View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by City Research Online

Cubi-Molla, P. & Herrero, C. (2012). Quality of life lost due to non-fatal road traffic injuries. Health Economics, 21(5), pp. 528-550. doi: 10.1002/hec.1729

CITY UNIVERSITY LONDON

EST 1894

City Research Online

Original citation: Cubi-Molla, P. & Herrero, C. (2012). Quality of life lost due to non-fatal road traffic injuries. Health Economics, 21(5), pp. 528-550. doi: 10.1002/hec.1729

Permanent City Research Online URL: http://openaccess.city.ac.uk/5371/

Copyright & reuse

City University London has developed City Research Online so that its users may access the research outputs of City University London's staff. Copyright © and Moral Rights for this paper are retained by the individual author(s) and/ or other copyright holders. All material in City Research Online is checked for eligibility for copyright before being made available in the live archive. URLs from City Research Online may be freely distributed and linked to from other web pages.

Versions of research

The version in City Research Online may differ from the final published version. Users are advised to check the Permanent City Research Online URL above for the status of the paper.

Enquiries

If you have any enquiries about any aspect of City Research Online, or if you wish to make contact with the author(s) of this paper, please email the team at publications@city.ac.uk.

Quality of Life Lost Due to Non-Fatal Road Traffic Injuries

Patricia Cubí-Mollá* — City University London

Carmen Herrero† — University of Alicante and IVIE

January 10, 2011

Abstract

The objective of this paper is to evaluate the effect of a non-fatal road crash on the health-related quality of life of injured people. A new approach based on the cardinalization of categorical Self-Assessed Health valuations, is suggested. Health losses have been estimated by using different Time Trade-off and Visual Analogue Scale tariffs, in order to assess the robustness of the results. The methodology is based on the existing literature about treatment effects. Our main contribution focuses on evaluating the loss of health up to one year after the non-fatal accident, for those who are non-institutionalized, which aids the appropriate estimation of the aggregated health losses in quality-of-life terms.

Keywords: Health-related quality of life; Health measurement; Road crashes; Scaling Methods.

* *Corresponding author:* Patricia Cubí-Mollá, Department of Economics, Social Sciences D309. City University. Northampton Square, London EC1V 0HB Phone: +44 (0) 207 040 4543. Email: p.cubi-molla@city.ac.uk.

†Departamento de Fundamentos de Análisis Económico. Facultad de Económicas. Universidad de Alicante. 03071 Alicante; SPAIN. Phones: office: 34-96590-3618. Secretary: 34-96590-3563. Fax: 34-96590-3898. e-mail: carmen.herrero@ua.es

1 Introduction

We aim at estimating the loss of health following a non-fatal road crash. The methodology is based on the definition of comparison groups, by using the existing literature regarding treatment effects. The main contribution of this paper is the evaluation of health losses due to injuries in terms of quality of life. Moreover, this paper applies a novel method for scaling categorical health measures, a powerful tool in health-related analysis.

The selection of the topic "road crashes" is not pointless. In 2001, injuries represented 12% of the global burden of disease (WHO, 2001). The category of injuries worldwide is dominated by those incurred in road crashes. In 2004, over 50% of deaths caused by road crashes were associated to young adults in the age range of 15–44 years, and traffic injuries were the second-leading cause of death worldwide among both children aged 5–14 years, and young people aged 15–29 years (WHO, 2004). In addition, road crashes are expected to be the main origin of the projected 40% increase in global deaths resulting from injury between 2002 and 2030 (WHO, 2007).

The effectiveness of policies aimed at limiting the magnitude of the problem should be estimated carefully, allowing for the different outcomes they could yield: a reduction of the number of crashes, fatalities, and severity of the nonfatal injuries. In this context, Bishai et al. (2006) demonstrated that observed patterns in rich countries show only a decline in fatalities, but no decline of crashes or injuries. In this respect, focusing on the impact of health and well-being of injuries, and the sequelae of events in people surviving road crashes, is extremely important. Nonetheless, few studies have dealt with this topic, in part due to the computational difficulties of properly addressing the issue. This study contributes to the computation of the benefits of road crashes prevention, and could serve as an input into a cost-benefit analysis.

The actual loss of health following road traffic injuries (*RTIs,* hereafter) equals the difference between the values associated to the post-injury health state and the potential health state (under the counterfactual scenario in which the accident does not happen).

The so-called *potential* health status is always unidentified and thus some assumptions must be made in order to approximate it. Regarding the *post-injury* health status, the ideal is to estimate the chronic sequelae that a traffic crash can have on the affected, and to evaluate the impact of these sequelae in their daily living. However, this is a challenging task, since it is hard to obtain a complete set of data that comprises all the required information. This setback is particularly relevant at evaluating the medium or long term health effects for those who have been seriously injured by a road crash, and who have been discharged from hospitals or similar health care institutions. The impracticality of performing a follow-up for these affected individuals makes it almost impossible to document any future health complications that could be indirectly caused by the *RTIs* suffered in the past. If such information is omitted however, we run the risk of underestimating the actual toll of road crashes on society.

The literature proposes few solutions to deal with this bias. Normally, the post injury health status is obtained from sources as police records, Hospital Discharge Registers, or databases from health care institutions. The nature of these databases is essentially linked to the estimation of the seriousness of the injuries, rather than focusing on the impact of the injury over the general health state of the individual. Particularly relevant is the methodology developed at estimating the potential health state. The earliest studies in this area directly assume the potential health state as that of being in "perfect health" (Sullivan et al., 2003; Redelmeier and Weinstein, 1999). More recently, a comparison group is taken as a proxy of the potential health state of the victim (Nyman et al, 2008). This methodology is mainly based on the use of population norms that provide some benchmark against which to compare post-injury outcomes. The common norms are stated in terms of changed health baselines for men/women and different age-groups.

We think that these methodologies can be improved. Notice that the assumptions stated above treat road crashes either as fortuitous events or completely based on few observable factors. However, data show that people affected by *RTIs* can be neither considered as randomly selected, nor as a perfectly targeted population. We should think about the existence of unobservable factors, such as the degree of risk aversion, the driving ability, and so on, that could affect the health state as well as the probability of having a road crash. Previous literature does not control for the possible existence of a selection bias in the results. Moreover, they fail to express the result in preference-based metrics, so that it cannot be extended to a policy or social framework. Nyman et al. (2008) make a first attempt to outperform the previous studies. Nonethless, there are some weak points in their approach: on the one hand, they use scaling methods with a lack of theoretical support; on the other hand, they consider road crashes as purely stochastic occurences. Thus, there exists room for improving the estimation of the impact of *RTIs* and sequelae in terms of quality of life (QoL, hereafter) lost.

In this work we estimate the loss of health (in *QoL* terms) that is due to a road crash, for those who suffer from *RTIs*, up to one year after the crash. The analysis is performed for non-institutionalized individuals. The methodology is based on the definition of comparison groups, by using the existing literature concerning treatment effects. We analyze whether the introduction of variables that could capture risk aversion modifies the results substantially, with respect to the outcomes derived by treating road crashes as fortuitous events. If so, the results could suggest the existence of unobserved factors that had not been captured by the controls introduced in the model.

In Section 2 the methodology is described, starting with the estimation of the direct loss of health, and following with the cardinalization of categorical variables. The data and variables used for the analysis will also be described in this section. In Section 3 we present the main results, and several robustness checks. Section 4 concludes.

2 Methodology

Two different steps are required for our analysis. The first one concerns the evaluation of health losses. We target the population of interest, and we establish the methodology and the identification strategy. The second step regards the measurement of health states. We discuss the selection of a proper metric, and the procedure for deriving the selected

QoL measure. The data and the variable definitions will be described in this section.

2.1 Evaluation of health losses

2.1.1 Target population

The paper focuses on evaluating the medium term health effects for those who have been injured by a road crash, up to one year after the crash. The analysis does not consider those individuals with more severe consequences from road crashes: either the institutionalized individuals or those non institutionalized who are not able to report their Self-Assessed Health (*SAH*, hereafter). The effect of *RTIs* on that population should be addressed by a different procedure, and by using alternative databases as could be hospital registers or institution-specific records.

The population of interest for the current analysis embraces those injured by a non-fatal road crash, who have already been discharged from hospitals or similar health care institutions. The difficulty of performing a follow-up for these affected individuals makes the estimation more challenging: it is quite complex to document any future health complications that could be indirectly caused by the *RTIs* suffered in the past. Besides, the data collection of those severely injured by *RTIs* is quite limited, so that our short-term objective consists in obtaining the finest estimation of health losses but with the restrain of the availability of data. Most likely the results are conservative estimations that indicate minimum thresholds of health related QoL lost.

We are interested in estimating the health effects on the following target groups:

[SV] Severely injured by a road crash.

[SL] Slightly injured by a road crash.

[INJ] Injured (slightly or severely) by a road crash.

Distinguishing between the three defined groups for the target population is not pointless. On one hand, the composition of the groups may result from different observable characteristics. For example, severe injuries may be attached to risky behavior towards driving, or the misuse of seat belts, what could be correlated to age, gender or experience. Considering these groups independently may provide more accurate results. On the other hand, it is worthy to mention that most of the international and country-level accident databases (e.g. CARE, IRTAD or PENDANT, to cite but a few) stratify the injured ones by the outcome of the crash and/or the severity of the injuries: dead, seriously injured, slightly injured, not injured. Therefore, the health effects to be estimated could be easily applied to different frameworks. Still, the analysis will be also performed over those who had a road crash, regardless the severity of the injuries. The results obtained form *[INJ]* will represent the overall effect of *RTIs*, and the estimates may be useful in contexts where the level of severity of *RTIs* is either not available or reliable.

2.1.2 Analysis of health losses

The analysis of health losses due to *RTIs* can be performed by using the *treatment effects* literature. In this context, the "treatment" is interpreted as the occurrence of a road crash that causes injuries to the affected individuals. In this paper we focus on the average loss of health as a result of a road crash, for those who had a non-fatal accident (average treatment effect on the treated or $ATET$)

Some notation is useful at this point. Let D_i indicate whether individual i had a road crash ($D_i = 1$) or not ($D_i = 0$). Let H_i represent the health status for individual i. This health state is measured after the road crash takes place. We define H^0 as the outcome that individual i would attain if he had not suffered from *RTIs*. Equivalently, H^1 is the outcome that individual i would realize if he had suffered the road crash. Thus, $H_i = D_i H_i^1 + (1 - D_i) H_i^0$ is the observed health status of the individual i. Let H be the set of all observed health statuses.

The $ATET$ cannot be identified using observational data since H^0 is only observed for those targeted by $D_i = 0$. A suitable solution is to approximate the average health state that injured people would have had in the absence of the road crash (potential health status) by the average health state observed in a comparable group of people that have not had an accident. As we mentioned in the Introduction, data show that traffic crashes are not completely random, but they are more likely to happen to people with particular traits (for instance men aged 15-29). Therefore the average health of injured (*affected group*, hereafter) and non-injured (*comparison group*, hereafter) individuals cannot be unconditionally compared. Thus, the validity of this approximation is likely to be higher once differences in the distribution of observed individual characteristics are controlled for. Let Z be a vector including information relative to individual i that is a priori thought to influence his probability of suffering a road crash.

We follow the procedure developed in Abadie (2005) for the conventional difference-in-differences estimator, now adapted to the situation where only cross-section data for the post-treatment period are available. The $ATET$ can be written as follows (the usual identification assumptions apply):

$$ATET = E\left[H^1 - H^0 | D = 1\right] = E\left[\frac{H}{P\left(D=1\right)} \cdot \frac{D - P(D=1|Z)}{1 - P(D=1|Z)}\right] \tag{1}$$

Equation (1) suggests a simple two-step method to estimate the $ATET$. First, the conditional probabilities are estimated using a logit model and the fitted values of $P\left(D=1|Z\right)$ are calculated for each individual in the sample. Second, the fitted values are plugged into the sample analog of Equation (1). Then, a simple weighted average of the outcome variable recovers the $ATET$. Finally, the asymptotic variance of the estimator is also calculated, following the procedure developed in Abadie (2005).

For the results derived from such methodology to be correct, it is necessary to consider one of the following assumptions: 1) there are no unobservable factors affecting both the outcome and the probability of having a crash; 2) if unobservable factors exist, these can be captured by the observable ones (e.g. risk-loving behavior is usually related to smokers); 3) if unobservable factors exist and are not reflected by the observables, its overall average impact is equal for both the affected and the comparison group.

People affected by *RTIs* can be neither considered as randomly selected, nor as a perfectly targeted population. The existence of some random component cannot be denied, mainly related to the occurrence or non-occurrence of a road crash, rather than the seriousness of the injures. For instance, being involved into a crash caused by a different individual, or an unexpected puncture on the road. However, many factors influence both the existence of *RTIs* and the seriousness of the injures. For instance, wearing seat-belts, airbags, driving carefully, not being drunk or using the pedestrian crossing, are factors that can be controlled by the individual. In fact, data show that the group of people that have a road crash includes higher proportion of male, aged 16-35, have unhealthy habits as smoking or consumption of alcoholic drinks, among other features. These characteristics, that may affect the health status, the probability of being injured by a road crash, as well as the severity of the injury, are observable.

There are also unobservable factors, as the degree of risk aversion or the driving ability that could also affect health as well as accident and injury causation. In order to ensure that the results provide an accurate estimation, the propensity score will be computed under different sets of controls. In a first set of variables (*xvars1*) only age-gender controls will be included. The second group of factors (*xvars2*) will add new observable characteristics of the individual. Finally, a third group (*xvars3*) will add controls that try to capture behavioral and psychological characteristics of the individuals, as proxies for the unobservable factors that could affect the probability of belonging to *[SV]*, *[SL]* or in general to *[INJ]*.

We will analyze whether the introduction of variables that could capture risk aversion modifies the *ATET* substantially, with respect to the results derived by treating road crashes as fortuitous events. If so, the results could suggest the existence of unobserved factors that had not been captured by the controls introduced in the model.

2.1.3 Main data and variable definitions for the evaluation of health losses

For estimating the impact of *RTIs* on population health, we use the survey about diseases, disabilities and health states (*Encuesta de Discapacidades, Deficiencias y Estados de Salud*), arranged by the Spanish National Institute of Statistics (*INE,* 1999). The survey includes 70,402 households (about 217,760 individuals), selected with a probability proportional to the size of each region. The survey is divided into two sections: Diseases and Disabilities Unit (*Módulo de Discapacidades y Deficiencias*), and Health Unit (*Módulo de Salud, MS* hereafter). We use data from the *MS* section. In that unit an individual in each household is randomly chosen - in total: 69,555 individuals; however, 840 observations from Ceuta and Melilla were dropped for computational feasibility.[1] For practical reasons, the analysis is performed over the population aged between 15 and 75. Observations with missing values are also dropped from the sample. The final sample size is 45,864 individuals. The interviewed is confronted with a battery of questions related to health habits, as well as demographic and socioeconomic information.

Two questions in *MS* have been selected to target those injured due to traffic accidents. The first question (F1) states as follows: "During the last 12 months, have you suffered from a traffic accident that has prevented you from performing any usual activity?". A total of 850 respondents report "Yes". The second question (F2) is "How has this traffic accident influenced in your daily life" From the individuals who give an affirmative answer to the first question, 297 answer "Seriously" or "Quite a lot", and 553 report "Slightly" in the second question.

The composition of the target groups is the following (note that the comparison groups selected can approximate better the potential health state of the affected individuals):

[SV] Affected group: 297 individuals answering "Seriously" or "Quite a lot" in F2. For the comparison group, those reporting "Slightly" in F2 are dropped from the sample.

[1] The particular geographical location of Ceuta and Melilla, as well as the area they cover and their population size, might result in different casualty and nature of traffic accidents. In addition, the observable characteristics of the residents of Ceuta y Melilla are, on average, different from those observed in the rest of the sample. Including them would make difficult to perform a direct comparison.

[SL] Affected group: 553 reporting "Slightly" in F2. For the comparison group, those reporting "Seriously" or "Quite a lot" in F2 are dropped from the sample.

[INJ] Affected group: 850 individuals answering "Yes" in F1.

For evaluating the propensity score we perform different logit models, in order to identify the nature of the selection bias. We must take special care for not including variables that could be themselves affected by treatment into the regression. The characteristics of the injured people are recorded up to one year after the accident, so that they could be reflecting the consequences of a road crash rather than the probability of suffering it. These sort of variables could introduce an additional problem, that is the endogeneity in the regression, what could modify the estimated effect of the treatment. Because of that, we dropped from the regression the individual characteristics that are likely to be a consequence rather than a factor related to the propensity to have a accident. For instance, the current labour status, number of hours of sleep, alcohol consumption and BMI, among others.

The following sets of controls have been included:

***xvars1*:** age-gender categories

***xvars2*:** resident location, educational level, household size, population size and logincome.

***xvars3*:** if the individual has suffered another sort of accident (not a road crash); if the individual has restricted his/her nocturnal outing during the last 12 months by fear of being robbed; if he/she has been a usual smoker one year ago.

The first variable to be used in *xvars3* tries to mark inattentive persons. The second one would capture the level of apprehension. Finally, smoking may be capturing the level of risk aversion.

Average characteristics for some key variables are given in **Table I.** A Mann-Whitney rank-sum test is used to compare the means of each variable for injured and non-injured (the null hypothesis is equality of means). For a start, severely injured individuals self-report lower health levels than non-injured ones, what is consistent with the hypothesis about the existence of chronic health losses following non-fatal road crashes. Those that were "slightly injured" report better health levels than the comparison group. However, no conclusion should be drawn from this preliminary analysis: the cohort embraces mainly men aged 15-34, whose potential health state could be even better than the reported one. In general, we observe that men are more likely to be seriously injured by a road crash than women. The table also suggests the existence of different features for those with severe *RTIs*, with respect to those slightly injured. For instance, the highest level of education completed differs mainly by the higher proportion in secondary studies, in contrast to a lower proportion of superior studies for those seriously injured.

(Table I around here)

The factors entered in *xvars3* (proxies for variables that are usually unobservable) seem to be significant not only for the occurrence of a road crash, but also for the outcome of the accident. The group of injured people, regardless of the seriousness of the *RTIs*, includes a higher proportion of heavy smokers. Nonetheless, only the group of those severely injured seems to contain a higher proportion of inattentive or apprehensive individuals. Thus, with the aim of obtaining a valid estimate of the *ATET*, it is necessary to control for these differences in the distribution of individual characteristics.

2.2 Measurement of quality of life

2.2.1 Selection of a quality of life measure

A wide variety of metrics are used to quantify the burden of illnesses and injuries to the population (an exhaustive description of these measures can be found in Seguí-Gomez and MacKenzie (2003), MacKenzie (2001) or Sturgis et al. (2001), among others). In general terms, we can talk about two different sort of measures, depending on the way of approaching the health status.

Measures in the first group focus on the impact of the injury over the general health state of the individual, developing a variety of indices or metrics that define "health". Measures as *Self-Assessed Health, Euroquol Time-Tradeoff tariff (EQ TTO tariff), Euroquol Visual Analogue Scale tariff (EQ VAS tariff)* or *Health Utility Index,* can be placed within such an approach. Metrics in the second group aim at estimating the seriousness of the injuries, either reflecting the degree of functional limitation of the injured individuals (*Functional Capacity Index*, *Disability weights*), or attending to the mortality risk or life threat (*Abbreviated Injury Scale*, *Injury Severity Score*, *ICD-9 Injury Severity Score*, *Anatomic Profile Score,* among others).

Scales belonging to the second group are the ones most commonly used to assess health losses due to injuries. However, several studies suggest that an individual's injury and acute psychological responses are strongly linked. Hence, both play important roles in determining *QoL* and disability outcomes (e.g. O'Donnell et al., 2005). Although measures of severity in the second group provide some understanding of the relative seriousness of injuries in terms of threat to life and resource utilization, they still fall short in measuring the long-term impact of non-fatal injuries on the person, his or her family, and the society at large. These considerations have challenged the field to move beyond counting injuries by severity alone to measuring their direct impact on health-related *QoL*.

In the present work we approach the problem from a *QoL* perspective, that is: we analyze the impact of non-fatal injuries on the *QoL* of the injured individuals, not only attending to the physical damage that the injury caused, but also contemplating the possible psychological consequences, as well as the potential impact on the well-being of those affected. To perform this exercise we use Spanish data. In order to check the robustness of the results, the analysis is performed by using different quality-related health state scores (*VAS tariff* and *TTO tariff),* that are obtained by applying the Spanish EQ-

5D index tariffs (see Badia et al., 1995 and Badia et al., 2001). Our analysis is performed by using the actual tariffs (that allow negative values, that is, health states worse than death). To simplify notation, we denote the outcomes as VAS and TTO.

2.2.2 Construction of the selected measure

The health scores are not directly available in the dataset *MS.* Thus, the health states will be derived from the respondent's assessment of her own health status. That piece of information about self-assessed health will be obtained from the categorical variable SAH : "In your opinion, how is your health in general?", where respondents must choose one of the following categories: "very good", "good", "fair", "bad" or "very bad". Since categorical measures of health are one of the most commonly used indicators in socioeconomic surveys, a wide variety of methods were developed with the aim of dealing with the cardinalization of ordinal health measures (e.g. Van Doorslaer and Wagstaff, 1994; Cutler and Richardson, 1997; Groot, 2000; Van Doorslaer and Jones, 2003). The interval regression model is shown to outperform other econometric approaches, in terms of validity and ability to mimic the distribution of scaling health measures (Van Doorslaer and Jones, 2003). However, it has been criticized for not admitting any skewness in the distribution of health (Cubí-Mollá, 2010), whereas it is well-known that the health of a general population sample has a very skewed distribution, with the great majority of respondents reporting their health in higher levels.

In this paper we apply a variation of the previous methodologies suggested by Cubí-Mollá (2010). This methodology combines the distribution of observed SAH (which comprises n categories) with external information on the distribution of a generic measure of ill-health h (defined as $h = 1 - y$, where y stands for the generic health measure). The aim is to construct a continuous standardized latent ill-health variable, h^*. Let us denote h_i^* the true, latent ill-health of the individual i in a range $(-\infty, 0]$, and assume that h_i^* has a lognormal distribution, for all i.

The methodology assumes that h_i^* can be represented by h_i in a $0 - 1$ scale, and the thresholds of the intervals determining SAH $(\eta_j, j = 0, 1, .., n)$ are obtained from external information and thus, are observable. The range of h_i and h_i^* is divided into n intervals, each one corresponding to a different value of SAH:

$$SAH_i = n - j + 1 \text{ iff } \eta_{j-1} < h_i \leq \eta_j, \; j = 1, 2, ..., n \tag{2}$$

where it is fixed that $\eta_0 = 0, \eta_n = \max(h),$[2] $\eta_j \leq \eta_{j+1}$, and h_i is assumed to be a linear function of a vector of factors X_i :

$$\log(h_i) = X_i \beta + u_i, \text{ with } u_i \sim N(0, \sigma^2) \tag{3}$$

The expression (2) represents the well-known ordered logit model, and (3) will allow us to use a nonparametric approach to estimate the thresholds of the model, by using the cumulative frequency of observations for each category of SAH to find the quantiles of the empirical distribution function for h. The setting of the thresholds allows us to identify the variance of the error term $\widehat{\sigma}^2$ and hence, the scale of h^* without having any scaling or identification problems (Cubí-Mollá, 2010).

Our aim is to estimate the average health valuation on a continuous 0-1 scale, for each individual by conditioning on X_i. Noticing that $\exp(u_i) \sim lognormal(0, \sigma^2)$, we obtain the expression:

$$H_i = E[h_i|x_i] \approx \exp\left(X_i\widehat{\beta}\right) \cdot \exp\left(\widehat{\sigma}^2/2\right) \tag{4}$$

where H_i captures the estimated average value of ill-health, ranging from 0 to 1, associated to the observable characteristics of individual i.

In order to evaluate the robustness of this methodology, the thresholds are determined in terms of different generic health measures obtained from external data. We use

[2] Since the continuous health measures are not rescaled, they may take values lower than 0, which implies ill-health values higher than 1.

TTO and *VAS tariffs* as the continuous self-assessed measures.

2.2.3 External data and variable definitions for the cardinalization of SAH

The required external information (i.e., $\eta_i, i = 1, 2, .., n-1$) is obtained from the Catalan health surveys *Enquesta de Salut de Catalunya* 2002 (*ESCA02* hereafter) and *Enquesta de Salut de Catalunya* 2006 (*ESCA06* hereafter), arranged by the Catalan government (*Generalitat de Catalunya*). A total of 8,400 individuals (in the former) and 18,126 individuals (in the latter) were selected for the surveys, which include different health measures as *VAS*, *EQ-5D* and *SAH*. From these variables, three cardinal health measures could be obtained: *VAS* (directly from the survey), *VAS tariff* and *TTO tariff* (estimated from *EQ-5D*). Both surveys include a battery of questions concerning the existence of *RTIs.* However, the limited number of injured persons is not sufficient to perform the analysis suggested in this paper.

In the *ESCA02* we dropped 1,837 observations from the sample: 19 observations because either *VAS* or *SAH* were not reported, 1748 observations related to individuals aged under 15 or over 75, and 66 proxy-respondent interviews (due to impairments). A total of $4,133$ observations ($3,896$ corresponding to individuals aged less than 15 or more than $75, 192$ proxy-respondent interviews and 45 observations that were considered untruthful by the interviewer) were dropped from the *ESCA06.*

Finally, some observations (3) presenting inconsistencies were discarded. Those have been detected based on the values provided by the variables *VAS* and *SAH.* Thus, several individuals reported "excellent" health or *VAS* close to 1, but negative values for the tariffs. Similarly, some individuals reported "bad" health or *VAS* close to 0, but tariff values close to 1. The analysis uses pooled individual data from both surveys (*ESCA02/06* hereafter).[3] The final sample size is $20,557$ individuals.

Different thresholds will be computed, conditional on observed characteristics in ESCA02

[3] Similar analysis have been performed over *ESCA02* and *ESCA06* separately. The results are very similar to those obtained in this paper.

/06. We will explore whether interval boundaries widely differ across demographic groups. If so, the model would be adapted to fit the possible response-category cut-point shifts, among different populations. Special attention will be provided to check whether road crashes influence the cut-points. If the interval boundaries are found to be different across those who had *RTIs* and those who had not, even controlling for additional factors as age and gender, that would support our hypothesis regarding the dissimilar composition of both groups. Furthermore, these thresholds may capture an effect of road crashes, and so the *QoL* lost is likely to be sensitive to making the interval boundaries *RTIs*-specific. The analysis will be also performed by using homogeneous thresholds, that is, not conditioning on individual characteristics.

It is important to notice that the SAH variable included in *ESCA02/06* is not identical to the SAH variable incorporated into *MS* (the dataset used for estimating the $ATET$). The dissimilarity lies in the five possible answers given to the respondents: the category "very bad" is not available in *ESCA02/06*, but "excellent" is incorporated. In order to define a single health index, the construction of SAH containing 4 categories is performed (the new variable will be called $SAH4$; the methodology explained in the previous section will apply for $n = 4$), following the approach adopted by several authors (e.g. Lindley and Lorgelly, 2007; Hernández-Quevedo et al., 2008). The collapsed categorizations are summarized in **Table II**.

(Table II about here)

For the interval regression, a wide range of factors are considered that can affect the self-valuation of the health state of an individual:[4]

(a) *Socioeconomic factors:* age, gender, marital status, education, labour (unemployed, disabled, retired, housekeeper, student, other), income, household size, residence location, population size, and citizenship.

[4]Several observations are dropped because of missing values in some of the regressors.

(b) *Health-related factors:* BMI (underweight, BMI < 18 / normal weight, $18 \leq$ BMI≤ 25 / overweight or obese, BMI> 25) existence of a chronic illness (bronchitis, allergy, epilepsy, diabetes, hypertension, heart injuries, cholesterol, cirrhosis, arthritis, ulcer, hernia, cardiovascular diseases, anaemias, nerves, migraines, menopause, other), existence of disorders (mental, visual, auditory, articulation, bones, nervous system, visceral, other), if the individual has had an accident during the last 12 months (serious road crash or other kinds of accident), if the individual is currently taking medicines.

(c) *Lifestyles:* if the person sleeps more than 8 hours, practices sports (working days / weekend), if the person is a usual smoker or a hard drinker.

3 Results

3.1 Interval thresholds

We explore whether interval boundaries widely differ across demographic groups. Pooled data from *ESCA02* and *ESCA06* are grouped by gender and age category; by the existence of (at least) one chronic illness; by the existence of disabilities, and by existence and severity degree of the *RTIs* (non-injured, slightly injured, severely injured). The small sample size of those injured by a road crash in *ESCA02/06* impedes the inclusion of additional factors.

Subjective thresholds are shown to be quite similar by disability status or existence of a chronic illness.[5] This pattern is also observed in different samples, by Van Doorslaer and Jones (2003). However, interval boundaries on the basis of age-gender groups and severity of *RTIs* present dissimilar cut-points. For instance, the threshold between the two lower categories of SAH is considerably higher for those injured than the threshold derived from the rest of subgroups, maybe with the exception of men and women aged

[5] The figures are not shown in this paper. The authors can provide the results upon request.

15-29. On the contrary, the thresholds between the higher categories of SAH are much lower than the rest. Different analysis have proved that this effect cannot be induced by the age-gender composition of the subsample. Thus, maybe these thresholds are capturing an effect of road crashes, and so the interval regression approach is likely to be sensitive to making the interval boundaries *RTIs*-specific. This response-category cut-point shift is taken into account at scaling the SAH answers in *MS.* The analysis is also performed by using age-gender groups and homogeneous thresholds. **Table III** illustrates the results regarding the VAS tariff (the thresholds by groups derived from alternative tariffs show the same pattern). The thresholds are presented in terms of ill-health.[6] The sample size for every subgroup is also provided, in order to assess the comparability of both databases.

(Table III about here)

Observe that when the actual health tariffs are used, the higher bound corresponds to one minus the minimum value of the tariff. Also, in the absence of *RTIs*, the boundaries are quite similar to the homogeneous thresholds. Thus, if the results obtained by using conditioned or unconditioned boundaries are rather different, this could highlight the importance of controlling for possible response-category cut-point shifts.

Observe that η_1 is very close to 0 for the *VAS tariff* (also for the *TTO tariff*). This is a direct consequence of the "ceiling effect" of these scores: a value of health equal to 1 is assigned to a great percentage of the population (63.7%). The interpolation used for estimating the thresholds avoids that $\eta_1 = 0$ for these metrics.

The values should be interpreted as follows: for instance, referring to VAS, and using the homogeneous thresholds: an individual who reports the best category of health is assumed to have a VAS level that belongs to the interval $[0, 0.075]$ in terms of ill-health, or $[0.925, 1]$ in terms of health. Similarly, the values for the remaining $SAH4$ categories are $(0.075, 0.237]$ for the "good" category, $(0.237, 0.586]$ for the "fair" category

[6] In order to assess the robustness of this assertion, the results have also been derived from all the established groups of thresholds. Since the $ATET$ derived from these conditional thresholds do not differ substantially from the $ATET$ derived from the homogeneous thresholds, the results have not been reported in the paper. The authors can provide the results upon request.

and $(0.586, 1.076]$ for the "bad" categories.

3.2 Cardinal health measurement

The specification for intervals is implemented into parallel regression models. The characteristics of the regressors as well as the parameter estimates of the interval regression model are found in **Appendix.** The health status of each individual is controlled for a wide range of socioeconomic variables, and most of the coefficients are significant (CI 95%). The McKelvey and Zavoina[7] pseudo-R^2 is computed for each model, and rounds 0.48, indicating that these predictors account for approximately 48% of the variability in the latent outcome variable. On average, 65% of the estimated health tariffs lay into the correct interval (settled by the reported answer to the SAH question). A Regression Error Specification Test (RESET test)[8] has been applied to each interval and logit regression model, and none of them shows evidence of misspecification.

It is important to remark that the estimated value of health is highly linked to the self-perception of health status, rather than the actual health status per se. A positive coefficient means that an individual has a higher value of latent ill-health and is more likely to report a lower category of self-assessed health. The regressors have been built so that the reference individual is a Spanish woman aged 25-35, who lives in La Rioja, single, employed, completed higher education, who did not suffer an injury during the last 12 months, no chronic illness, non-smoker, sleeps less than 8 hours per day, does not make any physical exercise and has a proper BMI (does not show underweight or obesity).[9]

As expected, the ill-health decreases with income, level of education, absence of chronic

[7]The McKelvey and Zavoina pseudo-R^2 is an attempt to measure model fit as the proportion of variance accounted for: $\text{var}(h)/[\text{var}(h) + \text{var}(u)]$.

[8]RESET test is popular means of diagnostic for correctness of functional form. I test: $H_0 : \gamma = 0$ against the alternative $H_1 : \gamma \neq 0$, in $\log(h_i) = X_i\beta + y_i\gamma + error$, where y_i is generated by taking powers of the predicted values $\widehat{\log(h_i)}$ in (4). A failure to reject H_0 says the test has not been able to detect any misspecification.

[9]In order to allow for some variability in the effect of a road crash in health, several interactions (e.g. with gender, age, education, labor status) were introduced in the preliminar models; since no interaction was significant, and they did not modify the results, they were finally dropped.

illness, and absence of injuries or limitations. Besides, ill-health is decreasing with sleeping more than 8 hours per day, exercising, living in cities with higher population. Students are healthier than any other employment condition, married and widowers are more likely to report a higher category of *SAH* (and thus higher value of true health) than single people. The results also provide evidence about the decline of *QoL* as age increases.

3.3 Estimation of health losses

The coefficients and standard errors corresponding to the logit models are included in **Appendix**. For every group, the propensity score is larger (CI 95%) for men aged between 15 and 24, with secondary studies. Age-gender categories are found to be essential control variables for establishing the propensity score. Remark that the coefficients of the behavioral factors are significantly different from zero for those with severe *RTIs*. In particular smoking is statistically significant even for those slightly injured.

It interesting to stress the main objective of the logit regression. From equation (1) we can write:

$$ATET = E_{comp}\left[w \cdot H\right] - E_{aff}\left[H\right]$$

where $E_{aff}\left[\cdot\right] = E\left[\cdot|D=1\right], E_{comp}\left[\cdot\right] = E\left[\cdot|D=0\right]$ and $w = \frac{P(D=1|Z)}{1-P(D=1|Z)} \cdot \frac{P(D=0)}{P(D=1)}$.

Thus, the logit model balances the samples of comparison and affected groups, by introducing a weight for each individual in the comparison group.

The average health effect under "selection of observables" is estimated in terms of decrease in health. The standard errors and confidence intervals are computed by bootstrapping. The number of iterations is 1,500, and the bias-corrected estimate has been considered, assuming that standard errors are normally distributed. It can be observed that the effects differ depending on the metric in which the ratio is expressed. The results of the estimation and the confidence interval are illustrated in **Figure 1 and Figure 2**. For a better comprehension, the results are expressed in terms of decrease in health, instead of increase in bad health. On average terms, *RTIs* cause a decrease in health from 0.028 (*TTO tariff*, with the global thresholds and using *xvars3*) to 0.047 (*VAS tariff*, being the thresholds obtained by *RTIs* and the propensity score from *xvars1).* For every health measure, the confidence interval embraces values strictly negative, which indicates the existence of a reduction in *QoL* for those injured by a road traffic crash.

The *QoL* lost is much higher for those severely injured. It raises up to 0.123 points of the VAS tariff (being the thresholds obtained by *RTIs* and the propensity score from *xvars1*). The results also suggest the existence of a minor decrease in health for those slightly injured, but the estimate is not statistically significant.

(Figure 1 about here)

(Figure 2 about here)

The differences between simple averages of health for affected and comparison group have been computed. The results differ from the estimated *ATET*, what supports the validity of the hypothesis about the existence of selection on observables. In order to highlight the real impact of the total loss of health on individuals' health state, we compute the following rate:

$$\Delta H = \frac{E_{aff}\left[H\right] - E_{comp}\left[w \cdot H\right]}{E_{comp}\left[w \cdot H\right]} = \frac{ATET}{E_{aff}\left[H\right] - ATET}$$

ΔH indicates the proportion of health that on average individuals have lost due to a road crash, with respect to the health state that, on average, individuals would have had if the accident had not happened, estimated by using adjusted comparison groups. The confidence interval of ΔH is also re-scaled. The results are also shown in **Tables IV-VI.**

(Table IV about here)

(Table V about here)

(Table VI about here)

RTIs involve a decline in health from 3.19% (*TTO tariff*, on the basis of homogeneous thresholds and being the comparison group averages adjusted by *xvars3*) to 5.61% (*VAS tariff*, thresholds by *RTIs*, and using *xvars1).* Health losses for those who had severe *RTIs* are estimated to be between 8.17% and 14.90%. Health losses following slight injuries, up to one year after the road crash, are not found to be statistically significant.

For every health measure, the *ATET* derived from the balanced data is considerably higher (in absolute terms) than the effect estimated by considering that road crashes are completely random. Thus, to control for a selection bias is a relevant factor in the analysis. The results also suggest that the potential health state of the injured is, on average, better than the health status of those who have not had a road crash. Such differences in *QoL* are barely reduced by introducing controls to capture observable and unobservable factors that could affect the probability of *RTIs* (*xvars2* and *xvars3*, respectively). In fact, the correlation between both results (random and non-random treatment) remains almost constant, even though the coefficient of the additional variables are significantly different from zero.

It is remarkable how the estimates change depending on the measurement of health indices. At a first step, the definition of the thresholds by *RTIs* does not imply a monotonous cut-point shift. However, the use of different thresholds for scaling self-assessed health for severely, slightly and non-injured individuals affects the estimation of the *ATET*. By deriving homogeneous thresholds, we could be excluding some psychological component that may be linked to the health self-perception among the affected, which seems to lead to a lower *ATET*. The consideration of different thresholds by *RTIs* can be interpreted as a way to control for another source of selection bias.

4 Discussion

The fact that road crashes represent an alarming threat to health has been reported by a majority of the studies that deal with *RTIs*, causes of death or the evaluation of the burden of diseases. The application of different policies aimed at reducing the magnitude of the problem is essential. The effectiveness of these policies should be estimated carefully, allowing for making a distinction among the different outcomes they could yield: a reduction of the number of crashes, fatalities and severity of the non-fatal *RTIs*. In order to pursue this task, and to allow a comparison among analysis of different interventions, we should express the total toll of deaths, injuries and sequelae derived from traffic accidents in a simple metric, that could estimate the total loss of health that could be avoided.

To our knowledge few studies evaluate health losses due to non-fatal *RTIs* in *QoL* terms. Redelmeier and Weinstein (1999) estimate that *RTIs* report a loss of health of 0.127 *QoL*. Sullivan et al. in 2003 estimated the morbidity caused by *RTIs* in 0.356. These authors consider the baseline *QoL* for calculating the decrement due to injury as 1.00 (this is, non-injured are always in perfect health). Also, Sullivan et al. (2003) do not express the result in preference-based metrics, so that it cannot be extended to a policy or social framework. More recently, Nyman et al. (2008) computes the *QoL* decrements as 0.015 for those who say they are fully recovered, 0.024 for those who are still recovering, and 0.038 for those who say that their injury is persistent. These authors do not take "perfect health" as baseline; however, they consider road crashes as stochastic occurrences, in contrast to our main hypothesis.

The methodology developed in this paper arises from the need to control for the possible existence of selection bias. This is done in two steps: firstly, with the computation of *RTIs*-specific thresholds in ESCA02/06. Secondly, by controlling for observables, as well as introducing proxies for unobservables, at estimating the *ATET*. For instance, let us focus on health losses following severe *RTIs*, in terms of VAS tariff. Note that when *RTIs* are treated completely as random (i.e. homogeneous thresholds, and not controlling for any variable at estimating the effect), the *QoL* decrement amounts to 0.055 units. Once we control for selection in just one step, we observe a decrement of either 0.069 (homogeneous thresholds, non-random effect) or 0.102 (*RTIs*-specific thresholds, random effect). Finally, if the control is set up in both steps (*RTIs*-specific thresholds, non-random effect), the individuals experiment a 0.114 decline in *QoL*. The results follow a coherent pattern, what could add credibility to the methodology.

The main drawback at dealing with health consequences of *RTIs* is the data availability. There is still much to do before there is a complete set of data that comprises all valuable information made available (details of the accident, joint with description of the health state of the injured individuals, etc.). Meanwhile, the short-term objective consists of obtaining the best estimation of health losses under the limitation of the lack of available data.

In this paper, several measures have been developed in this direction. To start with, monitoring health-related QoL has been enhanced by establishing equivalences between cardinal and categorical health variables, since the former are the preferred measures for cost-effectiveness analysis, but the latter is more frequently enclosed in surveys. Furthermore, typical limitations have been overcome. Firstly, the potential health status has not been assumed to be perfect health. Secondly, the methodology developed in this paper has contemplated the need of controlling for the possible existence of a selection bias. Different thresholds for scaling self-assessed health for injured and non-injured individuals have been defined with this aim. Also, the $ATET$ has been estimated under three different assumptions regarding the occurrence of a road crash: treating them as fortuitous events, completely based on different sets of observable factors, or involving additional behavioral or psychological features which are, usually, unobservable. The $ATET$ has been shown to increase significantly when allowing for non-random components, remain essentially stable when controlled for different sets of socioeconomic characteristics, and decrease slightly under controls for risk aversion. Hence, the results have been proved to be robust to the estimation of the propensity score in the first part of the procedure. The estimation also suggests that the potential health state of the injured is, on average, better than the health status of those who have not had a road crash.

Our research has limitations, mainly derived from the source of data. Due to the lack of available information, continuous measures of health have been partially obtained from external data. Despite the validity of the model, it may have introduced some bias, derived from different self-perceptions. Furthermore, both surveys are administered to non-institutionalized population, so that the analysis cannot be performed for those individuals, maybe the most seriously injured, that still remain in trauma centers. There is also missing information regarding possible *RTIs* that have occurred in the past (more than one year previous to the survey), and that may be affecting the actual health state of the individual but is not observed. Finally, the $ATET$ is likely to slightly decrease if additional unobserved factors could be incorporated in the analysis. Our results bring to light the relevance of the impact of road crashes in health-related QoL.

5 Acknowledgements

The authors are grateful to Ildefonso Mendez, Elena Martinez, Climent Quintana, Juan Oliva, Eduardo Sanchez and seminar participants in the Department of Economics at the University of Alicante, for their stimulating comments on a draft of this paper. The authors also acknowledge the suggestions of the anonymous referees. Usual disclaimers apply. We thank DGT and Generalitat de Catalunya for providing the data. Financial support from Fundacion BBVA is acknowledged. We also benefit from support from Feder Funds and Spanish Ministerio de Educacion y Ciencia under project SEJ2007-62656, and Generalitat Valenciana, under project GV06/275. No conflicts exist that may bias the results of this study.

6 Appendix

.

Interval regression models, by different thresholds (dependent variable: health indices VAS and TTO)

(Table AI about here)

Coefficients for logit regressions, by group of covariates

(Table AII about here)

References

[1] Abadie A. 2005. Semiparametric Difference-in-Differences Estimators. *Review of Economic Studies* **72**: 1-19.

[2] Badia X, Fernández E, Segura A. 1995. Influence of sociodemographic and health status variables on valuation of health states in a Spanish population. *European Journal of Public Health* **5**: 87-93.

[3] Badia X, Roset M, Herdman M, Kind P. 2001. A comparison of United Kingdom and Spanish general population time trade-off values for EQ-5D health states. *Medical Decision Making* **21**: 7-16.

[4] Bishai D, Quresh A, James P, Ghaffar A. 2006. National road casualties and economic developement. *Health Economics* **15**: 65-81

[5] Cubi-Molla P. 2010. Scaling Methods for Categorical Self-Assessed Health Measures. IVIE WP-AD 2010-01.

[6] Cutler D, Richardson E. 1997. Measuring the health of the United States Population. *Brookings Papers on Economic Activity: Microeconomics*: 217-271.

[7] Generalitat de Catalunya. 2007. Enquesta de salut de Catalunya 2006 (ESCA06). Departament de Salut.

[8] Generalitat de Catalunya. 2003. Enquesta de salut de Catalunya 2002 (ESCA02). Departament de Salut.

[9] Groot W. 2000. Adaptation and scale of reference bias in self-assessments of quality of life. *Journal of Health Economics* **19**: 403-420.

[10] Hernández Quevedo C, Jones AM, Rice N. 2008. Reporting bias and heterogeneity in self-assessed health. Evidence from the British Household Panel Survey. *Cuadernos Económicos de ICE* 75: 63-97.

[11] INE (Instituto Nacional de Estadística). 1999. Encuesta sobre discapacidades, deficiencias y estado de salud (EDDES). Madrid.

[12] Lindley J, Lorgelly PK 2007. What is the relationship between income inequality and health? Evidence from the BHPS. *Health Economics* **17**(2):249-265.

[13] MacKenzie E. 2001. Measuring Disability and Quality of Life Postinjury. In: *Injury Control: A Guide to Research and Programme Evaluation.* Cambridge University Press, Cambridge.

[14] Nyman JA, Barleen NA, Kirdruang P. 2008. Quality-Adjusted Life Years Lost from Nonfatal Motor Vehicle Accident Injuries. *Medical Decision Making* **28**: 819-828.

[15] O'Donnell M, Creamer M, Elliott P, Atkin C, Kossmann T. 2005. Determinants of Quality of Life and Role-Related Disability After Injury: Impact of Acute Psychological Responses. *Journal of Trauma* **59**:1328 –1335.

[16] Redelmeier DA, Weinstein MC. 1999. Cost-Effectiveness of Regulations against Using a Cellular Telephone while Driving. *Medical Decision Making* **19**: 1-8.

[17] Seguí-Gomez M, MacKenzie E. 2003. Measuring the public health impact of injuries. *Epidemiologic Reviews* **25**: 3-19.

[18] Sturgis P, Thomas R, Purdon S, Bridgwood A, Dodd T. 2001. *Comparative Review and Assessment of Key Health State Measures of the General Population*. Research funded by the Department of Health, UK.

[19] Sullivan PW, Follin SL, Nichol MB. 2003. Transitioning the second-generation antihistamines to over-the-counter status: a costeffectiveness analysis. *Medical Care* **41**: 1382–1395.

[20] Van Doorslaer E, Wagstaff A. 1994. Measuring inequalities in health in the presence of multiple-category morbidity indicators. *Health Economics* **3**: 281-291.

[21] Van Doorslaer E, Jones AM. 2003. Inequalities in self-reported health: validation of a new approach to measurement. *Journal of Health Economics* **22**: 61-87.

[22] WHO. 2001. Mental health: new understanding, new hope. In: *The world health report 2001*. Geneva.

[23] WHO. 2004. *World report on road traffic injury prevention*. Geneva.

[24] WHO. 2007. *World Health Statistics 2007*.

	Comparison	Affected		
		[SV]	*[SL]*	*[INJ]*
N	45,567	297	553	850
Health				
SAH = very bad / bad	5.12	10.77***	4.34	6.59*
SAH = fair	23.38	29.63***	16.27***	20.94*
SAH = good	57.73	48.82***	60.94	56.71
SAH = very good	13.77	10.77	18.44***	15.76*
age-gender group				
Men, 15-24	7.15	14.48***	24.59***	21.06***
Men, 25-34	8.86	14.14***	16.82***	15.88***
Men, 35-44	8.24	8.75	7.05	7.65
Men, 45-54	7.13	5.72	6.15	6
Men, 55-64	7.13	4.71	3.98***	4.24***
Men, 65-74	8.65	7.07	5.97**	6.35**
Women, 15-24	6.98	9.76*	12.3***	11.41***
Women, 25-34	9.03	9.76	7.96	8.59
Women, 35-44	8.24	6.06	4.34***	4.94***
Women, 45-54	7.73	4.38**	4.16***	4.24***
Women, 55-64	8.97	4.71**	3.25***	3.76***
Women, 65-74	11.9	10.44	3.44***	5.88***
Education				
Less prim.or primary	51.51	45.79**	38.7***	41.18***
Secondary	32.73	43.43***	46.29***	45.29***
More secondary	15.76	10.77**	15.01	13.53*
Accidents (not *RTIs*)	2.31	5.05***	2.17	3.18*
Fear	2.16	3.7*	1.45	2.24
Usual smoker	42.66	55.22***	49.19***	51.29***

(Standard deviation in brackets)* Sign. at 10% ** Sign. at 5% *** Sign. at 1%

Table I. Average characteristics for comparison and affected groups.

	SAH		SAH	
%	ESCA02/06	SAH4	MS	%
4.95	Bad	1	Very bad	0.62
			Bad	4.67
18.25	Fair	2	Fair	23.14
47.54	Good	3	Good	57.74
22.44	Very good	4	Very good	13.84
6.82	Excellent			

Table II. Definition of SAH4

	MS		ESCA02/06		Upper bound of interval			
	N	%	N	%	VG or Exc η_1	Good η_2	Fair η_3	Bad η_4
By age-group and gender								
men, 15-29	5,491	11.8%	2,731	13.3%	0.097	0.250	0.623	1.076
men, 30-44	5,936	12.8%	3,121	15.2%	0.076	0.242	0.582	1.076
men, 45-59	4,844	10.4%	2,543	12.4%	0.052	0.230	0.536	1.076
men, 60-75	5,738	12.4%	1,914	9.3%	0.045	0.227	0.573	1.076
women, 15-29	5,312	11.4%	2,578	12.6%	0.092	0.239	0.536	1.076
women, 30-44	5,945	12.8%	2,964	14.4%	0.080	0.259	0.583	1.076
women, 45-59	5,459	11.8%	2,559	12.5%	0.059	0.240	0.588	1.076
women, 60-75	7,692	16.6%	2,118	10.3%	0.048	0.244	0.624	1.076
By *RTIs*								
Non-injured	45,567	98.2%	19,865	96.8%	0.074	0.210	0.585	1.076
slightly injured	553	1.2%	564	2.8%	0.084	0.261	0.536	1.076
Severely injured	297	0.6%	99	0.5%	0.110	0.261	0.494	1.076
Homogeneous thresholds								
	46,417	100%	20,528	100%	0.075	0.237	0.586	1.076

Table III. Thresholds by subgroups of population. *VAS,* ill-health.

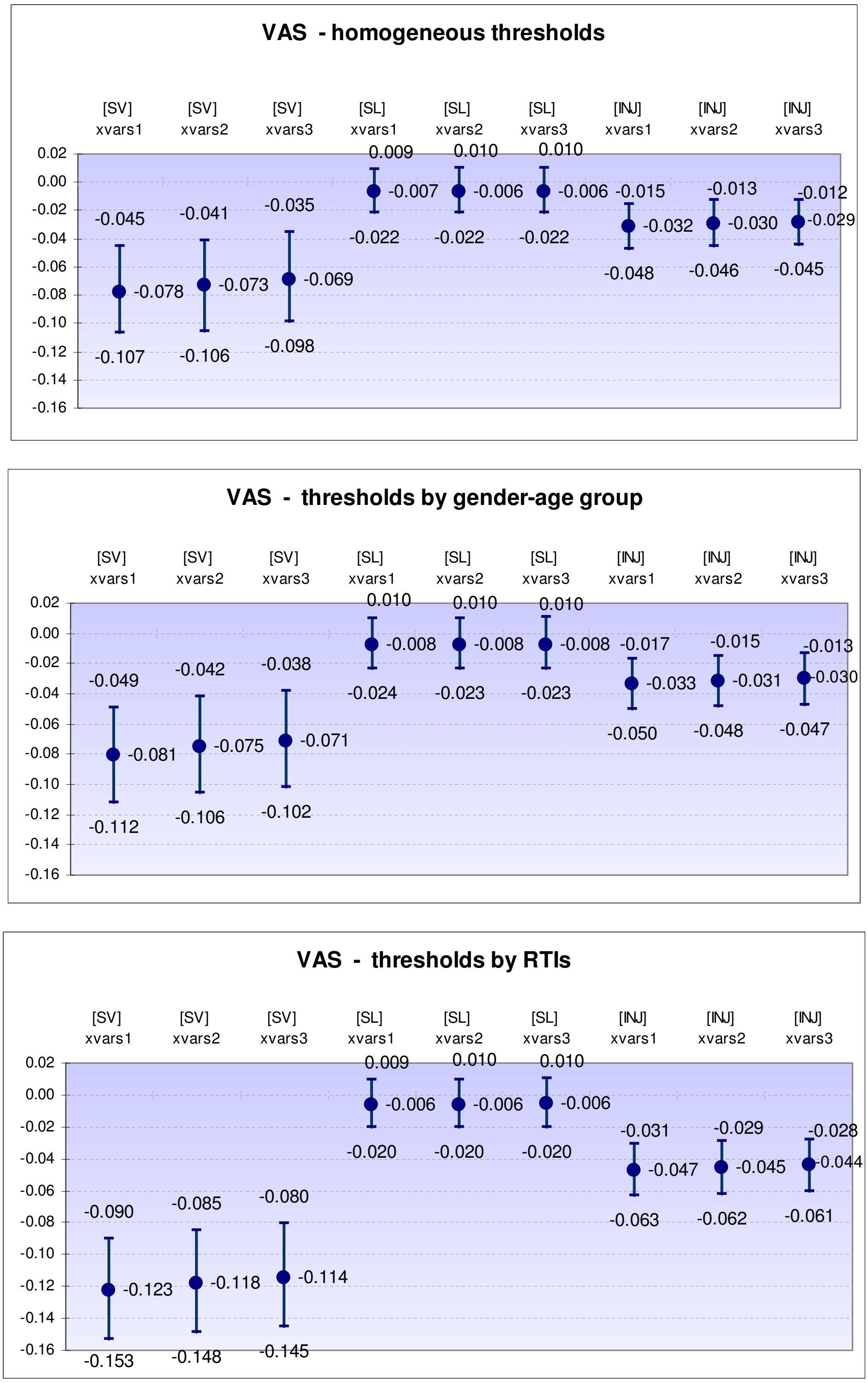

Figure 1. *ATET* for the VAS tariff, by different thresholds, affected groups and controls

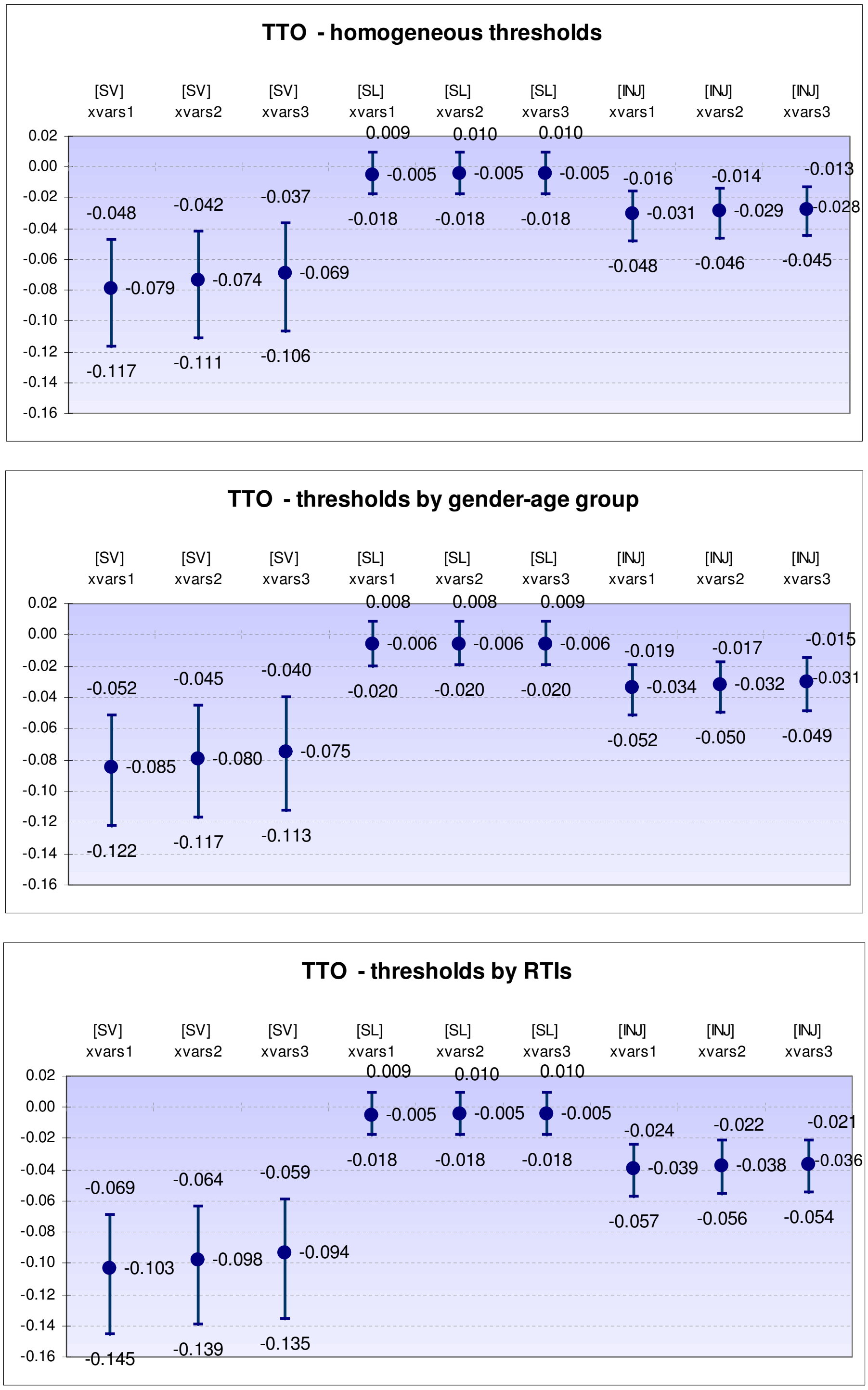

Figure 2. *ATET* for the TTO tariff (actual and re-scaled), by different thresholds, affected groups and controls

Thresholds	Controls	Random	Non-random	ΔH (%)	CI (95%)	
			VAS			
Homogeneous	*xvars1*	-0.055	-0.078	-9.64%	-12.75%	-5.84%
	xvars2	-0.055	-0.073	-9.06%	-12.60%	-5.36%
	xvars3	-0.055	-0.069	-8.65%	-11.85%	-4.58%
Gender-age	*xvars1*	-0.063	-0.081	-10.09%	-13.41%	-6.36%
	xvars2	-0.063	-0.075	-9.46%	-12.82%	-5.49%
	xvars3	-0.063	-0.071	-9.02%	-12.43%	-5.00%
RTIs	*xvars1*	-0.102	-0.123	-14.90%	-17.95%	-11.37%
	xvars2	-0.102	-0.118	-14.42%	-17.51%	-10.78%
	xvars3	-0.102	-0.114	-14.05%	-17.19%	-10.30%
			TTO			
Homogeneous	*xvars1*	-0.058	-0.079	-9.18%	-13.00%	-5.77%
	xvars2	-0.058	-0.074	-8.65%	-12.49%	-5.14%
	xvars3	-0.058	-0.069	-8.17%	-12.00%	-4.53%
Gender-age	*xvars1*	-0.067	-0.085	-10.01%	-13.79%	-6.31%
	xvars2	-0.067	-0.080	-9.42%	-13.23%	-5.59%
	xvars3	-0.067	-0.075	-8.90%	-12.83%	-4.92%
RTIs	*xvars1*	-0.082	-0.103	-12.01%	-16.11%	-8.40%
	xvars2	-0.082	-0.098	-11.50%	-15.56%	-7.80%
	xvars3	-0.082	-0.094	-11.03%	-15.18%	-7.26%

Table IV. *ATET* estimates for Severely Injured. .

Thresholds	Controls	Random	Non-random	ΔH (%)	CI (95%)	
			VAS			
Homogeneous	*xvars1*	0.039	-0.007	-0.81%	-2.58%	1.16%
	xvars2	0.039	-0.006	-0.76%	-2.55%	1.23%
	xvars3	0.039	-0.006	-0.75%	-2.55%	1.27%
Gender-age	*xvars1*	0.030	-0.008	-0.97%	-2.84%	1.21%
	xvars2	0.030	-0.008	-0.91%	-2.76%	1.28%
	xvars3	0.030	-0.008	-0.91%	-2.75%	1.28%
RTIs	*xvars1*	0.035	-0.006	-0.74%	-2.36%	1.14%
	xvars2	0.035	-0.006	-0.69%	-2.33%	1.18%
	xvars3	0.035	-0.006	-0.69%	-2.34%	1.21%
			TTO			
Homogeneous	*xvars1*	0.038	-0.005	-0.57%	-2.01%	1.07%
	xvars2	0.038	-0.005	-0.51%	-1.97%	1.11%
	xvars3	0.038	-0.005	-0.52%	-1.98%	1.10%
Gender-age	*xvars1*	0.032	-0.006	-0.73%	-2.30%	0.95%
	xvars2	0.032	-0.006	-0.68%	-2.23%	0.98%
	xvars3	0.032	-0.006	-0.68%	-2.23%	1.03%
RTIs	*xvars1*	0.038	-0.005	-0.58%	-2.02%	1.07%
	xvars2	0.038	-0.005	-0.52%	-1.99%	1.10%
	xvars3	0.038	-0.005	-0.52%	-1.99%	1.10%

Table V. *ATET* estimates for Slightly Injured.

Thresholds	Controls	Random	Non-random	ΔH (%)	CI (95%)	
			VAS			
Homogeneous	*xvars1*	0.006	-0.103	-3.84%	-5.67%	-1.88%
	xvars2	0.006	-0.098	-3.62%	-5.44%	-1.62%
	xvars3	0.006	-0.094	-3.49%	-5.31%	-1.52%
Gender-age	*xvars1*	-0.002	-0.033	-4.10%	-6.04%	-2.08%
	xvars2	-0.002	-0.031	-3.87%	-5.83%	-1.85%
	xvars3	-0.002	-0.030	-3.73%	-5.69%	-1.68%
RTIs	*xvars1*	-0.013	-0.047	-5.61%	-7.43%	-3.77%
	xvars2	-0.013	-0.045	-5.41%	-7.26%	-3.56%
	xvars3	-0.013	-0.044	-5.29%	-7.15%	-3.43%
			TTO			
Homogeneous	*xvars1*	0.004	-0.140	-3.53%	-5.39%	-1.89%
	xvars2	0.004	0.000	-3.33%	-5.21%	-1.66%
	xvars3	0.004	0.000	-3.19%	-5.04%	-1.51%
Gender-age	*xvars1*	-0.003	-0.034	-3.92%	-5.90%	-2.28%
	xvars2	-0.003	-0.032	-3.70%	-5.69%	-2.05%
	xvars3	-0.003	-0.031	-3.55%	-5.56%	-1.79%
RTIs	*xvars1*	-0.004	-0.039	-4.51%	-6.43%	-2.80%
	xvars2	-0.004	-0.038	-4.31%	-6.24%	-2.53%
	xvars3	-0.004	-0.036	-4.17%	-6.13%	-2.46%

Table VI. *ATET* estimates for Injured by a road crash (slightly or severely).

	VAS			TTO		
	All	Age-gender	*RTIs*	All	Age-gender	*RTIs*
male 15-24	-0.088	-0.003	-0.071	-0.077	0.022	-0.072
	(5.40)**	(-0.18)	(4.76)**	(4.45)**	(-1.31)	(4.15)**
male 25-34	0.008	0.014	0.012	0.019	0.027	0.02
	(-0.58)	(-1.1)	(-0.94)	(-1.29)	(-1.8)	(-1.37)
male 35-44	0.077	-0.001	0.069	0.076	-0.014	0.076
	(5.38)**	(-0.04)	(5.21)**	(4.97)**	(-0.91)	(4.93)**
male 45-54	0.134	-0.141	0.119	0.125	-0.174	0.125
	(8.81)**	(8.33)**	(8.34)**	(7.38)**	(9.29)**	(7.35)**
male 55-64	0.159	-0.103	0.139	0.145	-0.154	0.144
	(9.86)**	(5.68)**	(9.10)**	(7.69)**	(7.34)**	(7.62)**
male 65-75	0.094	-0.17	0.074	0.054	-0.256	0.054
	(5.06)**	(8.11)**	(4.10)**	(2.33)*	(9.98)**	(2.30)*
female 15-24	-0.071	-0.05	-0.06	-0.067	-0.075	-0.065
	(4.29)**	(3.36)**	(3.96)**	(3.83)**	(4.49)**	(3.68)**
female 35-44	0.064	0.032	0.054	0.051	0.058	0.051
	(4.78)**	(2.42)*	(4.33)**	(3.55)**	(3.86)**	(3.50)**
female 45-54	0.141	-0.049	0.127	0.139	-0.08	0.139
	(9.46)**	(3.07)**	(9.02)**	(8.19)**	(4.38)**	(8.15)**
female 55-64	0.161	-0.026	0.148	0.169	-0.046	0.169
	(10.23)**	(-1.49)	(9.81)**	(9.01)**	(2.24)*	(8.97)**
female 65-75	0.124	-0.059	0.11	0.115	-0.093	0.114
	(7.44)**	(3.22)**	(6.80)**	(5.61)**	(4.13)**	(5.56)**
Andalucia	-0.093	-0.098	-0.09	-0.114	-0.121	-0.114
	(3.15)**	(3.03)**	(3.18)**	(3.24)**	(3.15)**	(3.23)**
Aragon	-0.025	-0.022	-0.027	-0.04	-0.039	-0.041
	(-0.8)	(-0.65)	(-0.91)	(-1.07)	(-0.95)	(-1.07)
Asturias	0.04	0.044	0.037	0.04	0.046	0.041
	(-1.23)	(-1.24)	(-1.17)	(-1.03)	(-1.08)	(-1.04)
Canarias	0.028	0.031	0.027	0.031	0.036	0.033
	(-0.88)	(-0.89)	(-0.89)	(-0.82)	(-0.86)	(-0.85)
Cantabria	-0.015	-0.012	-0.017	-0.031	-0.028	-0.03
	(-0.46)	(-0.32)	(-0.53)	(-0.77)	(-0.63)	(-0.74)
CLM	-0.038	-0.038	-0.04	-0.058	-0.057	-0.058
	(-1.25)	(-1.14)	(-1.36)	(-1.59)	(-1.45)	(-1.58)
CYL	-0.004	0	-0.007	-0.016	-0.012	-0.016
	(-0.14)	(-0.01)	(-0.23)	(-0.46)	(-0.32)	(-0.44)

Robust z statistics in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

Table AI. Interval regression models, by thresholds (dependent variable: VAS and TTO tariffs)

	VAS			TTO		
	All	Age-gender	*RTIs*	All	Age-gender	*RTIs*
Catalunya	-0.029	-0.025	-0.029	-0.041	-0.038	-0.04
	(-0.97)	(-0.76)	(-1.02)	(-1.14)	(-0.98)	(-1.12)
CV	-0.08	-0.083	-0.079	-0.102	-0.107	-0.102
	(2.64)**	(2.49)*	(2.69)**	(2.81)**	(2.71)**	(2.80)**
Extremadura	-0.03	-0.027	-0.028	-0.034	-0.034	-0.033
	(-0.92)	(-0.77)	(-0.88)	(-0.86)	(-0.8)	(-0.83)
Baleares	-0.054	-0.063	-0.052	-0.07	-0.08	-0.069
	(-1.57)	(-1.68)	(-1.6)	(-1.72)	(-1.81)	(-1.7)
Galicia	0.095	0.101	0.088	0.103	0.112	0.104
	(3.12)**	(3.04)**	(3.04)**	(2.85)**	(2.86)**	(2.85)**
Madrid	-0.007	-0.005	-0.008	-0.013	-0.011	-0.012
	(-0.24)	(-0.14)	(-0.26)	(-0.36)	(-0.26)	(-0.34)
Murcia	-0.046	-0.058	-0.044	-0.054	-0.066	-0.054
	(-1.36)	(-1.58)	(-1.36)	(-1.35)	(-1.52)	(-1.33)
Navarra	0.006	0.006	0.003	0.002	0.001	0.001
	(-0.18)	(-0.16)	(-0.1)	(-0.05)	(-0.03)	(-0.03)
PVasco	0.036	0.041	0.032	0.032	0.038	0.033
	(-1.14)	(-1.17)	(-1.05)	(-0.85)	(-0.92)	(-0.86)
bronchitis	0.249	0.264	0.256	0.356	0.369	0.357
	(20.98)**	(21.32)**	(21.12)**	(20.89)**	(21.29)**	(20.94)**
allergy	0.032	0.028	0.031	0.035	0.033	0.036
	(3.76)**	(3.18)**	(3.72)**	(3.32)**	(3.03)**	(3.41)**
epilepsy	0.225	0.209	0.228	0.319	0.3	0.319
	(5.99)**	(5.54)**	(5.89)**	(5.76)**	(5.48)**	(5.73)**
diabetes	0.173	0.197	0.179	0.256	0.279	0.256
	(13.55)**	(14.08)**	(13.55)**	(13.40)**	(13.95)**	(13.38)**
blood pr.	0.037	0.053	0.039	0.057	0.072	0.058
	(4.05)**	(5.21)**	(4.26)**	(4.56)**	(5.28)**	(4.57)**
heart fails	0.218	0.248	0.231	0.34	0.364	0.341
	(17.30)**	(18.33)**	(17.52)**	(17.57)**	(18.39)**	(17.58)**
cholesterol	0.06	0.073	0.061	0.087	0.098	0.086
	(6.45)**	(6.98)**	(6.48)**	(6.67)**	(7.03)**	(6.61)**
cirrhosis	0.241	0.261	0.254	0.377	0.389	0.377
	(5.70)**	(5.90)**	(5.75)**	(5.73)**	(5.87)**	(5.73)**
arthritis	0.283	0.323	0.28	0.365	0.408	0.366
	(34.39)**	(35.35)**	(34.02)**	(32.71)**	(34.04)**	(32.73)**

Robust z statistics in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

Table AI (cont.). Interval regression models, by thresholds (dependent variable: VAS and TTO tariffs)

	VAS			TTO		
	All	Age-gender	*RTIs*	All	Age-gender	*RTIs*
ulcer	0.123	0.136	0.123	0.168	0.18	0.167
	(10.33)**	(10.54)**	(10.26)**	(10.15)**	(10.36)**	(10.11)**
hernia	0.126	0.14	0.127	0.171	0.185	0.172
	(10.30)**	(10.55)**	(10.19)**	(9.88)**	(10.20)**	(9.87)**
cardiovasc.	0.108	0.123	0.112	0.161	0.175	0.161
	(11.94)**	(12.56)**	(12.17)**	(12.55)**	(13.04)**	(12.53)**
anaemias	0.141	0.134	0.145	0.201	0.195	0.202
	(6.85)**	(6.38)**	(6.90)**	(6.83)**	(6.59)**	(6.85)**
nerves	0.218	0.225	0.223	0.313	0.319	0.313
	(22.66)**	(22.31)**	(22.69)**	(22.45)**	(22.45)**	(22.46)**
migraine	0.072	0.064	0.072	0.096	0.09	0.096
	(7.35)**	(6.33)**	(7.35)**	(7.28)**	(6.66)**	(7.30)**
menopause	-0.027	-0.019	-0.026	-0.03	-0.025	-0.03
	(-1.38)	(-0.89)	(-1.29)	(-1.08)	(-0.84)	(-1.09)
other	0.254	0.265	0.252	0.331	0.344	0.332
	(21.49)**	(21.46)**	(21.11)**	(20.18)**	(20.45)**	(20.20)**
mental handicap	0.205	0.185	0.215	0.313	0.286	0.314
	(7.22)**	(6.34)**	(7.28)**	(7.25)**	(6.65)**	(7.25)**
visual handicap	0.051	0.057	0.056	0.087	0.093	0.087
	(2.65)**	(2.71)**	(2.82)**	(3.06)**	(3.11)**	(3.06)**
auditory handicap	0.038	0.051	0.038	0.051	0.064	0.051
	(2.08)*	(2.58)**	(2.05)*	(-1.95)	(2.32)*	(-1.95)
articul. handicap	0.18	0.198	0.176	0.219	0.24	0.22
	(2.31)*	(2.34)*	(2.15)*	(-1.83)	(-1.93)	(-1.83)
bones handicap	0.243	0.26	0.266	0.425	0.432	0.424
	(17.11)**	(17.39)**	(17.80)**	(18.85)**	(19.06)**	(18.84)**
nervous handicap	0.345	0.365	0.372	0.576	0.58	0.576
	(10.53)**	(10.75)**	(10.79)**	(11.17)**	(11.27)**	(11.17)**
visceral handicap	0.238	0.255	0.269	0.44	0.441	0.442
	(8.53)**	(8.74)**	(9.07)**	(9.85)**	(9.86)**	(9.89)**
other handicap	0.134	0.151	0.153	0.257	0.266	0.256
	(4.26)**	(4.46)**	(4.55)**	(5.05)**	(5.15)**	(5.03)**
road crash	0.176	0.164	0.355	0.207	0.204	0.311
	(4.60)**	(4.41)**	(11.89)**	(4.28)**	(4.25)**	(7.02)**
other injuries	0.109	0.108	0.112	0.157	0.157	0.158
	(5.66)**	(5.43)**	(5.86)**	(6.14)**	(5.94)**	(6.15)**

Robust z statistics in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

Table AI (cont.). Interval regression models, by thresholds (dependent variable: VAS and TTO tariffs)

	VAS			TTO		
	All	Age-gender	*RTIs*	All	Age-gender	*RTIs*
sleep +8h	-0.033	-0.034	-0.033	-0.042	-0.044	-0.042
	(5.75)**	(5.47)**	(5.89)**	(6.15)**	(5.92)**	(6.16)**
exercise free time	-0.112	-0.112	-0.104	-0.12	-0.124	-0.12
	(14.78)**	(14.01)**	(14.61)**	(14.30)**	(13.70)**	(14.26)**
exercise wk. days	-0.029	-0.034	-0.029	-0.039	-0.044	-0.039
	(3.67)**	(3.95)**	(3.82)**	(4.26)**	(4.37)**	(4.21)**
BMI infra	0.067	0.059	0.063	0.074	0.069	0.075
	(2.97)**	(2.78)**	(2.96)**	(2.88)**	(2.70)**	(2.91)**
BMI supra	0.018	0.017	0.016	0.015	0.016	0.016
	(2.83)**	(2.55)*	(2.62)**	(2.06)*	(1.99)*	(2.09)*
medicines	0.244	0.261	0.225	0.253	0.28	0.254
	(35.57)**	(35.53)**	(34.79)**	(32.15)**	(33.07)**	(32.28)**
smoker	0.01	0.004	0.008	0.007	0.002	0.007
	(-1.45)	(-0.54)	(-1.31)	(-0.89)	(-0.23)	(-0.92)
alcohol	-0.02	-0.019	-0.019	-0.024	-0.025	-0.024
	(3.11)**	(2.84)**	(3.13)**	(3.24)**	(3.12)**	(3.18)**
married	-0.009	-0.026	-0.007	-0.004	-0.022	-0.004
	(-1.11)	(2.97)**	(-0.91)	(-0.44)	(2.09)*	(-0.44)
widow	-0.093	-0.116	-0.095	-0.129	-0.153	-0.13
	(6.83)**	(7.41)**	(7.06)**	(7.32)**	(7.73)**	(7.33)**
sep/div	0.043	0.023	0.044	0.06	0.043	0.061
	(2.42)*	(-1.2)	(2.59)**	(2.81)**	(-1.86)	(2.86)**
nostuds	0.272	0.306	0.261	0.32	0.36	0.322
	(21.67)**	(22.07)**	(21.63)**	(21.12)**	(21.73)**	(21.14)**
primstuds	0.171	0.185	0.157	0.174	0.194	0.175
	(16.94)**	(16.90)**	(16.62)**	(15.50)**	(15.82)**	(15.58)**
secndstuds	0.084	0.09	0.076	0.081	0.092	0.082
	(9.16)**	(9.37)**	(8.92)**	(8.26)**	(8.70)**	(8.29)**

Robust z statistics in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

Table AI (cont.). Interval regression models, by thresholds (dependent variable: VAS and TTO tariffs)

	VAS			TTO		
	All	Age-gender	*RTIs*	All	Age-gender	*RTIs*
unemployed	0.034	0.034	0.031	0.033	0.035	0.033
	(3.02)**	(3.01)**	(2.89)**	(2.59)**	(2.66)**	(2.60)**
unable	0.212	0.208	0.217	0.308	0.295	0.307
	(6.76)**	(6.43)**	(6.60)**	(6.35)**	(6.14)**	(6.32)**
retired	0.08	0.084	0.077	0.1	0.103	0.099
	(6.37)**	(5.85)**	(6.23)**	(6.01)**	(5.63)**	(5.96)**
housekeeper	0.055	0.058	0.05	0.056	0.062	0.055
	(5.28)**	(5.15)**	(4.99)**	(4.45)**	(4.59)**	(4.41)**
student	-0.064	-0.04	-0.059	-0.06	-0.044	-0.062
	(4.52)**	(3.08)**	(4.51)**	(4.01)**	(3.02)**	(4.12)**
other	0.062	0.066	0.059	0.073	0.079	0.072
	(3.10)**	(3.02)**	(2.96)**	(2.76)**	(2.78)**	(2.72)**
logincome	-0.085	-0.091	-0.08	-0.097	-0.105	-0.097
	(14.67)**	(14.76)**	(14.45)**	(14.03)**	(14.24)**	(14.00)**
househ. size	0.007	0.008	0.007	0.008	0.01	0.008
	(2.87)**	(3.27)**	(2.84)**	(2.86)**	(3.37)**	(2.82)**
municip. size	-0.02	-0.023	-0.019	-0.023	-0.027	-0.023
	(3.31)**	(3.67)**	(3.33)**	(3.29)**	(3.59)**	(3.30)**
nation	0.074	0.093	0.068	0.077	0.095	0.077
	(2.51)*	(2.93)**	(2.46)*	(2.36)*	(2.67)**	(2.36)*
constant	-1.432	-1.294	-1.515	-1.844	-1.665	-1.846
	(17.64)**	(14.97)**	(19.55)**	(19.33)**	(16.31)**	(19.30)**
Obs	46417	46417	46417	46417	46417	46417
Variance	0.235	0.235	0.235	0.372	0.372	0.372
% fit	65.6%	65.5%	65.4%	63.9%	64.3%	63.9%
pseudo-R2	0.469	0.420	0.480	0.498	0.453	0.497

Robust z statistics in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

McKelvey-Zavoina pseudo-R2 =[Var(predicted-h*)/Var(h*)]

Table AI (cont.). Interval regression models, by thresholds (dependent variable: VAS and TTO tariffs)

	[SV]			[SL]			[INJ]		
	xvars1	*xvars2*	*xvars3*	*xvars1*	*xvars2*	*xvars3*	*xvars1*	*xvars2*	*xvars3*
Age-gender groups (ref: male 15-24)									
Male 25-34	-0.238	-0.174	-0.339	-0.594	-0.661	-0.718	-0.496	-0.518	-0.608
	(-0.218)	(-0.23)	(-0.233)	(0.137)***	(0.142)***	(0.144)***	(0.116)***	(0.122)***	(0.124)***
Male 35-44	-0.645	-0.567	-0.794	-1.391	-1.449	-1.531	-1.154	-1.165	-1.293
	(0.250)***	(0.259)**	(0.264)***	(0.183)***	(0.188)***	(0.193)***	(0.147)***	(0.151)***	(0.155)***
Male 45-54	-0.926	-0.879	-1.105	-1.384	-1.484	-1.563	-1.253	-1.302	-1.428
	(0.288)***	(0.300)***	(0.306)***	(0.193)***	(0.200)***	(0.205)***	(0.161)***	(0.167)***	(0.171)***
Male 55-64	-1.119	-1.085	-1.293	-1.818	-1.976	-2.048	-1.6	-1.692	-1.808
	(0.309)***	(0.338)***	(0.338)***	(0.231)***	(0.245)***	(0.247)***	(0.184)***	(0.198)***	(0.198)***
Male 65-75	-0.908	-0.9	-1.101	-1.607	-1.8	-1.87	-1.389	-1.514	-1.625
	(0.267)***	(0.315)***	(0.319)***	(0.195)***	(0.221)***	(0.223)***	(0.157)***	(0.180)***	(0.182)***
Female 15-24	-0.369	-0.357	-0.357	-0.669	-0.66	-0.662	-0.588	-0.578	-0.58
	(-0.242)	(-0.242)	(-0.241)	(0.151)***	(0.151)***	(0.151)***	(0.129)***	(0.129)***	(0.129)***
Female 25-34	-0.627	-0.517	-0.652	-1.362	-1.395	-1.441	-1.13	-1.113	-1.187
	(0.241)***	(0.251)**	(0.255)***	(0.175)***	(0.179)***	(0.180)***	(0.141)***	(0.145)***	(0.147)***
Female 35-44	-1.012	-0.937	-1.089	-1.876	-1.932	-1.986	-1.591	-1.601	-1.688
	(0.282)***	(0.290)***	(0.292)***	(0.223)***	(0.225)***	(0.225)***	(0.173)***	(0.176)***	(0.177)***
Female 45-54	-1.274	-1.247	-1.258	-1.855	-1.984	-1.978	-1.682	-1.757	-1.756
	(0.317)***	(0.337)***	(0.334)***	(0.227)***	(0.234)***	(0.234)***	(0.184)***	(0.192)***	(0.191)***
Female 55-64	-1.349	-1.363	-1.256	-2.249	-2.445	-2.389	-1.949	-2.082	-2.012
	(0.309)***	(0.347)***	(0.348)***	(0.252)***	(0.270)***	(0.269)***	(0.193)***	(0.211)***	(0.211)***
Female 65-75	-0.837	-0.851	-0.718	-2.478	-2.709	-2.637	-1.785	-1.942	-1.855
	(0.237)***	(0.294)***	(0.298)***	(0.246)***	(0.272)***	(0.273)***	(0.161)***	(0.187)***	(0.189)***
Resident location (ref: La Rioja)									
Canary Islands		1.777	1.813		0.581	0.597		0.954	0.975
		(1.021)*	(1.022)*		(-0.534)	(-0.536)		(0.471)**	(0.473)**
Other regional dummies		Not sign.	Not sign.		Not sign.	Not sign.		Not sign.	Not sign.
Education (ref: more than secondary)									
Less than primary or primary		0.496	0.444		0.308	0.285		0.36	0.323
		(0.230)**	(0.226)**		(0.151)**	(0.150)*		(0.126)***	(0.125)***
Secondary		0.523	0.472		0.13	0.113		0.253	0.224
		(0.205)**	(0.204)**		(-0.135)	(-0.134)		(0.113)**	(0.113)**
Additional SE factors									
Household size		-0.018	-0.005		-0.061	-0.056		-0.046	-0.038
		(-0.06)	(-0.058)		(0.036)*	(-0.036)		(-0.032)	(-0.031)
Population size		0.052	0.034		0.106	0.102		0.086	0.076
		(-0.123)	(-0.123)		(-0.092)	(-0.093)		(-0.074)	(-0.074)
Logincome		-0.038	-0.038		0.103	0.1		0.054	0.053
		(-0.112)	(-0.112)		(-0.085)	(-0.085)		(-0.068)	(-0.068)
Behavioural									
Accidents (not *RTIs*)			0.815			0.027			0.39
			(0.273)***			(-0.296)			(0.202)*
Fear			0.643			0.064			0.36
			(0.316)**			(-0.363)			(-0.238)
Usual smoker			0.61			0.226			0.351
			(0.139)***			(0.095)**			(0.079)***
Constant	-4.328	-5.428	-5.697	-3.176	-4.607	-4.662	-2.901	-4.144	-4.268
	(0.154)***	(1.726)***	(1.723)***	(0.088)***	(1.129)***	(1.129)***	(0.077)***	(0.936)***	(0.936)***
Observations	45864	45864	45864	46120	46120	46120	46417	46417	46417

Robust standard errors in parentheses. * Significant at 10% level; ** Significant at 5% level;*** significant at 1% level

Table AII. Coefficients for logit regressions *[SV]*, *[SL]* and *[INJ]*, by group of covariates