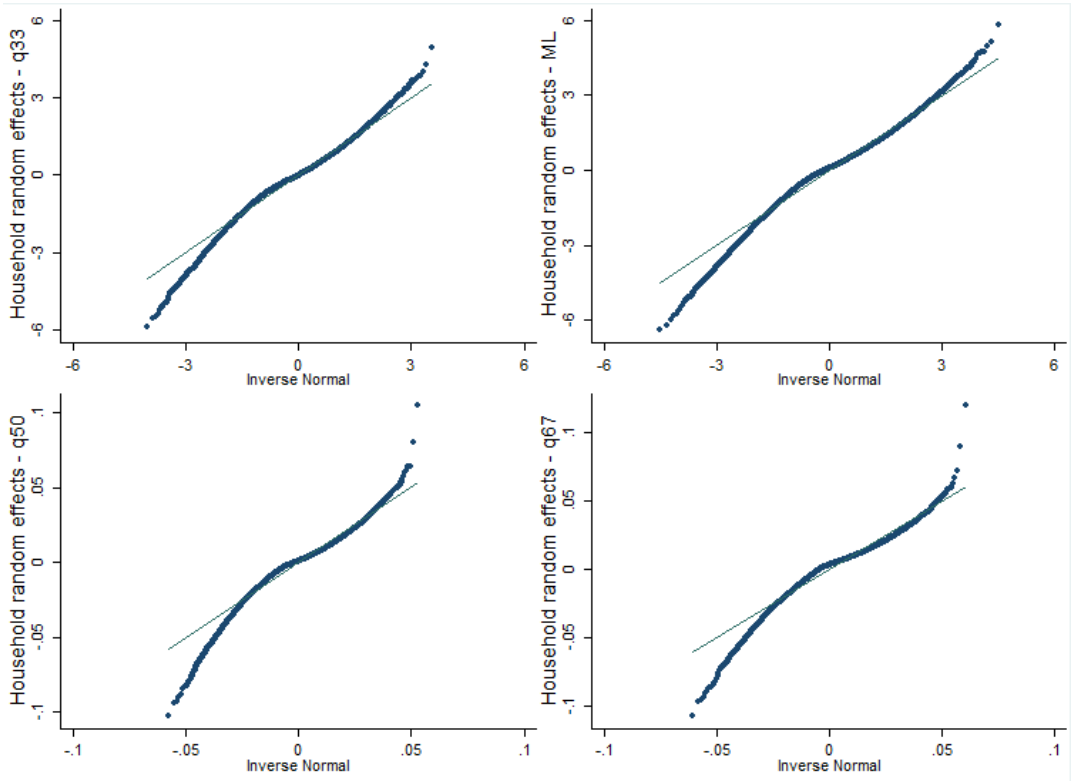


Figure 4.9: Normal quantile-quantile plot of MLQR (quantiles 0.33, 0.50 and 0.67) and MLM household random effects of model 2.

In Figure 4.8, AIC values show better fitting for quantiles 0.67 and 0.75 corresponding to PCS distribution with higher frequencies. MLQR and MLM AIC values intersect at quantile 0.33.
For fixed part we will confine our comment to only some of the covariates. Coefficient standard errors performed with 100 bootstrap replication seem be enough. Observing all covariate panels (Fig. from 4.1 to 4.6) we note for almost all covariates, with the exception of Sport, Disabled, Acute disease and Arthrosis, that MLQL estimates and the MLM estimate intersect at quantile 0.33. This is a sign that the MLM estimates are pulled towards left tail of the distribution, and their location are at about a third of the distribution.
The first panel in Figure 4.1 display intercept estimates and may be interpreted as the estimated conditional quantile function of the PCS score distribution of a male having 18 years old, without education qualification, employed, normal-weight, do not practice sport, not disabled, without health problem and chronic disease, living alone with sufficient economic resource in a little town. Others panel in this figure show that the effect of age and gender are similar to MLM estimate along the distribution.
Education covariates is associated with modest increase of PCS score and the positive effect tends to decrease towards upper quantiles. The same decreasing magnitude happens for several other covariate: Unemployed, Not working, Unable to work, Well-being (VT), Arthrosis, Urban Degree, sometime losing significance. Infarct, Stroke, Ulcer, Stones, Hypertension and Insufficient econominc resource are covariate having an impact constant along the conditional distribution, and then exerting a pure location shift effect. For these variables MLQR results are quite consistent with MLM results.
Of particular interest are Acute disease, Well-being (VT) and Arthrosis-Arthritis quantile coefficient estimates. Clearly in these covariates the impact on PCS score is different across the conditional distribution and MLM estimates result an intermediate effect. As evidenced by quantile regression coefficients (Figure 4.3) an Acute disease involved greater disparity in lower quantile than in upper quantile; about 52 percent of subjects with acute disease have a negative difference on PCS score two points higher than the one estimated by MLM (-7.8 and -6 respectively for quantile 0.33).
Similar evaluation may be done for Arthrosis coefficient (figure 4.4): 55 percent of subjects affected by this pathology have perceived a negative effect on physical health higher than the one estimated by MLM.
The MLQR analysis brings out an interesting differential effects of several covariates along conditional distribution of response variable taking into account the dependence between observations.

Chapter 5

Conclusions and perspectives

In this chapter we summarize the main results of this work and some possible analysis to realize in the future.
This thesis aims to perform a Multilevel linear model for data collected by the Istat survey on Health conditions in 2004-2005. The response variable chosen for the analysis is the Physical Component Summary (PCS) index based on SF-12 questionnaire. For Multilevel analysis three level of hierarchy are considered: individual (level 1), household (level 2) and municipality (level 3).
Our study shows a high degree of homogeneity within level 1 units belonging from the same group, with an intraclass correlation of 27% in a level-2 null model. Considering a level-3 null model, the largest amount of variance lies at level 1, almost 73%, a percentage of almost 25 is at level 2 and a residual quote is present at level 3, more than 2%.
Considerable heteroscedasticity on age are detected and modelled on level-1 residuals by using a dummy variable that distinguishes for age class and sex. Three final models were estimated which are distinguished by the presence of Vitality (well-being) index or for level-2 covariate on household composition rather than household typology. Almost all variance is explained by level 1 covariates. In fact, in our model the explanatory variables having more impact on the outcome are disability, unable to work, age and chronic diseases (18 pathologies), while socio-economic factors (measured by the adequacy economic resources available as declared by householder) and contextual variables (such as urban degree and household typology) have a little effect, although significant.
Others external contextual variables are tested without success. Significant positive effect in mean scores are detected increasing with education level.
The percentage of variance explained by the final models is considerable: about 48% and 45% at municipality and household level respectively; percentages varying between 24% and 89% at level 1 and increasing from younger age groups to older ones were explained.

Model-based and design based analysis are performed for one of the final models to assess for informativeness of the survey weights. The results obtained have confirmed the non-informativeness of the weights.
The not normal distribution of residuals and level-2 random effects as well as the markedly skewed of the response variable led us to search for more robust estimation methods. Quantile regression have a good property of robustness and is distribution free. In recent years, with the aim to extend the capability of quantile regression for independent data to deal with hierarchical data, M. Geraci and M. Bottai (2007, 2011) propose a new approach that brings together these two regressions: quantile regression and multilevel regression.
An additional analysis on two final models are performed by using the R procedure `lqmm` developed by M. Geraci (2012) for a "Linear Quantile Mixed Model", here named "Multilevel Linear Quantile Regression", estimate.
This novel procedure of analysis give us the possibility to describe more generally the conditional distribution of the response through the estimation of its quantiles, while accounting for the dependence among the observations. This has represented a great advantage of our models with respect to classic multilevel regression. The median regression with random effects reveals to be more efficient than the mean regression in representation of the outcome central tendency. A more detailed analysis of the conditional distribution of the response on other quantiles highlighted a differential effect of some covariate along the distribution.
This model perform well especially in the fixed part estimates showing stability in the estimated values.
Some anomalous results seem to be present in random part estimates. It is probably due to particular hierarchical structure of the data characterized by many small groups: for some quantiles the intercept variance estimate is equal to zero. In order to perform more appropriate evaluations on capability of this procedure to capture the variability of level 2 could be interesting re-estimate the model raising the second hierarchical level (for example from household to municipality or large area) also allowing the comparison with other studies performed on these data and using a territorial level less detailed than in this study.
As well as in the simple multilevel analysis, the random effects present not normal distribution; perform a new analysis assuming random effects not-normal (e.g. asymmetric Laplace distributed) could improve the results.

Appendice A

Table 1: Descriptive statistics by covariate included in the models.

VARIABLE	N	Mean	Stand.Dev.	Median	IQR
Age					
18-19	2,594	55.29	4.83	56.58	1.97
20-24	7,035	55.08	4.89	56.26	2.04
25-29	7,995	54.51	5.34	56.02	2.68
30-34	9,409	53.77	6.08	55.91	3.77
35-39	10,085	53.36	6.45	55.59	4.46
40-44	10,085	53.01	6.62	55.50	4.63
45-49	9,115	52.14	7.29	55.19	5.63
50-54	8,441	50.89	8.22	54.32	7.78
55-59	8,768	49.70	8.73	53.26	10.12
60-64	7,134	48.03	9.54	51.50	12.59
65-69	7,193	46.13	10.17	49.52	15.75
70-74	6,474	43.55	10.77	45.71	17.87
75-79	5,336	40.58	11.24	41.77	19.36
80-84	3,762	37.68	11.16	37.35	18.89
85-89	1,484	34.05	10.87	32.05	17.53
¿=90	934	31.46	9.91	28.50	13.96
Gender					
Maschi	50,452	51.01	8.81	55.13	7.76
Femmine	55,392	48.87	10.28	53.49	12.59
Qualification					
University	10,105	53.22	6.79	55.61	4.41
Upper Secondary	27,946	53.11	6.97	55.59	4.50
Lower Secondary	37,246	51.27	8.39	54.97	7.21
Primary	24,637	44.70	11.14	47.88	18.47
No Educ. Qual.	5,910	41.88	12.15	43.24	22.57

Continued on next page

– *continued from previous page*

VARIABLE	N	Mean	Stand.Dev.	Median	IQR
Employment status					
Employed	50,198	53.10	6.46	55.50	4.63
Unemployed	5,963	53.33	7.01	55.91	4.77
Not working	47,908	46.80	10.85	50.86	16.33
Unable to work	1,775	30.87	10.14	27.62	11.96
Body-mass Index					
IMC	105,844	49.89	9.67	54.32	10.12
Sport					
No	54,824	48.06	10.85	53.01	14.24
Yes	51,020	51.86	7.74	55.26	6.73
Disabled					
No	100,036	51.05	8.33	54.83	8.27
Yes	5,808	29.95	9.40	27.19	11.00
Well-being (VT) Index					
VT	105,844	49.89	9.67	54.32	10.12
Infarct					
No	103,610	50.15	9.46	54.53	9.61
Yes	2,234	37.82	11.22	36.77	19.66
Heart disease					
No	101,113	50.48	9.17	54.78	8.95
Yes	4,731	37.26	11.33	35.94	19.19
Stroke					
No	104,382	50.11	9.45	54.39	9.68
Yes	1,462	34.12	11.74	30.90	18.88
Arthrosis,Arthrit.					
No	82,251	52.66	7.02	55.30	5.20
Yes	23,593	40.23	11.29	40.78	19.55
Osteoporosis					
No	99,007	50.70	8.99	54.83	8.55
Yes	6,837	38.12	11.38	37.36	19.70
Hepatic cirrhosis					
No	105,519	49.92	9.64	54.32	10.09
Yes	325	39.29	12.14	38.71	22.34
Parkinson					
No	105,448	49.96	9.60	54.32	10.00
Yes	396	31.43	10.46	28.38	14.13

Continued on next page

– *continued from previous page*

VARIABLE	N	Mean	Stand.Dev.	Median	IQR
Ulcer					
No	102,747	50.13	9.51	54.54	9.62
Yes	3,097	42.09	11.57	43.60	19.96
Cancer					
No	104,637	50.03	9.55	54.32	9.79
Yes	1,207	37.67	11.64	36.30	20.73
Cancer in the past					
No	104,271	50.00	9.60	54.32	9.87
Yes	1,573	42.76	11.45	44.44	19.36
Alzheimer					
No	105,258	50.00	9.55	54.32	9.90
Yes	586	29.34	8.71	26.81	8.62
Stones					
No	102,899	50.11	9.52	54.51	9.67
Yes	2,945	42.34	11.59	43.71	20.35
Thyroid disease					
No	101,487	50.09	9.56	54.54	9.71
Yes	4,357	45.35	10.92	48.66	17.93
Asthmatic					
No	101,916	50.18	9.45	54.54	9.54
Yes	3,928	42.47	12.11	44.08	22.20
Diabetic					
No	99,963	50.43	9.26	54.78	9.05
Yes	5,881	40.65	11.62	41.56	20.36
Hypertension					
No	88,028	51.29	8.62	55.25	7.26
Yes	17,816	42.96	11.43	44.95	19.91
Bronchitis,Emphisema					
No	100,011	50.50	9.19	54.78	8.84
Yes	5,833	39.47	11.56	39.13	20.28
Household composition					
1 component	13,471	46.17	11.48	50.50	18.33
2 components	26,506	47.29	10.67	51.50	15.04
3 components	26,765	50.94	8.83	54.84	7.87
4 components	27,172	52.36	7.61	55.37	5.45
5 or more comp.	11,930	51.87	8.35	55.42	6.06

Continued on next page

– *continued from previous page*

VARIABLE	N	Mean	Stand.Dev.	Median	IQR
Household typology					
Living alone	13,471	46.17	11.48	50.50	18.33
Childless couple	21,689	46.86	10.66	50.92	15.60
Couple with child.	57,303	52.05	7.87	55.26	6.04
Other typology	13,381	49.30	10.41	54.30	12.00
Economic resources					
Insufficient	73,205	50.72	8.98	54.84	8.16
Good	32,639	48.02	10.83	52.71	14.84
Urban degree					
Low	25,349	49.23	10.06	53.94	11.63
Intermediate	43,989	49.97	9.60	54.32	9.93
High	36,506	50.25	9.44	54.73	9.50
Depression					
No	99,800	50.41	9.26	54.78	8.93
Yes	6,044	41.25	11.87	40.73	20.91
Disease					
No health problem	72,245	52.39	7.55	55.30	5.38
Illness	18,694	47.88	9.91	51.28	13.61
Acute disease	14,905	40.27	11.67	39.94	20.19

*Coefficient not significant: *** "p¿0.05"; ** "p¿0.01"; * "p¿0.001"*

Appendix B

SF-36 and SF-12 description

The SF-36 questionnaire is a generic measure instrument of health status and one of the most widely used. Was constructed to satisfy psychometric standards necessary for comparisons. Accordingly , this questionnaire has proven useful in comparing general and specific populations, estimating the relative burden of different diseases or differentiating the health benefits produced by a wide range of different treatments. SF-12 was developed as an alternative to the SF-36 for the purpose of monitoring large samples of general and patient populations and in response to need for shorter health survey measures. The SF-12 is a subset of items from the SF-36. Like the SF-36, the SF-12 measures two broad health status domains: Physical well- being (the Physical Component Summary –PCS) and psychological well-being (the Mental Component Summary – MCS). The SF-12 scoring protocol based on a complicated algorithms. Although the short length of this questionnaire involves limit in sub-domains assess. In fact, the eight health status sub-domains which compose SF-36, cannot be analyzed separately in SF-12 questionnaire. The eight sub-domains of the SF-36 are:

a) Physical Functioning (PF), Role-Physical (RP), Bodily Pain (BP), General Helath (GH) which contribute to PCS score;

b) Vitality (VT), Social Functioning (SF), Role-Emotional (RE), Mental Health (MH) which contribute to MCS score.

Istat survey contains the SF-12 questionnaire, plus questions involved in scoring of VT and MH indexes. These questions are listed below separately for each index. Note: questions are not comparable with English version questionnaires being a free translation of the Italian version questions.

Questions forming SF-12 questionnaire , Vitality (VT) and Mental Health (MH) indexes, Disease variable

- Physical Component Summary (PCS) questions:

  1. In general, how is your health?
     - Very well
     - Well
     - Fair
     - Bad
     - Very bad
  2. Is now your health limiting you in performing activities of moderate physical effort (such as moving a table, use a vacuum cleaner, bowling or take a bike ride, etc.)?
     - Yes, its limits a lot
     - Yes, its limits a little
     - No, it does not limit at all
  3. Is now your health limiting you in climbing several flights of stairs?
     - Yes, its limits a lot
     - Yes, its limits a little
     - No, it does not limit at all
  4. During the past 4 weeks, have you been less efficient than you would like in your work or other regular daily activities as a result of your physical health?
     - No
     - Yes
  5. During the past 4 weeks, have you had to limit some kind of work or other activities as a result of your physical health?
     - No
     - Yes

6. During the past 4 weeks, how much has pain interfered with your normal work (including both indoor and outdoor)?
    - Not at all
    - A little bit
    - Moderately
    - Quite a bit
    - A lot

- Mental Component Summary (MCS) questions:

1. During the past 4 weeks, have you been less efficient than you would like in your work or other regular daily activities as a result of any emotional problems (such as feeling depressed or anxious)?
    - No
    - Yes
2. During the past 4 weeks, have you had a loss of concentration on work or other daily activities as a result of your emotional problems (such as feeling depressed or anxious)?
    - No
    - Yes
3. During the past 4 weeks, how often have you felt calm and/or peaceful?[1]
    - All of the time
    - Nearly always
    - For a long time
    - A little of the time
    - Hardly ever
    - Never

[1]Question also included in MH index.

4. How often during the past 4 weeks did you feel full of energy?[2]
    - All of the time
    - Nearly always
    - For a long time
    - A little of the time
    - Hardly ever
    - Never
5. How often during the past 4 weeks did you feel downhearted and sad?[3]
    - All of the time
    - Nearly always
    - For a long time
    - A little of the time
    - Hardly ever
    - Never

- Vitality (VT)(Well-being in this work) questions:

1. How often during the past 4 weeks did you feel full of energy?[4]
    - All of the time
    - Nearly always
    - For a long time
    - A little of the time
    - Hardly ever
    - Never
2. During the past 4 weeks, how often did you feel lively and brilliant?
    - All of the time
    - Nearly always
    - For a long time
    - A little of the time
    - Hardly ever
    - Never

[2]Question also included in VT index.
[3]Question also included in MH index.
[4]Question also included in MCS index.

3. During the past 4 weeks, how often did you feel worn out?
    - All of the time
    - Nearly always
    - For a long time
    - A little of the time
    - Hardly ever
    - Never
4. During the past 4 weeks, how often did you feel tired?
    - All of the time
    - Nearly always
    - For a long time
    - A little of the time
    - Hardly ever
    - Never

- Mental Health (MH) questions:

1. During the past 4 weeks, how often have you felt calm and/or peaceful?[5]
    - All of the time
    - Nearly always
    - For a long time
    - A little of the time
    - Hardly ever
    - Never
2. How often during the past 4 weeks did you feel downhearted and sad?[6]
    - All of the time
    - Nearly always
    - For a long time
    - A little of the time
    - Hardly ever
    - Never

[5]Question also included in MCS index.
[6]Question also included in MCS index

3. During the past 4 weeks, how often did you feel very restless?
   - All of the time
   - Nearly always
   - For a long time
   - A little of the time
   - Hardly ever
   - Never
4. During the past 4 weeks, how often have you felt so down in the dumps that nothing could cheer you up?
   - All of the time
   - Nearly always
   - For a long time
   - A little of the time
   - Hardly ever
   - Never
5. During the past 4 weeks, how often did you feel happy?
   - All of the time
   - Nearly always
   - For a long time
   - A little of the time
   - Hardly ever
   - Never

- Disease variable questions:

1. In the last 4 weeks have you had any illness or health problem?
   We are interested in all causes which may have disrupted your health during this period, both serious (such as pneumonia, appendicitis, ...) and less serious (such as cold, headache, cough, toothache, intestinal disorders, rheumatism, etc.).
   Any chronic disease should be considered, even if arising before the four weeks provided that they have given rise to health problems during these 4 weeks
   - No
   - Yes, please specify

2. In addition to what may be stated in response to the previous question, in the past 4 weeks, your health was compromised by wounds, fractures, bruises, dislocations, sprains, burns, or other problems due to traumas, poisoning, suffocation, etc.?
   - No
   - Yes
3. Diseases or disorders mentioned led to a limitation of your usual activities (at home, at school, in leisure time, at work, etc ...) during the last 4 weeks?
   - No
   - Yes, for how many days?

Bibliography

Asparouhov, T. (2006). "General multi-level modeling with sampling weights". In: *Communications in Statistics—Theory and Methods* 35.3, pp. 439–460.

Cade, B.S. and B.R. Noon (2003). "A gentle introduction to quantile regression for ecologists". In: *Frontiers in Ecology and the Environment* 1.8, pp. 412–420.

Carle, A.C. (2009). "Fitting multilevel models in complex survey data with design weights: Recommendations". In: *BMC Medical Research Methodology* 9.1, p. 49.

Cherepanov, D. et al. (2010). "Gender differences in health-related quality-of-life are partly explained by sociodemographic and socioeconomic variation between adult men and women in the US: evidence from four US nationally representative data sets". In: *Quality of life research* 19.8, pp. 1115–1124.

Costa, G. et al. (2003). "Individual and contextual determinants of inequalities in health: the Italian case". In: *International journal of health services* 33.4, pp. 635–667.

Falorsi, P.D. and Falorsi S. (1995a). "Il metodo di stima generalizzato per le indagini sulle famiglie e sulle imprese". In: *Rapporto di ricerca* 13.

— (1995b). "Il metodo di stima generalizzato per le indagini sulle famiglie e sulle imprese". In: *Rapporto di ricerca* 13.

Fontaine, KR and I. Barofsky (2001). "Obesity and health-related quality of life". In: *Obesity reviews* 2.3, pp. 173–182.

Gelman, A. and J. Hill (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Geraci, M. (2012). "lqmm: Linear Quantile Mixed Models". In:

Geraci, M. and M. Bottai (2007). "Quantile regression for longitudinal data using the asymmetric Laplace distribution". In: *Biostatistics* 8.1, pp. 140–154.

— (2011). "Linear Quantile Mixed Models". In:

Geraci, M. and M.M. Geraci (2012). "Package 'lqmm'". In:

Goldstein, H. (2011). *Multilevel statistical models*. Vol. 922. Wiley.

Goldstein, H. and M.J.R. Healy (1995). "The graphical presentation of a collection of means". In: *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, pp. 175–177.

Graubard, B.I. and E.L. Korn (1996). "Modelling the sampling design in the analysis of health surveys". In: *Statistical Methods in Medical Research* 5.3, pp. 263–281.

Grilli, L. and M. Pratesi (2004). "Weighted estimation in multilevel ordinal and binary models in the presence of informative sampling designs". In: *Survey methodology* 30.1, pp. 93–104.

Hox, J.J. (2010). *Multilevel analysis: Techniques and applications*. Taylor & Francis.

Istat (2005a). "Condizioni di salute, fattori di rischio e ricorso ai servizi sanitari". In: *Statustiche in breve*.

— (2005b). "Nota metodologica dell'Indagine "Condizioni di salute, fattori di rischio e ricorso ai servizi sanitari—Anno 2005". In:

Koenker, R. (2005). *Quantile regression*. Vol. 38. Cambridge university press.

Koenker, R. and G. Bassett Jr (1978). "Regression quantiles". In: *Econometrica: journal of the Econometric Society*, pp. 33–50.

Koenker, R. and K. Hallock (2001). "Quantile regression: An introduction". In: *Journal of Economic Perspectives* 15.4, pp. 43–56.

Koenker, R. and J.A.F. Machado (1999). "Goodness of fit and related inference processes for quantile regression". In: *Journal of the american statistical association* 94.448, pp. 1296–1310.

König, H.H. et al. (2010). "Health status of the advanced elderly in six European countries: results from a representative survey using EQ-5D and SF-12". In: *Health and quality of life outcomes* 8.1, pp. 1–11.

Mackenbach, J.P. et al. (2008). "Socioeconomic inequalities in health in 22 European countries". In: *New England Journal of Medicine* 358.23, pp. 2468–2481.

Olsen, K.M. and S.A. Dahl (2007). "Health differences between European countries". In: *Social science & medicine* 64.8, pp. 1665–1678.

Pfeffermann, D. et al. (1998). "Weighting for unequal selection probabilities in multilevel models". In: *Journal of the Royal Statistical Society: series B (statistical methodology)* 60.1, pp. 23–40.

Pirani, E. and S. Salvini (2012). "Place of living and health inequality: a study for elderly Italians". In: *Statistical Methods & Applications*, pp. 1–16.

Rabe-Hesketh, S. and A. Skrondal (2006). "Multilevel modelling of complex survey data". In: *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 169.4, pp. 805–827.

— (2012). *Multilevel and longitudinal modeling using Stata*. STATA press.

Raudenbush, S.W. (2008). "Many small groups." In:

Roux, A.V.D. (2001). "Investigating neighborhood and area effects on health". In: *Journal Information* 91.11.

Smith, G.D., M. Bartley, and D. Blane (1990). "The Black report on socioeconomic inequalities in health 10 years on." In: *BMJ: British Medical Journal* 301.6748, p. 373.

Snijders, T.A.B.A.B. and R.J. Bosker (2011). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Sage Publications Limited.

Taft, C., J. Karlsson, and M. Sullivan (2001). "Do SF-36 summary component scores accurately summarize subscale scores?" In: *Quality of Life Research* 10.5, pp. 395–404.

Tian, M. and G. Chen (2006). "Hierarchical linear regression models for conditional quantiles". In: *Science in China Series A: Mathematics* 49.12, pp. 1800–1815.

Ware, J.E. and B. Gandek (1998). "Overview of the SF-36 health survey and the international quality of life assessment (IQOLA) project". In: *Journal of clinical epidemiology* 51.11, pp. 903–912.

West, B., K.B. Welch, and A.T. Galecki (2006). *Linear mixed models: a practical guide using statistical software*. Chapman & Hall/CRC.

Yu, K., Z. Lu, and J. Stander (2003). "Quantile regression: applications and current research areas". In: *Journal of the Royal Statistical Society: Series D (The Statistician)* 52.3, pp. 331–350.