Health misperception and healthcare utilisation of older Europeans

Submission to the European Population Conference (EPC 2020)

Preliminary draft – please do not quote or circulate without permission by the authors

Sonja Spitzer
sonja.spitzer@iiasa.ac.at
International Institute for Applied Systems Analysis,
Laxenburg, Austria

Mujaheed Shaikh
shaikh@hertie-school.org
Hertie School of Governance,
Berlin, Germany

Abstract

Understanding the drivers of healthcare utilisation in Europe is of utmost importance in the context of rapid population ageing and increasing public health expenditure. This paper explores individual health perception biases as a potential determinant of doctor visits and concomitant out-of-pocket expenditure. Based on longitudinal data from the Survey of Health, Ageing and Retirement in Europe, we observe how biased beliefs about health status affect healthcare utilisation of the population 50+ in 15 European countries. Using biomarkers and their self-reported equivalents, we find that individuals who underestimate their health visit the doctor more often than individuals who correctly assess their health. The higher healthcare utilisation is accompanied by larger out-of-pocket payments. By contrast, individuals that overestimate their health visit the doctor less often and have lower out-of-pocket payments. The effects are larger for men, which is particularly relevant given the well documented gender differences in healthcare seeking behaviour.

Keywords: Healthcare utilisation, health perception, overconfidence and underconfidence, SHARE data, doctor visits, out-of-pocket expenditure

JEL classification: D83, I11, J14, H51

1 Introduction

Belief or confidence is a strong predictor of behaviour in different domains of life. It has significant implications in areas such as education, labour market decisions and outcomes, savings and investment choices, and political decisions among others as shown by prior work (Anderson et al. 2017, Ortoleva & Snowberg 2015, Reuben et al. 2017). It is particularly relevant for health as it can directly affect the risk of accident and injury (Preston & Harris 1965) and can have serious long lasting effects on health and mortality; recent work in this domain shows that overconfidence is related to engaging in risky health behaviors (Arni et al. 2019). Individual's (mis)perception of own health can also affect health seeking behaviour and subsequent utilisation of healthcare services such as timely screenings, immunisations, annual health checks, and doctor visits.

The large literature that assesses the determinants of healthcare use, while highly profuse, leaves substantial scope to examine individuals beliefs about own health and abilities. In explaining variation in health expenditure and healthcare utilisation, this literature focuses on either the supply side i.e. provider (more specifically physician) confidence and precision (Baumann et al. 1991, Berner & Graber 2008, Cutler et al. 2013, Meyer et al. 2013), or it focuses on easily observable demand characteristics such as age, gender, income, social class, employment and education (Biro 2013, Cameron et al. 2010, Tavares & Zantomio 2017, Vallejo-Torres & Morris 2013, Van Doorslaer et al. 2004, Zhang et al. 2018). Concurrently, the large literature in psychology and economics that examines the multitude of outcomes affected by over- or underconfidence, has barely scratched the surface with respect to health outcomes and healthcare utilisation.

In this paper, we combine the above two streams and contribute to existing literature by focusing on a hitherto ignored dimension on the demand side to provide new evidence on the relationship between health misperception and healthcare utilisation across 15 European countries. The European setting is particularly interesting for at least two reasons. First, utilisation of health services is conditional on having access to such services; a fair comparison of utilisation requires that the entities being compared have no initial variation in accessibility. Due to the very principal of universal coverage in these countries, it is fair to say that almost everyone has access to the health system, unlike other systems such as the United States. Second, Europe is a policy relevant setting due to (i) the growing pressure on the healthcare systems to reduce expenditures and unnecessary care and (ii) a rapidly ageing population.

The existing literature has repeatedly shown that individuals frequently over- or underestimate their own health status (Bago d'Uva et al. 2008, Beaudoin & Desrichard 2011, Coman & Richardson 2006, Furnham 2001, Jürges 2007). Additionally, health perception differs by sociodemographic characteristics such as race (Jackson et al. 2017), age (Srisurapanont et al. 2017, Crossley & Kennedy 2001), gender (Schneider et al. 2012, Merrill et al. 1997), country of residence (Spitzer & Weber 2019, Capistrant et al. 2014, Jürges 2007), and education (Bago d'Uva et al. 2008). Others that are more closely related to this paper, have shown that the difference between subjective and predicted survival probability affects healthcare utilisation (Biró 2016*a*), and individuals with lower time preference and higher life expectancy are more likely go for cancer screening (Picone et al. 2004).

Despite the limited empirical evidence showing the direction in which misperception affects utilisation, it is a priori ambiguous how over- or under-confidence might be related to healthcare use. On the one hand, individuals that overestimate their health are less likely to visit the doctor when necessary, seek medical attention, or attend timely screenings due to their belief of perfect health. Health perception is likely to affect physical activity, which is shown to decrease healthcare utilisation (Rocca et al. 2015). On the other hand the same individuals might engage in activities or behaviour detrimental to health and thus end up at the hospital more often. Older individuals that overestimate their mobility are more prone to suffer fall-induced injuries (Sakurai et al. 2013). Similar line of reasoning suggests that individuals that underestimate their own health overutilise healthcare services by seeking care and purchasing relatively more medication when not necessary. Assessing the relationship between health perception and healthcare utilisation thus remains an empirical task which we aim to undertake in this paper.

We use longitudinal data on 15 European countries from the Survey of Health, Ageing and Retirement (henceforth SHARE) conducted across Europe. We categorise individuals into those that overestimate their current health, those that underestimate their current health and those that achieve concordance based on the objective biomeasure data (a form of physical performance measurement data) and the same subjective health performance measure. To measure healthcare utilisation, we use the self-reported 'annual number of doctor visits' data. Using count models, a rich set of controls, and the longitudinal feature of our data, we find that relative to individuals that achieve concordance (in other words accurately estimate their health), individuals that underestimate their health visit the doctor more often (approx. 2 more visits). In contrast, individuals that overestimate their health visit the doctor less often. Heterogeneity by gender shows that the effects are larger in size for men than women. We also analyse concomitant out-of-pocket (OOP) expenditure via log-gamma models and find that individuals that underestimate their health have higher expenses while individuals that overestimate their health have lower expenses. Our results are not biased by other individual characteristics such as education, age, employment or marital status. Neither are they a manifestation of the reverse relationship between healthcare utilisation and estimation of one's health (visiting the doctor more often allows an accurate assessment of own health) since we estimate the relationship between current perception and future health care. The results are robust to different model specifications, different estimation methods and also to different measures of health perception.

The paper is structured as follows. In the next Section, we describe the data and variables. In Section 3, we introduce our methodology. Section 4 presents and discusses the results and Section 5 concludes the paper.

2 Data and descriptives

We analyse the effect of health perception on healthcare utilisation based on SHARE, a representative cross-country panel study of non-institutionalised individuals aged 50 and older as well as their younger spouses (Börsch-Supan et al. 2013).[1] The survey provides rich information on health, socio-economic background, employment and social networks based on about 380,000 interviews from around 140,000 individuals. It is particularly well suited to study European countries, since the data is ex ante harmonised. Also, it focuses on older individuals, who generally face higher healthcare needs than the young, making it the ideal data source for our analysis. SHARE was previously used to analyse healthcare utilisation by, among others, Bíró (2014), Bolin et al. (2009), Paccagnella et al. (2013), and Tavares & Zantomio (2017).

2.1 Sample construction

The chair stand test, which is used to obtain objective biomeasure data, is conducted in Wave 2 (2006/2007) and Wave 5 (2013), which is why our analysis mostly relies on these waves (Börsch-Supan 2019*b*,*d*). In addition, we utilise Wave 4 (2010-2012) and Wave 6 (2015) to obtain annual numbers of doctor visits and concomitant OOP expenditure in the subsequent wave[2] (Börsch-Supan 2019*c*,*e*). Hence, Wave 2 (w) is matched with Wave 4 ($w+1$) and Wave 5 (w) is matched with Wave 6 ($w+1$). We treat the data as pooled cross-sectional.

Our main analysis focuses on the effect of health perception at wave w on doctor visits at wave w+1, which is why we drop all observations that do not provide information on doctor visits at wave w+1. This concerns mostly respondents that participated in Wave 2 but not in the subsequent Wave 4, or respondents that participated in Wave 5 but not in the subsequent Wave 6. We also exclude all respondents younger than 50 years. Overall, this results in 62,696 observations from 15 European countries covering various European regions, namely Austria, Belgium, Czechia, Denmark, Estonia, France, Germany, Italy, Luxembourg, the Netherlands, Poland, Slovenia, Spain, Sweden and Switzerland.

For the regression analyses, the sample is split into individuals that are unimpaired and individuals that are impaired based on the objective biomeasure data. According to the chair stand test, 49,501 observations are unimpaired and 9,562 observations are impaired. The sample analysing OOP payments is smaller, since OOP payments were not captured in Wave 4 (Section 2.2.2).

[1]This paper uses data from SHARE Waves 1, 2, 4, 5, and 6 (DOIs: 10.6103/SHARE.w1.700, 10.6103/SHARE.w2.700, 10.6103/SHARE.w4.700, 10.6103/SHARE.w5.700, 10.6103/SHARE.w6.700). The SHARE data collection has been funded by the European Commission through FP5 (QLK6-CT-2001-00360), FP6 (SHARE-I3: RII-CT-2006-062193, COMPARE: CIT5-CT-2005-028857, SHARELIFE: CIT4-CT-2006-028812), FP7 (SHARE-PREP: GA N211909, SHARE-LEAP: GA N227822, SHARE M4: GA N261982) and Horizon 2020 (SHARE-DEV3: GA N676536, SERISS: GA N654221) and by DG Employment, Social Affairs & Inclusion. Additional funding from the German Ministry of Education and Research, the Max Planck Society for the Advancement of Science, the U.S. National Institute on Aging (U01_AG09740-13S2, P01_AG005842, P01_AG08291, P30_AG12815, R21_AG025169, Y1-AG-4553-01, IAG_BSR06-11, OGHA_04-064, HHSN271201300071C) and from various national funding sources is gratefully acknowledged(see www.share-project.org).

[2]SHARE Wave 3 focuses on people's life histories and thus cannot be utilised for our analysis.

For the analysis of effect heterogeneities, we split the sample by gender, by country of residence and by the number of chronic diseases (Section 4.2). Also, for the analysis of additional specifications of health perceptions based on cognition and walking ability, we add more waves to the analysis (Section 4.4).

2.2 Outcome variables

In line with the literature, we proxy health care utilisation by the number of annual doctor visits (see Bago d'Uva & Jones 2009, Bíró 2016*b*, Bolin et al. 2009, Lugo-Palacios & Gannon 2017, Tavares & Zantomio 2017, Zhang et al. 2018, among others). By analysing the number of annual doctor visits, we are able to capture the effects of health perception on public expenditure, since doctor visits are frequently subsidised by the public. In addition, doctor visits are a good indicator for healthcare seeking behaviour in general and, in particular, for preventative health care and screenings. In addition to doctor visits, we analyse annual OOP payments for doctor visits, which allows us to analyse the effect of health perception on private healthcare expenses.

2.2.1 Annual doctor visits

The annual number of doctor visits, emergency room visits and outpatient clinic visits is ascertained by the question "Now please think about the last 12 months. About how many times in total have you seen or talked to a medical doctor or qualified/registered nurse about your health? Please exclude dentist visits and hospital stays, but include emergency room or outpatient clinic visits". The survey question is phrased almost identical in both waves 4 and 6, however, the words "or qualified/registered nurse" are excluded in Wave 4.

The number of doctor visits is top coded at 98 visits per year. On average, individuals in our sample visit the doctor 7.6 times per year. The median, however, is much lower (5 times), demonstrating the variable's strong right-skewness (Table 2). Naturally, individuals that suffer from chronic diseases or activity limitations visits the doctor more frequently than healthy individuals; thus, the number of doctor visits also increases with age. Furthermore, women and the less educated visit the doctor more often. (Table 4).

2.2.2 Out-of-pocket expenditure for doctor visits

If participants report that they have seen or talked to a doctor, they are asked "Did you pay anything yourself for your doctor visits (in the last twelve months)? Please also include expenses for diagnostic exams, such as imaging or laboratory diagnostics". If they answer the question with 'yes', the participants are then asked "Overall, how much did you pay yourself for your doctor visits (in the last twelve months), that is how much did you pay without getting reimbursed by (a health insurance/ your national health system/ a third party payer)?". The amount of OOP payments is based on the latter question and set to zero if the respondent did not visit a doctor at all or if they state zero payments for doctor visits. All values are presented in Euros. Implausibly large values are set to missing, as suggested by SHARE (Jürges 2015). This concerns 3,207 observations.

OOP payments are available in Wave 6, but not in Wave 4; thus, we assess the effect of health perception at Wave 5 (w) on OOP expenditure in Wave 6 ($w + 1$) only. Consequently, the sample is smaller when we analyse OOP payments than it is when we analyse doctor visits. Since potential deductibles would not only consider expenditure for doctor visits, but also for other health care services such as dentist visits and hospital stays, deductibles are not considered when preparing the OOP expenditure variable.

Mean OOP expenditure is 74 Euros per year, however, 62% of the participants have zero OOP payments in Wave 6; thus, the median is zero (Table 2). Interestingly, OOP payments appear not to increase with the number of chronic diseases or activity limitations, yet educational attainment has a strong positive correlation with OOP expenditure. Furthermore, mean OOP payments vary substantially between countries and are highest in Luxembourg, Switzerland, Italy, and Austria (Table 5).

2.3 Explanatory variable: health perception

Our measure of health perception is strongly related to the concept of belief and confidence, which is shown to have substantial impact on human behaviour. In particular, our measure relates to the most common interpretation of over- and underconfidence, namely over- and underestimating of one's performance, actual ability, chance of success, or level of control (Moore & Healy 2008). Assuming an underlying true level of health, we group individuals according to their perception of their health status. More specifically, we differentiate between individuals that perceive their health status correctly (concordance), those who believe that they are healthier than they really are (overestimating) and those who believe that they are unhealthier than they really are (underestimating). The true level of health is proxied by objective biomeasure data based on physical performance measures. This objective information about the respondent's health is matched with the respondent's subjective assessment of his or her health, thus revealing whether their beliefs are correct or not.

SHARE provides several biomeasures which can be utilised to proxy true health. The most suited measure for analysing deviations between objective and subjective health is the ability to stand up from a chair, since the self-assessed variable relates directly to its tested equivalent. Furthermore, the ability to stand up from a chair has a binary outcome, i.e. impairment if the individual is unable to stand up and unimpairment otherwise, which facilitates the comparison of the self-assessed with the tested measure. Hence, we use the subjective and objective ability to stand up from a chair to capture health perception, which was already done in previous work (Spitzer & Weber 2019). In additional analyses, we also observe the deviation between subjective and objective cognition as well as between subjective and objective walking ability (Section 4.4).

To evaluate the subjective ability to get up from a chair, survey participants are asked whether

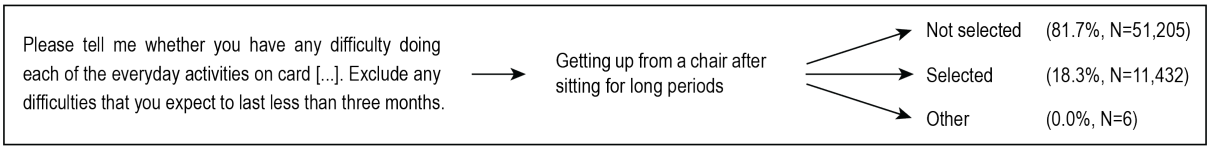


Figure 1: Survey question ascertaining subjective impairment (response category proportions in brackets)

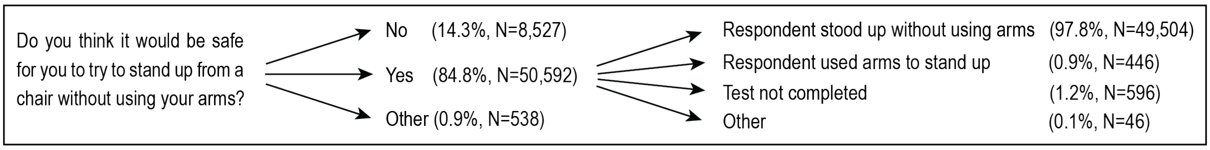


Figure 2: Sequence of questions ascertaining objective impairment (response category proportions in brackets)

they have difficulties in getting up from a chair. Figure 1 provides the detailed survey question. Individuals are considered subjectively impaired if they report difficulties in getting up from a chair and subjectively unimpaired if they do not. Overall, 18.3% of the survey participants in our sample are considered subjectively impaired.

Following this self-assessment, individuals are asked to physically stand up from a chair. The chair stand test is introduced with the interviewer saying "The next test measures the strength and endurance in your legs. I would like you to fold your arms across your chest and sit so that your feet are on the floor; then stand up keeping your arms folded across your chest. Like this...". The exact sequence of questions leading to the chair stand test is visualised in Figure 2. Individuals are considered objectively unimpaired if they stand up without using their arms. Individuals are considered objectively impaired if they are unable to stand up from the chair, if they have to use their arms to stand up and if they think it is unsafe to try to stand up from the chair. In addition, we conduct a robustness analysis in which individuals feeling unsafe to attempt the test are excluded from the sample (Section 4.3). Overall, 16.2% of the survey participants in our sample are considered objectively impaired. It is important to note that the chair stand test in Wave 2 was only conducted among those younger than 76 years. Thus, all countries that only participated in Wave 2 do not have chair stand test results for participants

Table 1: Overview health perception categories

	Objective impairment			
Subjective impairment	Unimpaired	Impaired	Missing values	**Total**
Unimpaired	Concordance: 43,348	Overestimating: 5,276	3	**49,504**
Impaired	Underestimating: 6,153	Concordance: 4,286	7	**9,569**
Missing values	2,581	993	49	**3,623**
Total	**51,205**	**11,432**	**59**	**62,696**

Note: No weights applied

aged 76 and older.

Three health perception outcomes are possible for each survey participant based on the self-assessed and the tested mobility impairment measure. These options are visualised in Table 1. If the self-assessed measure coincides with the tested measure, concordance is achieved. This is the case for the majority of the observations. If, however, survey participants report no difficulty to get up from a chair, but are unable to perform the chair stand test, they are considered to overestimate their health. On the contrary, participants are considered to underestimate their health if they report difficulty to get up from a chair, but are able to perform the test.

Overall, individuals that perceive their health correctly have lower numbers of yearly doctor visits (6.8 times) than individuals that suffer from health misperception, i.e. overestimating or underestimating (9.7 times). This is plausible since health perception is correlated with a range of factors that are also related to health care utilisation. For example, older individuals are more likely to be impaired and thus visit the doctor more often (Table 4) and they are also more likely to under- or overestimate their health (Table 6). Thus, it is important to analyse impaired and unimpaired individuals separately.

Individuals that are objectively unimpaired (i.e. able to stand up from a chair) and report their level of impairment correctly (i.e. positive concordance) visit the doctor 6.3 times per year. This is much lower than the number of doctor visits by people that are objectively unimpaired, but underestimate their health (10.2 visits). On the contrary, individuals that are objectively impaired and report their level of impairment correctly (i.e. negative concordance) visit the doctor 12.4 times per year, which is much higher than the number of doctor visits by individuals that are objectively impaired but overestimate their health (9.3 visits). These descriptive findings indicate that we have to differentiate between positive and negative concordance, or in other words, between objectively impaired and unimpaired individuals. In terms of regression analysis, this means that we will split the sample into impaired and unimpaired individuals.

2.4 Additional control variables

We control for a range of variables that might otherwise confound our results. Summary statistics for these control variables are provided in Table 2 and cross tabulations of control variables, doctor visits, health expenditure and health perception in Tables 4 to 6. Most importantly, we control for other health dimensions at wave w. In particular, we include the number of chronic diseases and the number of limitations in instrumental activities of daily living (IADLs) in our model. Chronic conditions that are considered are heart problems, high blood pressure or hypertension, high blood cholesterol, a stroke or cerebral vascular disease, diabetes, chronic lung diseases, cancer, stomach or duodenal ulcer, Parkinson disease, cataracts, hip fractures, other fractures and Alzheimer's disease. 36% of the sample have zero chronic diseases at wave w, the weighted mean is at 1.2 diseases. IADLs that are considered are difficulties in dressing, walking across a room, bathing or showering, eating and cutting up food, getting in our out of bed, using the toilet, using a map, preparing a hot meal, shopping for groceries, making a telephone call, taking medications, doing work around the house or garden, and managing money. 82% of the sample have zero IADLs at wave w, the weighted mean is at 0.5.

We also control for socio-demographic characteristics, since they are expected to influence health perception as well as health care utilisation. In particular, we include age and age squared, gender and educational attainment according to the International Standard Classification of Education (Eurostat 2018). Since pensioners appear to have higher healthcare utilisation (Bíró 2016*b*, Zhang et al. 2018), we also consider whether an individual is retired as opposed to all other employment options (employed, self-employed, unemployed, permanently sick or disabled, homemaker, other). Also, we control for whether the survey participant is married or in a registered partnership as opposed to never married, divorced or widowed.

The effect of economic resources on healthcare utilisation is considered via equivalised household income. Since household income has many missing values in SHARE, the dataset comes with two additional imputed variables. We use one of these imputed variables in our model and conduct a robustness analysis with the second imputed variable (Section 4.3). We have equivalised household income applying the square root scale, where household income is divided by the square root of the household size. Furthermore, we have applied a cube root transformation to normalise the skewed income distribution (Cox 2011). Standard log normalisation was not feasible due to a substantial number of zero values: 928 observations have zero income according to the first imputed income variable and 1,289 observations have zero income according to the second imputed income variable. We run a robustness analysis for which we use equivalised household income that was not normalised (Section 4.3).

Risk aversion is also controlled for, since risk adverse individuals appear to have a higher demand for medical tests (Picone et al. 2004). The control variable is based on the following survey question: "When people invest their savings they can choose between assets that give low return with little risk to lose money, for instance a bank account or a safe bond, or assets with a high return but also a higher risk of losing, for instance stock s and shares. Which of the statements on the card comes closest to the amount of financial risk that you are willing to take when you save or make investments?", with possible answers "substantial", "above average", "average" and "no". Most individuals say, that they take no risk (76.5%). In Wave 2, this question is answered by a financial household respondent only. Thus, we assume that the risk aversion of the financial respondent is representative for the entire household. Finally, we include country and wave dummies in our model.

Table 2: Summary statistics

	N	Mean	Std. Dev.	Min.	Max.	Median
Annual number of doctor visits at w+1	62553	7.603	9.926	0	98	5
Annual out-of-pocket expenditure for doctor visits at w+1	42544	74.307	311.762	0	47500	0
Health perception	58924	.303	.642	0	2	0
Health perception (individuals that felt 'unsafe' dropped)	50419	.255	.653	0	2	0
Subjective impairment	62494	.183	.387	0	1	0
Objective impairment	58934	.172	.378	0	1	0
Objective impairment (individuals that felt 'unsafe' dropped)	50423	.026	.159	0	1	0
Number of chronic diseases at w	62436	1.178	1.242	0	12	1
Number of chronic diseases at w+1	62534	1.25	1.265	0	10	1
Number of activity limitations at w	62490	.522	1.645	0	13	0
Number of ctivity limitations at w+1	62537	.768	2.141	0	13	0
Age	62553	65.128	10.179	50	103	64
Gender	62553	1.546	.498	1	2	2
Education	61667	1.745	.76	1	3	2
Is retired	62083	.516	.5	0	1	1
Is married	61295	.676	.468	0	1	1
Household income	62553	46144.25	75562.65	0	1200000	24000
Equivalised household income (not normalised)	62553	32630.39	54155.95	0	1080000	16970.56
Equivalised household income (cube root normalisation)	62553	27.749	10.534	0	102.599	25.698
Equivalised household income (cube root normalisation) 2	62553	24.951	8.586	0	118.504	24.858
Risk aversion	60972	3.722	.55	1	4	4
Survey wave	62553	3.621	1.495	2	5	5
Country	62553	16.211	4.575	11	35	16
Household size	62553	2.205	1.087	1	12	2

Note: Calibrated cross-sectional individual weights are applied.

3 Method

Ideally, we would like to randomly assign health perception to individuals to elicit causal effects of (mis)perception on healthcare utilisation and expenditure. In the absence of such random assignment, we rely on the panel dimension of the SHARE survey and control for a rich set of variables to account for confounding effects and bias due to reverse causation. Health perception is expected to affect healthcare utilisation, but the opposite mechanism that health care utilisation precedes health perception appears plausible too. For example, individuals that frequently visit the doctor might achieve concordance more likely, since they receive more information about their health status. To overcome potential endogeneity problems, we analyse the effect of current health perception (wave w) on future health care utilisation (wave $w + 1$).

The main outcome variable – annual doctor visits – is strongly right-skewed, yet without severe mass at zero. To accommodate these features, we apply a negative binomial model with mean dispersion, which is used frequently in the health care literature. We abstain from employing a simple Poisson model, since the outcome variable's variance is much larger than its mean. Thus, the number of doctor visits of individual i at wave $w + 1$ ($\text{DOCTOR}_{i,w+1}$) is assumed to follow a Poisson distribution, but with a negative binomial specification for which each individual unit has a separate, gamma distributed mean. More specifically

$$\text{DOCTOR}_{i,w+1} \sim \text{Poisson}(\mu_{i,w+1}), \tag{1}$$

where

$$\mu_{i,w+1} = \exp(\beta \times \text{HEALTH PERCEPTION}_{i,w} + \gamma \times \text{HEALTH}_{i,w} + \delta \times \text{X}_{i,w} + \nu_i), \qquad (2)$$

and

$$\exp(\nu_i) \sim \text{Gamma}(1/\alpha, \alpha) \qquad (3)$$

HEALTH PERCEPTION is a 3-category variable indicating whether individual i achieved concordance, overestimated or underestimated his or her health at wave w. The vector HEALTH includes the number of chronic diseases in period w as well as the number of IADLs in period w; thus, in the same period as health perception. The vector of control variables $\text{X}_{i,w}$ includes age and age squared, the individual's gender, its educational attainment, household income, risk aversion and control dummies for the survey wave as well as for the country of residence. The terms β, γ and δ represent coefficients.

When analysing the effect of health perception on OOP expenditure, we apply a non-linear model with log link and Gamma family instead of the negative binomial model to account for the continuous character of the outcome variable as well as for the excess zeros. The specification of the variables included, however, remains identical to that described in Equation 2.

As discussed earlier, the sample is split into (i) individuals that are objectively unimpaired in period w, i.e. able to stand up from the chair, and (ii) individuals that are objectively impaired in period w, i.e. unable to stand up from the chair. Hence, for the first sample we analyse the effect of overestimating on healthcare utilisation and for the second sample we analyse the effect of underestimating on healthcare utilisation. For heterogeneity and mediation analyses, we further split the sample by gender, country and the number of chronic diseases.

For the main analysis, health perception is based on the self-assessed and tested ability to stand up from a chair. In Section 4.4, we explore whether our results are robust to different specifications of health perception. In particular, we estimate Equation 2 using cognition and the ability to walk as basis for the health perception variable.

4 Results

4.1 Main results

The main regression results for counts of annual doctor, emergency room and outpatient clinic visits at wave $w + 1$ are presented in Table 3 along with the results on OOP expenditures. Columns 1 and 2 show estimated coefficients for the sample that is objectively unimpaired and columns 3 and 4 for the sample that is objectively impaired. We find strong and significant ef-

Table 3: Annual number of doctor visits and OOP expenditure for doctor visits at w+1

	(1) Unimpaired Doctor visits	(2) Unimpaired OOP	(3) Impaired Doctor visits	(4) Impaired OOP
Health perception (ref.: concordance)				
Underestimating	0.250***	0.193*		
	(0.018)	(0.077)		
Overestimating			-0.156***	-0.299*
			(0.027)	(0.144)
Chronic diseases	0.180***	0.127***	0.133***	0.174***
	(0.005)	(0.031)	(0.009)	(0.051)
Activity limitations	0.090***	0.064	0.032***	0.019
	(0.008)	(0.035)	(0.005)	(0.027)
Age	-0.002	0.058	0.023	0.196*
	(0.010)	(0.059)	(0.017)	(0.085)
Age squared	0.000	-0.000	-0.000	-0.002**
	(0.000)	(0.000)	(0.000)	(0.001)
Woman	0.036**	0.165**	0.003	0.413***
	(0.012)	(0.058)	(0.027)	(0.104)
Educ. group (ref.: medium)				
Low	-0.007	-0.409***	0.024	-0.107
	(0.016)	(0.071)	(0.032)	(0.120)
High	-0.008	0.479***	-0.078	0.448**
	(0.016)	(0.092)	(0.042)	(0.156)
Retired	0.031	-0.045	0.015	0.152
	(0.017)	(0.103)	(0.031)	(0.239)
Married	-0.025	0.020	0.014	0.084
	(0.015)	(0.071)	(0.028)	(0.148)
Equiv. hh income (cube root)	-0.001	0.011***	-0.001	0.011
	(0.001)	(0.003)	(0.001)	(0.007)
Risk aversion (ref.: no risk)				
Substantial	0.067	-0.164	-0.064	-0.164
	(0.063)	(0.134)	(0.130)	(0.277)
Above average	-0.134***	0.414*	-0.106	1.319
	(0.033)	(0.196)	(0.108)	(0.771)
Average	-0.009	0.172	-0.061	0.176
	(0.015)	(0.089)	(0.040)	(0.147)
Wave 5	-0.091***		-0.042	
	(0.015)		(0.037)	
Constant	1.525***	1.595	1.331*	-2.832
	(0.349)	(2.001)	(0.592)	(2.906)
Control variables country	Yes	Yes	Yes	Yes
N	47,377	33,575	8,780	6,413
Pseudo R2	0.024		0.019	
AIC	269,248	305,417	57,293	57,996
BIC	269,520	305,644	57,512	58,178
SE	cluster	cluster	cluster	cluster

Note: "Unimpaired" refers to the sample that is objectively unimpaired, i.e. able to stand up from the chair and "Impaired" refers to the sample that is objectively impaired, i.e. unable to stand up from the chair. The dependent variable "doctor visits" is based on the annual number of doctor visits, visits to emergency rooms and outpatient clinic visits at wave w+1, i.e. Wave 4 or Wave 6. All explanatory variables are taken from wave w, i.e. Wave 2 or Wave 5 respectively. The estimated coefficients are based on a negative binomial regression model with mean dispersion. The dependent variable "OOP" is based on annual out-of-pocket payments for doctor visits at wave w+1, i.e. Wave 6. All explanatory variables are taken from wave w, i.e. Wave 5. The coefficients are estimated based on a generalised linear model model with log link and a Gamma family. Standard errors are clustered at the household level and presented in parentheses. * p<0.05, ** p<0.01, *** p<0.001

fects for health perception on healthcare utilisation. Individuals that underestimate their health visit the doctor +28.4% more often in the subsequent period than individuals that achieve concordance. Computing marginal effects at means shows that this results in +1.6 doctor visits per year. On the contrary, individuals that overestimate their health go to the doctor less often than those that achieve concordance. Overestimating health at wave w results in -14.4% doctor visits at wave $w + 1$ compared to individuals that perceive their health correctly. The marginal effect at means of overestimating health on healthcare utilisation is -1.4 doctor visits per year.

Individuals that underestimate their health not only visit the doctor more often but also have significantly larger OOP expenses. On average, expenditures are 19.3% higher for those who underestimate their health as compared to those who achieve concordance. While we only observe OOP payments, similar effects are expected for public spending and hence our estimates underestimate the impact of health misperception on health expenditure. On the contrary, individuals that overestimate their health spend 30% less on doctor visits.

The results for doctor visits in Table 3 are based on a negative binomial model with mean dispersion. Figures 3 and 4 show that this model has the best fit, compared to a simple Poisson model, a negative binomial model with constant dispersion, and a zero-inflated Poisson model. The results for OOP payments are based on a log-gamma model. According to the Akaike information criterion and the Bayesian information criterion, the log-gamma model has a better fit than a log-Gaussian model or a log-Poisson model.

4.2 Effect heterogeneity

We asses heterogeneity of our main results in several ways. First, prior literature has shown differences in health perception by individual characteristics, most importantly, by gender (Merrill et al. 1997, Schneider et al. 2012, Spitzer & Weber 2019). Following such evidence, we assess if the relationship between health (mis)perception and utilisation also differs between men an women. Separate analyses by gender reveal that the association of health misperception on the number of annual doctor visits is slightly larger in magnitude for men than for women (Table 7). Marginal effects at means show that men that underestimate their health visit the doctor an additional 1.8 times compared to men that achieved concordance. For women, the difference is an additional 1.5 doctor visits. Men that overestimate their health have 1.5 less annual doctor visits compared to men that achieve concordance. For women, it is 1.3 visits less. Although the differences are not large enough to merit significant attention based on our coefficient sizes, they contribute to the previous literature that documents similar results. Gender differences in the effect of health beliefs on healthcare utilisation might explain parts of the well documented differences in healthcare seeking behavior between men and women, as men tend to have lower healthcare use (Galdas et al. 2005, Mansfield et al. 2003, Schlichthorst et al. 2016).

Second, reporting biases in health by countries in Europe are well documented (Capistrant et al. 2014, Jürges 2007, Spitzer & Weber 2019). To ensure that our finding are not driven by such differential reporting due to cultural biases in reporting health and oversampling of certain countries in the SHARE survey, we re-run our analysis for each country separately. By and large, we find similar results for all countries with the exception of a few where we do not find statistical significance due to the small sample sizes (Tables 8 and 9).

Finally, an important concern stems from the fact that individuals that underestimate or overestimate their health may well have health differences such that it is health that drives their utilisation. The descriptive statistics in Table 6 indicate a slight decrease of concordance with the number of chronic diseases, however, this trend is far from obvious and might also be due to the correlation between health and age. To disentangle these effects, we run separate regressions for those individuals that do not have any chronic diseases in wave $w + 1$ (healthy) and those that report one or more chronic diseases in wave $w + 1$ (unhealthy). The results are reported in Table 10. Marginal effects show that there is no substantial difference between the healthy and the unhealthy sub-samples with respect to the relationship between over- or underestimation and doctor visits. Since we categorise based on health, in other words, standardise for health, we can conclude that the results are not driven by objective health differences – both the healthy and the unhealthy group's healthcare utilisation is affected by their health perception in the same direction and similar magnitude.

4.3 Robustness analyses

We conduct a range of robustness analyses to observe whether our results are sensitive to model specifications and sample composition. These results are presented in Tables 11 and 12 along with the original model specification (Column 1). First, we utilise different income variables. We exchange the first imputed income variable provided by SHARE with the second imputed income variable (Column 2) and we use income that was not normalised with the cube root method but only equivalised (Column 3). These adjustments no effects on the results.

Second, we use a different specifications of our main explanatory variable health perception. For this, all individuals that felt unsafe to try the test in the first place are excluded from the sample (Column 4). This modification has no impact on the unimpaired sample and thus does not alter the estimated coefficients for underestimating health. However, this robustness analysis reduces the sample of the impaired to 957 observation only and the estimated coefficient for overestimating does not appear significant anymore.

Third, we separate the sample by survey wave to explore whether the slight change in the phrasing of the survey question for doctor visits in Wave 6 (Section 2.2.1) or the restriction of the chair stand test to those younger than 76 years in Wave 2 (Section 2.3) affect the results. The estimates in Table 13 reveal that the effect of health misperception on healthcare utilisation is slightly stronger in Wave 5 than in Wave 2. If the effect of underestimating health on health care utilisation is stronger for the elderly, this could potentially explain the different results, since the sample in Wave 2 is younger.

4.4 Additional measures of health perception

For the main analysis, health perception was operationalised based on tested and self-reported ability to stand up from a chair. In this section, we analyse whether the results hold for other health dimensions, in particular, health perception concerning cognition and walking ability.

4.4.1 Cognition

Similar to previous work, we use the deviation between subjective and objective cognition as an additional measure of health perception (Spitzer & Weber 2019). Objective cognition is operationalised based on a memory test, which is conducted in waves 4 to 6. In particular, individuals are asked to recall a list of 10 words in any order within a minute.

Subjective cognition is based on the question "How would you rate your memory at the present time?" which is answered on a Likert scale with categories "excellent", "very good", "good", "fair", and "poor". Since the subjective cognition variable has more than 80% missing values in Wave 6, we only utilise waves 4 and 5. Hence, the estimates for cognition are based on a different sample. For the main results presented in Section 4, health perception from Waves 2 and 5 were matched with health care utilisation from Waves 4 and 6. For the results on cognition, health perception from Waves 4 and 5 is matches with healthcare utilisation from Waves 5 and 6.

Defining impairment for cognition is not as straightforward as it is for the ability to stand up from a chair. While the chair stand variables are binary and therefore clearly indicate whether an individual is impaired or not, both the subjective and the objective cognition variables are categorical. Thus, we rely on previous literature to define the threshold marking cognitive impairment. Participants are considered objectively impaired if they recall only three words or less (Grodstein et al. 2001, Purser et al. 2005) . Additionally, in robustness analyses, individuals are considered impaired if they recall only two or fewer words. Individuals are considered subjectively impaired if they report fair or poor memory (Gardner et al. 2017).

Tables 14 provides regression results for the new specification of health perception. The results confirm our earlier findings. Individuals that underestimate their cognitive ability at wave w are more likely to visit the doctor at wave w+1 than individuals that achieve concordance between objective and subjective memory measures. By contrast, survey participants that overestimate their health have lower annual doctor visits than those that achieve concordance. Modifying the threshold for objective impairment from three to two words changes the magnitude of the coefficient for overestimating, but not its sign. The magnitude of the coefficient for overestimating remains virtually identical.

4.4.2 Walking ability

We also operationalise health perception based on walking ability. Objective walking ability is based on a walking speed test for which participants have to walk a distance of two and a half meters. Individuals are considered to be objectively impaired if their walking speed is 0.4

meters per second or slower. This threshold is in line with previous literature (Jürges 2007, Steel et al. 2003). Since the test is only conducted in waves 1 and 2, the analysis is restricted to those waves (Börsch-Supan 2019*a*). The walking speed test is supposed to be conducted only for individuals older than 75 years. However, the dataset includes information for those aged 75 and younger too. The variable has many missing values (∼90%) and thus needs to be handled with caution.

Subjective walking impairment is based on the question "Please look at card [...]. We need to understand difficulties people may have with various activities because of a health or physical problem. Please tell me whether you have any difficulty doing each of the everyday activities on card [...]. Exclude any difficulties that you expect to last less than three months". Participants are coded to have subjectively impaired walking ability if they report difficulties in walking for 100 meters.

When analysing health perception based walking ability, IADLs are not controlled for, since the ability to walk across a room is considered an IADL itself. Furthermore, risk aversion is not controlled for, since it was not covered in Wave 1. Also, the second imputed income variable is used for this analysis, since the first one was not available in Wave 1. The robustness analysis in Section 4.3 has shown, however, that both variables lead to the same results.

Results for the effect of health perception on the annual number of doctor visits based on walking ability measures are provided in Table 15. The coefficients in Table 15 confirm once again that individuals that underestimate their health have higher annual doctor visits than those that assess their health correctly. The results also show that those who overestimate their health have lower doctor visits. Thus, our results are robust to different specifications of health perception.

5 Conclusion

We utilised rich longitudinal data for 15 European countries from the Survey of Health, Ageing and Retirement to explore the effect of health (mis)perception on healthcare utilisation. Our results based on count models and log-gamma models suggest that individuals who underestimate their health visits the doctor more often and have higher out-of-pocket expenditure than those who assess their health correctly. By contrast, survey participants who overestimate their health visit the doctor less often and have smaller out-of-pocket payments. Heterogeneity analyses by gender show that the the effects are larger in size for men that for women, indicating that health perception could be an important explanation for the well documented gender differences in healthcare seeking behaviour.

Our results are robust to a range of sensitivity analyses with different model specifications, sample compositions, estimation methods and health dimensions. In addition, we account for potential endogeneity problems by exploiting the panel structure of our data. The main limitation of this paper is related to panel attrition. Individuals that suffer from diseases are less likely to participate in consecutive survey waves and thus are less likely to be included in our sample. However, we addressed this limitation by running our analyses separately by the number of diseases that a participant is suffering from and found no difference in the results between the healthy and the unhealthy, indicating that panel attrition is no concern for our conclusion.

A natural question to ask next is what do these results mean for policy? First, if individuals' own perception of health is what drives healthcare demand beyond actual health and other socioeconomic characteristics, then equipping them with the necessary tools and information through personalised or public health campaigns to accurately assess own health and determine the need to seek healthcare is perhaps a valuable long term strategy to reduce unnecessary use. This is a particularly relevant measure for countries with an ageing population that suffers cognitive dissonance thereby increasing health misperception (Brandtstädter & Greve 1994, Frieswijk et al. 2004, Henchoz et al. 2008, Idler 1993, Spitzer & Weber 2019). Reaching out to those who overestimate their health by providing information about the benefits of screening and preventive care might in addition increase their health and thus prevent suffering and costs in the long run. Second, waiting time is often used as a rationing measure by policy makers. Reducing unnecessary visits to the doctor can have important implications for such rationing mechanisms to work effectively. Not only will it free up physician time, but can also directly ensure timely visits for other patients that are in need for urgent intervention. Finally, addressing rising health expenditures has been a top priority on policy makers' agenda across countries. Addressing sources of waste and inefficiency in healthcare either on the demand or the supply side is therefore important in this direction.

References

Anderson, A., Baker, F. & Robinson, D. T. (2017), 'Precautionary savings, retirement planning and misperceptions of financial literacy', *Journal of Financial Economics* **126**(2), 383–398.

Arni, P., Dragone, D., Goette, L. & Ziebarth, N. R. (2019), 'Biased Health Perceptions and Risky Health Behaviors — Theory and Evidence', *Preliminary Draft* .

Bago d'Uva, T. & Jones, A. M. (2009), 'Health care utilisation in Europe: New evidence from the ECHP', *Journal of Health Economics* **28**(2), 265–279.

Bago d'Uva, T., O'Donnell, O. & Van Doorslaer, E. (2008), 'Differential health reporting by education level and its impact on the measurement of health inequalities among older Europeans', *International Journal of Epidemiology* **37**(6), 1375–1383.

Baumann, A. O., Deber, R. B. & Thompson, G. G. (1991), 'Overconfidence among physicians and nurses: The 'micro-certainty, macro-uncertainty' phenomenon', *Social Science and Medicine* **32**(2), 167–174.

Beaudoin, M. & Desrichard, O. (2011), 'Are memory self-efficacy and memory performance related? A meta-analysis.', *Psychological Bulletin* **137**(2), 211–241.

Berner, E. S. & Graber, M. L. (2008), 'Overconfidence as a Cause of Diagnostic Error in Medicine', *American Journal of Medicine* **121**(5 SUPPL.), S2–S23.

Biro, A. (2013), 'Discount rates and the education gradient in mammography in the UK', *Health economics* (22), 1021–1036.

Bíró, A. (2014), 'Supplementary private health insurance and health care utilization of people aged 50+', *Empirical Economics* **46**(2), 501–524.

Biró, A. (2016*a*), 'Differences between subjective and predicted survival probabilities and their relation to preventive care use', *B.E. Journal of Economic Analysis and Policy* **16**(2), 807–835.

Bíró, A. (2016*b*), 'Outpatient visits after retirement in Europe and the US', *International Journal of Health Economics and Management* **16**(4), 363–385.

Bolin, K., Lindgren, A., Lindgren, B. & Lundborg, P. (2009), 'Utilisation of physician services in the 50+ population: The relative importance of individual versus institutional factors in 10 European countries', *International Journal of Health Care Finance and Economics* **9**(1), 83–112.

Börsch-Supan, A. (2019*a*), 'Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 1. Release version: 7.0.0. SHARE-ERIC. Data set.'.
URL: *10.6103/SHARE.w1.700*

Börsch-Supan, A. (2019*b*), 'Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 2. Release version: 7.0.0. SHARE-ERIC. Data set.'.
URL: *10.6103/SHARE.w2.700*

Börsch-Supan, A. (2019*c*), 'Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 4. Release version: 7.0.0. SHARE-ERIC. Data set.'.
URL: *10.6103/SHARE.w4.700*

Börsch-Supan, A. (2019*d*), 'Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 5. Release version: 7.0.0. SHARE-ERIC. Data set.'.
URL: *10.6103/SHARE.w5.700*

Börsch-Supan, A. (2019*e*), 'Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 6. Release version: 7.0.0. SHARE-ERIC. Data set.'.
URL: *10.6103/SHARE.w6.700*

Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., Korbmacher, J., Malter, F., Schaan, B., Stuck, S. & Zuber, S. (2013), 'Data resource profile: The Survey of Health, Ageing and Retirement in Europe (SHARE)', *International Journal of Epidemiology* **42**(4), 992–1001.

Brandtstädter, J. & Greve, W. (1994), 'The aging self: Stabilizing and protecitve processes', *Developmental Review* (14), 52–80.

Cameron, K. A., Song, J., Manheim, L. M. & Dunlop, D. D. (2010), 'Gender disparities in health and healthcare use among older adults', *Journal of Women's Health* **19**(9), 1643–1650.

Capistrant, B. D., Glymour, M. M. & Berkman, L. F. (2014), 'Assessing mobility difficulties for cross-national comparisons: Results from the WHO study on AGEing and Adult Health', *Journal of the American Geriatrics Society* **62**(2), 329–335.

Coman, L. & Richardson, J. (2006), 'Relationship between self-report and performance measures of function: a systematic review.', *Canadian journal on aging* **25**(3), 253–270.

Cox, N. (2011), 'Stata tip 96: Cube roots', *Stata Journal* **11**(1), 149–154.

Crossley, T. F. & Kennedy, S. (2001), 'The reliability of self-assessed health status.', *Journal of Health Economics* **21**, 643–658.

Cutler, D., Skinner, J., Stern, A. D. & Wennberg, D. (2013), 'Physician beliefs and patient preferences : A new look at regional variation in health care spending', *NBER Working Paper Series* (19320).

Eurostat (2018), 'International Standard Classification of Education (ISCED)'.
URL: *http://ec.europa.eu/eurostat/statistics-explained/index.php/International_Standard_Classification_of_Education_%28ISCED%29#Correspondence_between_ISCED_2011_and_ISCED_1997*

Frieswijk, N., Buunk, B. P., Steverink, N. & Slaets, J. P. (2004), 'The effect of social comparison information on the life satisfaction of frail older persons', *Psychology and Aging* **19**(1), 183–190.

Furnham, A. (2001), 'Self-estimates of intelligence : culture and gender difference in self and other estimates of both general (g) and multiple intelligences', *Personality and Individual Differences* **31**, 1381–1405.

Galdas, P. M., Cheater, F. & Marshall, P. (2005), 'Men and health help-seeking behaviour: Literature review', *Journal of Advanced Nursing* **49**(6), 616–623.

Gardner, R. C., Langa, K. M. & Yaffe, K. (2017), 'Subjective and objective cognitive function among older adults with a history of traumatic brain injury: A population-based cohort study', *PLoS Medicine* **14**(3), 1–16.

Grodstein, F., Chen, J., Wilson, R. S. & Manson, J. E. (2001), 'Type 2 diabetes and cognitive function in community-dwelling elderly women', *Diabetes Care* **24**(6), 1060–1065.

Henchoz, K., Cavalli, S. & Girardin, M. (2008), 'Health perception and health status in advanced old age: A paradox of association', *Journal of Aging Studies* **22**(3), 282–290.

Idler, E. L. (1993), ‘Age differences in self-assessments of health: Age changes, cohort differences, or survivorship?’, *Journal of Gerontology* **48**(6), 289–300.

Jackson, J. D., Rentz, D. M., Aghjayan, S. L., Buckley, R. F., Meneide, T. F., Sperling, R. A. & Amariglio, R. E. (2017), ‘Subjective cognitive concerns are associated with objective memory performance in Caucasian but not African-American persons’, *Age and Ageing* **46**(6), 988–993.

Jürges, H. (2007), ‘True health vs response style: exploring cross-country differences in self-reported health’, *Health economics* **16**, 163–178.

Jürges, H. (2015), Health care utilization and out-of-pocket expenses, *in* F. Malter & A. Börsch-Supan, eds, ‘SHARE Wave 5 : Innovations & Methodology’, Munich Center for the Economics of Ageing (MEA) at the Max Planck Institute for Social Law and Socia, Munich, pp. 37–42.

Lugo-Palacios, D. G. & Gannon, B. (2017), ‘Health care utilisation amongst older adults with sensory and cognitive impairments in Europe’, *Health Economics Review* **7**(1), 1–15.

Mansfield, A., Addis, M. & Mahalik, J. (2003), ‘"Why Won’t He Go to the Doctor?": The Psychology of Men’s Help Seeking’, *International Journal of Men’s Health* **2**(2), 93–109.

Merrill, S. S., Seeman, T. E., Kasl, S. V. & Berkman, L. F. (1997), ‘Gender differences in the comparison of self-reported disability and performance measures’, *Journals of Gerontology - Series A Biological Sciences and Medical Sciences* **52**(1), 19–26.

Meyer, A. N., Payne, V. L., Meeks, D. W., Rao, R. & Singh, H. (2013), ‘Physicians’ diagnostic accuracy, confidence, and resource requests: A vignette study’, *JAMA Internal Medicine* **173**(21), 1952–1961.

Moore, D. A. & Healy, P. J. (2008), ‘The Trouble With Overconfidence’, *Psychological Review* **115**(2), 502–517.

Ortoleva, P. & Snowberg, E. (2015), ‘Overconfidence in political behavior’, *American Economic Review* **105**(2), 504–535.

Paccagnella, O., Rebba, V. & Weber, G. (2013), ‘VOLUNTARY PRIVATE HEALTH INSURANCE AMONG THE OVER 50s IN EUROPE† OMAR’, *Health economics* **22**.

Picone, G., Sloan, F. & Taylor, D. (2004), ‘Effects of risk and time preference and expected longevity on demand for medical tests’, *Journal of Risk and Uncertainty* **28**(1), 39–53.

Preston, C. E. & Harris, S. (1965), ‘Psychology of drivers in traffic accidents’, *Journal of Applied Psychology* **49**(4), 284–288.

Purser, J. L., Fillenbaum, G. G., Pieper, C. F. & Wallace, R. B. (2005), ‘Mild cognitive impairment and 10-year trajectories of disability in the Iowa established populations for epidemiologic studies of the elderly cohort’, *Journal of the American Geriatrics Society* **53**(11), 1966–1972.

Reuben, E., Wiswall, M. & Zafar, B. (2017), 'Preferences and Biases in Educational Choices and Labour Market Expectations: Shrinking the Black Box of Gender', *Economic Journal* **127**(604), 2153–2186.

Sakurai, R., Fujiwara, Y., Ishihara, M., Higuchi, T., Uchida, H. & Imanaka, K. (2013), 'Age-related self-overestimation of step-over ability in healthy older adults and its relationship to fall risk', *BMC Geriatrics* **13**(44), 1–9.

Schlichthorst, M., Sanci, L. A., Pirkis, J., Spittal, M. J. & Hocking, J. S. (2016), 'Why do men go to the doctor? Socio-demographic and lifestyle factors associated with healthcare utilisation among a cohort of Australian men', *BMC Public Health* **16**(Suppl 3), 81–90.
URL: *http://dx.doi.org/10.1186/s12889-016-3706-5*

Schneider, U., Pfarr, C., Schneider, B. S. & Ulrich, V. (2012), 'I feel good! Gender differences and reporting heterogeneity in self-assessed health', *European Journal of Health Economics* **13**(3), 251–265.

Spitzer, S. & Weber, D. (2019), 'Who is telling the truth? Biases in self-reported physical and cognitive health status of older Europeans', *IIASA Working Paper* **WP-19-002**.

Srisurapanont, M., Suttajit, S., Eurviriyanukul, K. & Varnado, P. (2017), 'Discrepancy between objective and subjective cognition in adults with major depressive disorder', *Scientific Reports* **7**(1), 1–7.

Steel, N., Huppert, F. & Melzer, D. (2003), Physical and cognitive function, *in* M. Marmot, J. Banks, R. Blundell, C. Lessof & J. Nazroo, eds, 'Health, Wealth and Lifestyles of the Older Population in England: The 2002 English Longitudinal Study of Ageing', IFS, London, chapter Physical a.

Tavares, L. P. & Zantomio, F. (2017), 'Inequity in healthcare use among older people after 2008: The case of southern European countries', *Health Policy* **121**(10), 1063–1071.

Vallejo-Torres, L. & Morris, S. (2013), 'Income-related inequity in healthcare utilisation among individuals with cardiovascular disease in England - account for vertical inquity', *Health economics* (22), 533–553.

Van Doorslaer, E., Koolman, X. & Jones, A. M. (2004), 'Explaining income-related inequalities in doctor utilisation in Europe', *Health Economics* **13**(7), 629–647.

Zhang, Y., Salm, M. & Soest, A. V. (2018), 'The effect of retirement on healthcare utilization : Evidence from China', *Journal of Health Economics* **62**, 165–177.

Table 4: Crosstable mean doctor visits at $w + 1$ (weighted)

	Health perception			
	Concordance Mean doctor visits	**Overestimating** Mean doctor visits	**Underestimating** Mean doctor visits	**Total** Mean doctor visits
No. chronic diseases at w				
0	4.9	6.3	7.2	5.3
1	6.7	9.1	9.6	7.4
2	8.8	11.2	10.6	9.4
3	9.6	13.5	11.6	10.5
4	12.0	14.8	16.0	13.4
5	14.3	11.1	13.0	13.2
6	13.9	15.9	18.9	14.5
7	22.2	13.2	14.6	16.7
8	13.2	10.1	19.5	14.3
9	20.0			20.0
10	11.7			11.7
12	20.0			20.0
Total	6.8	9.2	10.1	7.6
No. activity limitations at w				
0	6.2	8.1	8.9	6.8
1	8.9	10.9	10.5	9.6
2	11.0	13.4	13.9	11.9
3	11.9	14.8	11.6	12.5
4	12.0	17.0	30.1	14.7
5	11.6	18.5	11.3	12.8
6	13.0	15.1	10.3	12.4
7	12.1	13.1	17.6	12.8
8	14.6	13.4	8.5	14.2
9	19.6	10.6	25.5	15.5
10	17.9	12.9	30.0	18.9
11	10.8	14.8	14.1	11.1
12	12.1	18.1	9.1	13.7
13	14.2	3.5	9.7	12.3
Total	6.8	9.2	10.1	7.6
5-year age groups				
50-54	5.4	7.6	10.4	6.0
55-59	5.9	7.3	10.7	6.4
60-64	6.4	9.4	9.1	6.9
65-69	7.3	9.5	9.6	7.8
70-74	8.2	9.5	10.9	8.7
75-79	8.9	12.3	9.0	9.6
80-84	9.2	10.7	11.4	9.9
85-89	9.5	12.5	11.5	9.9
90+	11.0	12.3	7.6	11.0
Total	6.8	9.2	10.1	7.6
Gender				
Men	6.5	8.7	9.9	7.2
Women	7.1	9.7	10.2	7.9
Total	6.8	9.2	10.1	7.6

Note: Calibrated cross-sectional individual weights are applied.

Table 4, continued: Crosstable mean doctor visits at $w + 1$ (weighted)

	Health perception			
	Concordance Mean doctor visits	**Overestimating** Mean doctor visits	**Underestimating** Mean doctor visits	**Total** Mean doctor visits
Education				
Low	7.4	10.1	10.3	8.3
Medium	6.7	8.5	9.7	7.3
High	5.9	7.5	10.8	6.6
Total	6.8	9.2	10.1	7.6
Is retired				
0	6.0	8.4	10.1	6.7
1	7.7	9.8	10.2	8.4
Total	6.8	9.2	10.1	7.6
Is married				
0	7.3	10.0	10.1	8.2
1	6.7	8.9	10.2	7.4
Total	6.8	9.2	10.1	7.6
Risk aversion				
Substantial	6.8	7.2	16.8	7.7
Above average	4.7	6.5	8.6	5.1
Average	6.0	7.6	10.3	6.6
No	7.1	9.3	10.1	7.9
Total	6.8	9.2	10.1	7.5
Country				
Austria	6.8	8.7	8.7	7.3
Germany	7.5	11.4	11.3	8.4
Sweden	4.2	5.4	6.1	4.4
Netherlands	5.4	8.3	7.2	5.7
Spain	5.9	8.5	8.5	6.7
Italy	8.1	10.5	13.3	9.3
France	6.0	7.2	7.8	6.6
Denmark	4.6	6.8	8.2	5.1
Switzerland	4.6	8.1	7.2	5.3
Belgium	7.8	10.1	10.4	8.5
Czechia	7.1	8.9	9.7	7.7
Poland	7.1	6.9	9.9	7.4
Luxembourg	8.9	12.0	11.1	9.4
Slovenia	5.0	7.4	7.3	5.4
Estonia	5.5	6.6	7.6	5.9
Total	6.8	9.2	10.1	7.6
Survey wave				
Wave 2	7.1	8.6	10.0	8.0
Wave 5	6.6	9.9	10.2	7.3
Total	6.8	9.2	10.1	7.6

Note: Calibrated cross-sectional individual weights are applied.

Table 5: Crosstable mean OOP expenditure in Euros at $w+1$ (weighted)

	Health perception			
	Concordance Mean OOP	**Overestimating** Mean OOP	**Underestimating** Mean OOP	**Total** Mean OOP
No. chronic diseases at w				
0	64.0	71.7	94.5	66.1
1	71.7	64.7	100.2	73.4
2	72.8	83.5	83.8	75.2
3	87.0	102.3	88.0	88.0
4	107.7	90.9	75.8	99.9
5	84.6	45.2	144.2	91.1
6	169.1	5.7	181.6	167.3
7	18.2	30.0	97.8	37.4
8	48.7	6.1	334.6	48.3
10	0.0			0.0
12	300.0			300.0
Total	72.1	75.0	93.7	74.4
No. activity limitations at w				
0	73.0	74.9	90.5	74.3
1	63.6	61.6	97.4	70.2
2	57.5	105.1	68.9	66.8
3	109.4	128.2	133.2	114.5
4	33.0	40.9	96.7	44.8
5	72.4	57.7	52.0	64.7
6	60.9	29.1	58.7	62.6
7	37.3	216.5	87.4	59.2
8	104.2	14.1	0.3	80.3
9	10.6	28.2	18.9	57.2
10	100.3	4.0	0.0	78.1
11	25.3	181.9	319.8	68.6
12	43.1	4.8	279.1	61.5
13	153.9	6.4	301.1	167.9
Total	72.1	75.0	93.7	74.3
5-year age groups				
50-54	60.4	43.5	143.1	64.7
55-59	67.7	100.8	81.0	70.6
60-64	70.9	74.9	78.0	71.4
65-69	75.4	86.1	83.6	77.7
70-74	88.2	64.2	100.9	87.5
75-79	77.7	87.5	95.7	80.6
80-84	61.8	83.2	89.5	67.7
85-89	60.5	59.6	75.8	65.2
90+	198.0	56.4	21.5	142.6
Total	72.1	75.0	93.7	74.3
Gender				
Men	75.6	63.6	70.2	73.9
Women	68.9	83.5	106.3	74.7
Total	72.1	75.0	93.7	74.3

Note: Calibrated cross-sectional individual weights are applied.

Table 5, continued: Crosstable mean OOP expenditure in Euros at $w + 1$ (weighted)

	Health perception			
	Concordance Mean OOP	**Overestimating** Mean OOP	**Underestimating** Mean OOP	**Total** Mean OOP
Education				
Low	64.0	67.0	92.7	67.2
Medium	71.2	70.5	91.5	73.3
High	90.2	119.7	108.9	93.1
Total	72.1	75.0	93.7	74.7
Is retired				
0	63.8	81.8	101.6	67.7
1	80.5	71.8	89.2	81.0
Total	72.1	75.0	93.7	74.6
Is married				
0	71.6	65.3	77.1	71.4
1	73.6	82.9	99.8	76.7
Total	72.1	75.0	93.7	74.9
Risk aversion				
Substantial	92.3	62.6	40.2	85.9
Above average	89.9	299.0	124.0	98.1
Average	89.1	101.0	111.6	91.6
No	66.1	68.9	89.4	68.7
Total	72.1	75.0	93.7	74.3
Country				
Austria	125.6	63.0	140.4	121.1
Germany	58.0	74.8	44.8	57.3
Sweden	68.1	66.4	77.0	68.7
Spain	18.4	14.3	40.7	19.8
Italy	137.2	134.7	266.9	146.4
France	28.1	30.0	36.7	29.1
Denmark	5.4	1.3	3.6	5.1
Switzerland	397.7	372.6	450.8	399.8
Belgium	94.1	91.1	181.4	104.0
Czechia	7.5	10.4	11.3	8.5
Luxembourg	177.4	194.7	207.9	181.4
Slovenia	13.3	3.8	19.8	12.9
Estonia	15.5	13.8	11.7	15.2
Total	72.1	75.0	93.7	74.3
Survey wave				
Wave 5	72.1	75.0	93.7	74.3
Total	72.1	75.0	93.7	74.3

Note: Calibrated cross-sectional individual weights are applied.

Table 6: Crosstable health perception (weighted)

	Health perception						
	Concordance		**Overestimating**		**Underestimating**		**Total**
	Row %	95% CI	Row %	95% CI	Row %	95% CI	Row %
Objective impairment							
Unimpaired (n=49,380)	87.9	[87.4,88.4]	0.0		12.1	[11.6,12.6]	100.0
Impaired (n=9,544)	40.6	[39.0,42.2]	59.4	[57.8,61.0]	0.0		100.0
Total (n=58,924)	79.7	[79.2,80.3]	10.2	[9.8,10.7]	10.0	[9.6,10.4]	100.0
Pearson: Uncorrected chi2(2) =	3.26e+04						
Design-based F(1.99, 117023.69) =	6375.6161	Pr =	0.000				
No. chronic diseases at w							
0 (n=21,298)	84.6	[83.8,85.4]	9.6	[8.9,10.3]	5.8	[5.3,6.3]	100.0
1 (n=18,578)	80.0	[79.0,81.0]	9.9	[9.2,10.7]	10.0	[9.3,10.8]	100.0
2 (n=10,799)	75.9	[74.4,77.2]	10.4	[9.4,11.5]	13.7	[12.6,14.9]	100.0
3 (n=5,097)	71.5	[69.4,73.5]	12.7	[11.2,14.4]	15.8	[14.2,17.5]	100.0
4 (n=2,071)	68.7	[65.1,72.0]	12.2	[9.9,14.9]	19.2	[16.4,22.3]	100.0
5 (n=733)	64.2	[57.7,70.3]	11.5	[8.1,15.9]	24.3	[19.0,30.6]	100.0
6 (n=204)	70.0	[59.2,78.9]	15.4	[8.7,26.0]	14.6	[9.0,22.8]	100.0
7 (n=56)	68.8	[47.2,84.5]	12.2	[4.2,30.6]	18.9	[7.6,39.9]	100.0
8 (n=22)	83.3	[55.9,95.1]	13.4	[3.3,41.0]	3.3	[0.5,20.1]	100.0
9 (n=1)	100.0		0.0		0.0		100.0
10 (n=2)	100.0		0.0		0.0		100.0
12 (n=1)	100.0		0.0		0.0		100.0
Total (n=58,862)	79.7	[79.2,80.3]	10.2	[9.8,10.7]	10.0	[9.6,10.5]	100.0
Pearson: Uncorrected chi2(22) =	1252.4362						
Design-based F(16.95, 997454.56) =	26.2923	Pr =	0.000				
No. activity limitations at w							
0 (n=48,842)	82.9	[82.3,83.5]	9.2	[8.8,9.7]	7.8	[7.4,8.2]	100.0
1 (n=4,998)	62.2	[59.9,64.5]	15.4	[13.7,17.3]	22.4	[20.5,24.5]	100.0
2 (n=1,837)	59.9	[56.1,63.5]	14.6	[12.1,17.3]	25.6	[22.5,28.9]	100.0
3 (n=975)	59.3	[54.0,64.4]	16.1	[12.5,20.6]	24.6	[20.4,29.3]	100.0
4 (n=586)	63.6	[56.3,70.3]	19.8	[15.0,25.8]	16.6	[11.5,23.4]	100.0
5 (n=434)	64.3	[56.3,71.6]	19.7	[13.7,27.4]	16.1	[11.3,22.4]	100.0
6 (n=310)	69.3	[60.0,77.3]	17.5	[11.2,26.2]	13.2	[8.3,20.4]	100.0
7 (n=253)	71.7	[61.0,80.4]	9.6	[5.4,16.5]	18.7	[11.3,29.5]	100.0
8 (n=157)	76.1	[64.5,84.7]	11.8	[6.2,21.5]	12.1	[6.1,22.6]	100.0
9 (n=149)	87.2	[77.6,93.1]	6.6	[2.9,14.3]	6.2	[2.4,14.8]	100.0
10 (n=100)	82.2	[69.1,90.5]	15.8	[8.1,28.4]	2.1	[0.3,13.3]	100.0
11 (n=63)	77.2	[53.8,90.7]	18.8	[6.4,44.1]	4.0	[1.2,12.9]	100.0
12 (n=68)	76.4	[60.2,87.4]	13.7	[5.6,29.9]	9.9	[4.0,22.6]	100.0
13 (n=146)	74.4	[64.4,82.3]	4.1	[1.4,11.5]	21.6	[14.5,30.9]	100.0
Total (n=58,918)	79.7	[79.2,80.3]	10.2	[9.8,10.7]	10.0	[9.6,10.4]	100.0
Pearson: Uncorrected chi2(26) =	2431.9614						
Design-based F(24.66, 1.45e+06) =	37.5124	Pr =	0.000				
5-year age groups							
50-54 (n=7,804)	84.1	[82.6,85.5]	8.8	[7.7,10.0]	7.1	[6.3,8.1]	100.0
55-59 (n=10,970)	82.3	[81.0,83.5]	8.9	[8.0,9.9]	8.8	[7.9,9.8]	100.0
60-64 (n=11,472)	80.9	[79.7,82.1]	8.5	[7.7,9.4]	10.5	[9.7,11.5]	100.0
65-69 (n=10,645)	78.6	[77.3,79.8]	10.7	[9.8,11.8]	10.6	[9.7,11.6]	100.0
70-74 (n=8,570)	76.6	[75.1,78.1]	11.1	[10.1,12.3]	12.2	[11.1,13.4]	100.0
75-79 (n=4,742)	75.3	[73.2,77.3]	11.3	[9.9,12.9]	13.4	[11.8,15.1]	100.0
80-84 (n=3,020)	71.5	[68.6,74.3]	15.8	[13.6,18.3]	12.6	[10.7,14.8]	100.0
85-89 (n=1,337)	68.7	[64.4,72.6]	20.2	[16.8,24.0]	11.2	[8.7,14.2]	100.0
90+ (n=364)	69.1	[61.0,76.1]	24.6	[18.0,32.5]	6.3	[3.9,10.2]	100.0
Total (n=58,924)	79.7	[79.2,80.3]	10.2	[9.8,10.7]	10.0	[9.6,10.4]	100.0
Pearson: Uncorrected chi2(16) =	701.4072						
Design-based F(15.08, 888608.22) =	15.9828	Pr =	0.000				
Gender							
Men (n=25,929)	82.5	[81.7,83.3]	10.1	[9.5,10.8]	7.4	[6.9,7.9]	100.0
Women (n=32,995)	77.3	[76.6,78.1]	10.3	[9.8,10.9]	12.3	[11.7,12.9]	100.0
Total (n=58,924)	79.7	[79.2,80.3]	10.2	[9.8,10.7]	10.0	[9.6,10.4]	100.0
Pearson: Uncorrected chi2(2) =	407.0583						
Design-based F(2.00, 117566.06) =	68.9117	Pr =	0.000				

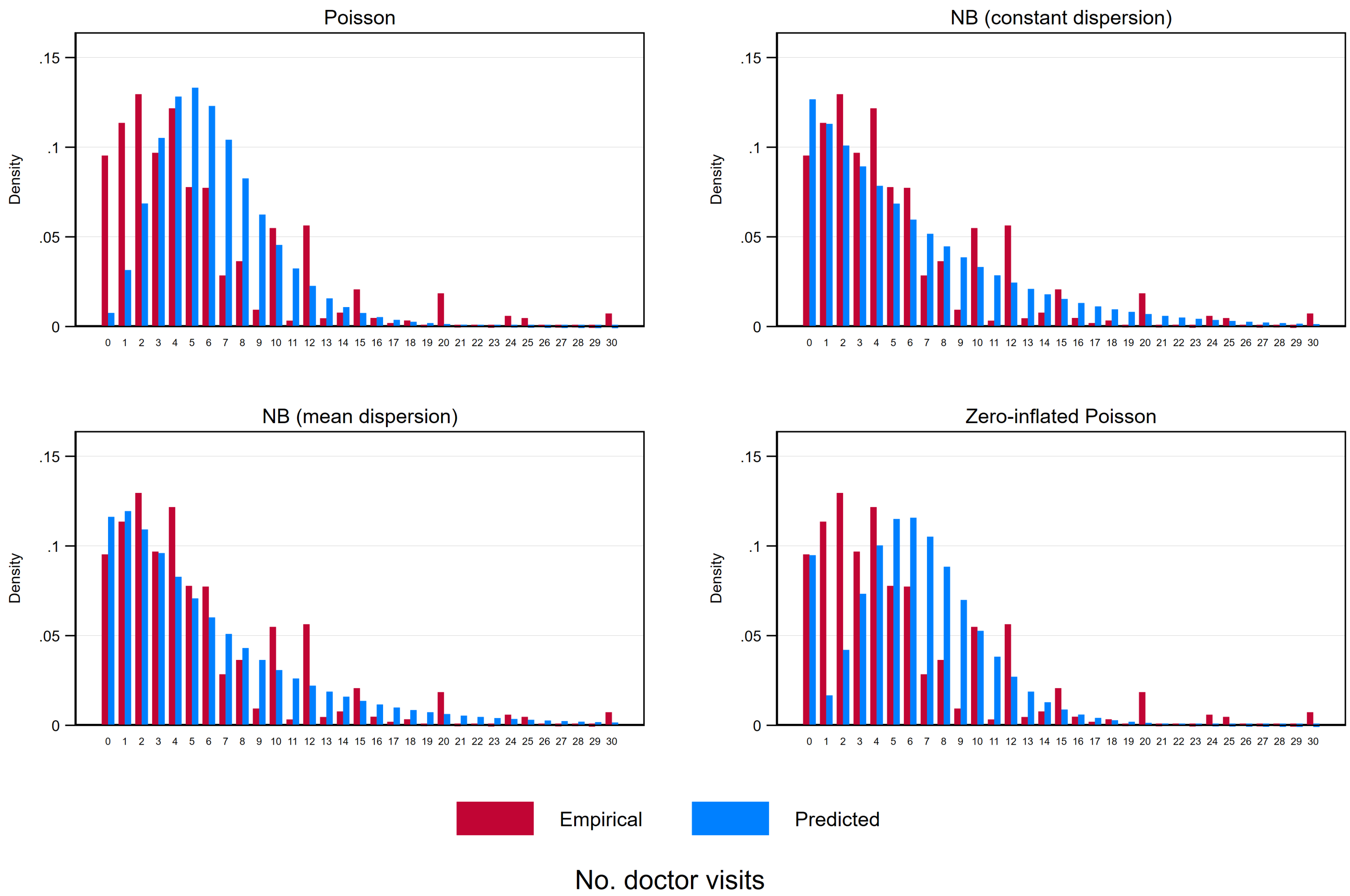
Note: Calibrated cross-sectional individual weights are applied.

Table 6, continued: Crosstable health perception (weighted)

	Health perception						
	Concordance		**Overestimating**		**Underestimating**		**Total**
	Row %	95% CI	Row %	95% CI	Row %	95% CI	Row %
Education							
Low (n=23,020)	76.4	[75.5,77.3]	12.8	[12.1,13.5]	10.8	[10.2,11.5]	100.0
Medium (n=21,963)	80.5	[79.6,81.5]	9.2	[8.5,9.9]	10.3	[9.6,11.0]	100.0
High (n=13,137)	85.6	[84.5,86.6]	6.7	[5.9,7.5]	7.8	[7.0,8.6]	100.0
Total (n=58,120)	79.8	[79.2,80.3]	10.2	[9.8,10.6]	10.0	[9.6,10.4]	100.0
Pearson: Uncorrected chi2(4) =	491.1841						
Design-based F(3.97, 230807.44) =	41.3845	Pr =	0.000				
Is retired							
0 (n=26,272)	82.0	[81.2,82.8]	9.3	[8.7,9.9]	8.8	[8.2,9.3]	100.0
1 (n=32,310)	77.5	[76.8,78.3]	11.1	[10.5,11.7]	11.3	[10.8,11.9]	100.0
Total (n=58,582)	79.8	[79.2,80.3]	10.2	[9.8,10.6]	10.0	[9.6,10.4]	100.0
Pearson: Uncorrected chi2(2) =	182.5597						
Design-based F(1.99, 116867.76) =	30.5986	Pr =	0.000				
Is married							
0 (n=15,680)	77.4	[76.2,78.5]	11.4	[10.5,12.4]	11.2	[10.4,12.1]	100.0
1 (n=42,048)	80.7	[80.1,81.3]	9.8	[9.3,10.3]	9.6	[9.1,10.0]	100.0
Total (n=57,728)	79.7	[79.1,80.2]	10.3	[9.8,10.7]	10.1	[9.7,10.5]	100.0
Pearson: Uncorrected chi2(2) =	83.6959						
Design-based F(2.00, 115173.49) =	12.7176	Pr =	0.000				
Risk aversion							
Substantial (n=595)	82.1	[76.2,86.8]	9.9	[6.8,14.2]	8.0	[4.6,13.4]	100.0
Above average (n=2,269)	85.9	[81.8,89.1]	7.3	[4.5,11.7]	6.8	[5.4,8.6]	100.0
Average (n=12,780)	84.1	[82.9,85.2]	6.8	[6.1,7.7]	9.1	[8.2,10.0]	100.0
No (n=42,224)	78.6	[77.9,79.2]	10.9	[10.4,11.5]	10.5	[10.0,11.0]	100.0
Total (n=57,868)	80.0	[79.4,80.5]	10.0	[9.5,10.4]	10.1	[9.7,10.5]	100.0
Pearson: Uncorrected chi2(6) =	253.6202						
Design-based F(4.59, 265461.56) =	12.4776	Pr =	0.000				
Country							
Austria (n=3,413)	79.5	[77.7,81.2]	9.2	[8.0,10.5]	11.3	[10.0,12.7]	100.0
Germany (n=5,391)	80.5	[79.3,81.8]	6.9	[6.1,7.7]	12.6	[11.6,13.7]	100.0
Sweden (n=4,802)	84.2	[83.0,85.3]	6.2	[5.4,7.1]	9.6	[8.7,10.6]	100.0
Netherlands (n=1,416)	86.9	[84.9,88.6]	4.7	[3.6,6.0]	8.4	[7.1,10.1]	100.0
Spain (n=5,995)	79.9	[78.3,81.5]	11.6	[10.4,13.0]	8.5	[7.5,9.6]	100.0
Italy (n=5,167)	76.5	[75.1,77.8]	15.3	[14.1,16.5]	8.3	[7.5,9.2]	100.0
France (n=4,531)	80.4	[79.1,81.8]	10.6	[9.5,11.8]	9.0	[8.1,9.9]	100.0
Denmark (n=4,602)	88.3	[87.3,89.2]	3.7	[3.1,4.3]	8.0	[7.3,8.9]	100.0
Switzerland (n=3,399)	86.2	[84.9,87.4]	6.3	[5.5,7.2]	7.5	[6.6,8.5]	100.0
Belgium (n=5,919)	82.0	[80.9,83.1]	6.8	[6.1,7.6]	11.2	[10.3,12.1]	100.0
Czechia (n=5,315)	79.4	[77.6,81.2]	10.4	[9.0,12.0]	10.1	[9.0,11.4]	100.0
Poland (n=1,274)	70.9	[68.1,73.5]	15.1	[13.1,17.4]	14.0	[12.1,16.2]	100.0
Luxembourg (n=1,065)	79.1	[76.3,81.6]	8.5	[6.7,10.6]	12.4	[10.5,14.7]	100.0
Slovenia (n=2,321)	79.2	[77.1,81.0]	10.2	[8.8,11.9]	10.6	[9.3,12.1]	100.0
Estonia (n=4,314)	77.1	[75.7,78.4]	9.6	[8.7,10.6]	13.3	[12.3,14.5]	100.0
Total (n=58,924)	79.7	[79.2,80.3]	10.2	[9.8,10.7]	10.0	[9.6,10.4]	100.0
Pearson: Uncorrected chi2(28) =	987.6800						
Design-based F(13.63, 802984.75) =	28.8396	Pr =	0.000				
Survey wave							
Wave 2 (n=15,205)	77.8	[76.8,78.8]	12.3	[11.6,13.1]	9.9	[9.2,10.6]	100.0
Wave 5 (n=43,719)	81.1	[80.4,81.7]	8.8	[8.3,9.3]	10.2	[9.7,10.7]	100.0
Total (n=58,924)	79.7	[79.2,80.3]	10.2	[9.8,10.7]	10.0	[9.6,10.4]	100.0
Pearson: Uncorrected chi2(2) =	195.1308						
Design-based F(2.00, 117821.25) =	31.9995	Pr =	0.000				

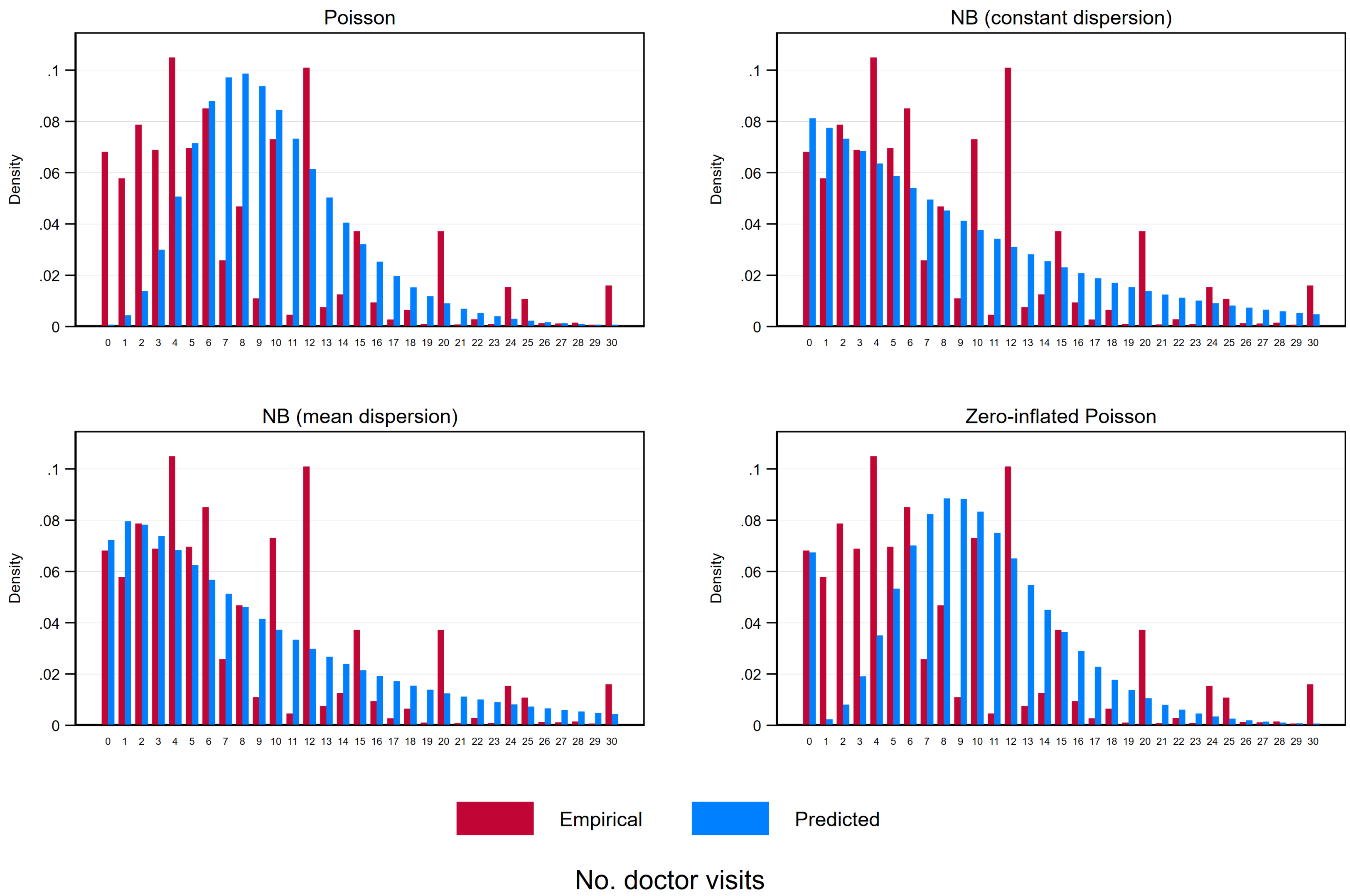
Note: Calibrated cross-sectional individual weights are applied.

Figure 3: Count model comparison for the annual number of doctor visits in the unimpaired sample, i.e. able to stand up from the chair



Note: Doctor visits are top coded at 98 visits by year. This figure shows only the first 30 doctor visits for better visualisation. Red bars represent the empirically observed numbers of doctor visits and blue bars represent the predicted values based on the respective count model.

Figure 4: Count model comparison for the annual number of doctor visits in the impaired sample, i.e. unable to stand up from the chair



Note: Doctor visits are top coded at 98 visits by year. This figure shows only the first 30 doctor visits for better visualisation. Red bars represent the empirically observed numbers of doctor visits and blue bars represent the predicted values based on the respective count model.

Table 7: Annual number of doctor visits at $w+1$ by gender

	(1) Unimpaired Men Doctor visits	(2) Impaired Men Doctor visits	(3) Unimpaired Women Doctor visits	(4) Impaired Women Doctor visits
Health perception (ref.: concordance)				
Underestimating	0.276***		0.232***	
	(0.035)		(0.020)	
Overestimating		-0.161***		-0.142***
		(0.044)		(0.033)
Chronic diseases	0.190***	0.162***	0.175***	0.118***
	(0.008)	(0.013)	(0.007)	(0.011)
Activity limitations	0.082***	0.027**	0.095***	0.035***
	(0.015)	(0.009)	(0.010)	(0.006)
Age	0.027	0.034	-0.019	0.024
	(0.017)	(0.028)	(0.013)	(0.020)
Age squared	-0.000	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Educ. group (ref.: medium)				
Low	-0.038	0.078	0.023	-0.013
	(0.025)	(0.049)	(0.020)	(0.039)
High	-0.028	-0.079	0.006	-0.094
	(0.024)	(0.056)	(0.022)	(0.060)
Retired	0.035	0.041	0.028	-0.022
	(0.031)	(0.066)	(0.021)	(0.037)
Married	-0.036	0.030	-0.037*	-0.012
	(0.025)	(0.051)	(0.018)	(0.033)
Equiv. hh income (cube root)	-0.000	-0.002	-0.001	0.000
	(0.001)	(0.002)	(0.001)	(0.002)
Risk aversion (ref.: no risk)				
Substantial	0.061	0.203	0.106	-0.500**
	(0.088)	(0.161)	(0.090)	(0.164)
Above average	-0.119**	0.025	-0.152**	-0.162
	(0.041)	(0.146)	(0.053)	(0.159)
Average	0.008	0.011	-0.022	-0.110*
	(0.022)	(0.059)	(0.020)	(0.049)
Wave 5	-0.113***	0.062	-0.072***	-0.104*
	(0.022)	(0.055)	(0.019)	(0.046)
Constant	0.414	0.681	2.288***	1.501*
	(0.562)	(0.986)	(0.435)	(0.719)
Control variables country	Yes	Yes	Yes	Yes
N	21,557	3,355	25,820	5,425
Pseudo R2	0.023	0.021	0.025	0.020
AIC	121,792	21,782	147,255	35,458
BIC	122,032	21,965	147,500	35,656
SE	cluster	cluster	cluster	cluster

Note: "Unimpaired" refers to the sample that is objectively unimpaired, i.e. able to stand up from the chair and "Impaired" refers to the sample that is objectively impaired, i.e. unable to stand up from the chair. The dependent variable "doctor visits" is based on the annual number of doctor visits, visits to emergency rooms and outpatient clinic visits at wave $w+1$, i.e. Wave 4 or Wave 6. All explanatory variables are taken from wave w, i.e. Wave 2 or Wave 5 respectively. The estimated coefficients are based on a negative binomial regression model with mean dispersion. Standard errors are clustered at the household level and presented in parentheses. * p<0.05, ** p<0.01, *** p<0.001

Table 8: Underestimating health and annual number of doctor visits at $w+1$ by country

	(1) Austria	(2) Belgium	(3) Czechia	(4) Denmark	(5) Estonia	(6) France	(7) Germany	(8) Italy
Health perception (ref.: concordance)								
Underestimating	0.168**	0.198***	0.194***	0.503***	0.240***	0.147**	0.315***	0.217***
	(0.063)	(0.049)	(0.045)	(0.087)	(0.070)	(0.048)	(0.053)	(0.062)
Chronic diseases	0.176***	0.156***	0.180***	0.207***	0.228***	0.173***	0.187***	0.211***
	(0.020)	(0.016)	(0.014)	(0.019)	(0.020)	(0.014)	(0.015)	(0.018)
Activity limitations	0.081**	0.150***	0.045	0.075*	0.108***	0.090***	0.105***	0.036*
	(0.029)	(0.026)	(0.027)	(0.035)	(0.026)	(0.025)	(0.030)	(0.014)
Age	-0.003	-0.039	0.031	0.017	0.002	-0.040	-0.053	0.063
	(0.040)	(0.026)	(0.035)	(0.036)	(0.040)	(0.027)	(0.027)	(0.034)
Age squared	0.000	0.000	-0.000	-0.000	-0.000	0.000	0.000*	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Woman	0.087	0.111**	0.048	-0.003	-0.042	0.054	0.047	0.176***
	(0.051)	(0.036)	(0.035)	(0.047)	(0.047)	(0.034)	(0.036)	(0.041)
Educ. group (ref.: medium)								
Low	-0.072	0.037	-0.011	0.017	-0.042	0.035	-0.027	0.033
	(0.061)	(0.046)	(0.036)	(0.070)	(0.057)	(0.038)	(0.058)	(0.055)
High	-0.005	-0.010	0.006	0.007	0.047	0.041	-0.024	-0.153
	(0.055)	(0.048)	(0.051)	(0.050)	(0.057)	(0.052)	(0.038)	(0.092)
Retired	0.091	0.064	-0.015	0.008	0.040	0.015	0.089	0.012
	(0.063)	(0.045)	(0.066)	(0.067)	(0.066)	(0.048)	(0.053)	(0.058)
Married	-0.027	-0.045	0.064	-0.127*	-0.034	-0.020	-0.017	0.002
	(0.052)	(0.041)	(0.038)	(0.051)	(0.050)	(0.039)	(0.044)	(0.056)
Equiv. hh income (cube root)	-0.000	-0.003	-0.001	-0.004	0.012	-0.002	-0.001	-0.001
	(0.002)	(0.002)	(0.003)	(0.003)	(0.007)	(0.002)	(0.002)	(0.002)
Risk aversion (ref.: no risk)								
Substantial	0.075	0.219	0.046	-0.232*	0.094	0.263	0.148	0.129
	(0.181)	(0.223)	(0.140)	(0.115)	(0.149)	(0.361)	(0.228)	(0.217)
Above average	0.195	-0.039	0.003	-0.095	-0.102	-0.182*	-0.150	0.033
	(0.174)	(0.105)	(0.115)	(0.073)	(0.163)	(0.089)	(0.108)	(0.212)
Average	-0.015	-0.055	-0.013	0.008	-0.144	-0.036	0.033	-0.129**
	(0.063)	(0.039)	(0.037)	(0.048)	(0.080)	(0.039)	(0.040)	(0.048)
Constant	1.156	2.712**	0.274	0.852	1.076	2.887**	3.394***	-0.882
	(1.336)	(0.874)	(1.175)	(1.231)	(1.348)	(0.932)	(0.887)	(1.133)
Control variables wave	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	2,734	4,947	4,056	4,148	3,242	3,653	4,668	3,830
Pseudo R2	0.017	0.018	0.019	0.020	0.021	0.019	0.021	0.020
AIC	16,334	29,867	24,257	21,503	17,832	19,929	28,072	23,799
BIC	16,435	29,978	24,364	21,610	17,929	20,035	28,181	23,905
SE	cluster	cluster	cluster	cluster	cluster	cluster	cluster	cluster

Note: The coefficients are based on the sample that is objectively unimpaired, i.e. able to stand up from the chair. The dependent variable "doctor visits" is based on the annual number of doctor visits, visits to emergency rooms and outpatient clinic visits at wave $w+1$, i.e. Wave 4 or Wave 6. All explanatory variables are taken from wave w, i.e. Wave 2 or Wave 5 respectively. The estimated coefficients are based on a negative binomial regression model with mean dispersion. Standard errors are clustered at the household level and presented in parentheses. * p<0.05, ** p<0.01, *** p<0.001

Table 8, continued: Underestimating health and annual number of doctor visits at $w+1$ by country

	(1) Luxembourg	(2) Netherlands	(3) Poland	(4) Slovenia	(5) Spain	(6) Sweden	(7) Switzerland
Health perception (ref.: concordance)							
Underestimating	0.051	0.325	0.123	0.284***	0.166**	0.329***	0.234**
	(0.101)	(0.166)	(0.081)	(0.079)	(0.055)	(0.077)	(0.080)
Chronic diseases	0.108***	0.150***	0.249***	0.223***	0.189***	0.119***	0.212***
	(0.030)	(0.044)	(0.026)	(0.028)	(0.018)	(0.022)	(0.025)
Activity limitations	0.228***	0.029	0.020	0.016	0.028	0.159***	0.227**
	(0.047)	(0.067)	(0.029)	(0.052)	(0.021)	(0.033)	(0.069)
Age	-0.094	-0.016	0.172	-0.008	0.017	0.027	0.008
	(0.072)	(0.126)	(0.103)	(0.056)	(0.033)	(0.040)	(0.045)
Age squared	0.001	0.000	-0.001	0.000	-0.000	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Woman	0.091	-0.061	0.075	-0.061	0.090*	-0.115*	0.006
	(0.091)	(0.088)	(0.070)	(0.061)	(0.044)	(0.050)	(0.056)
Educ. group (ref.: medium)							
Low	-0.231*	0.325**	-0.048	0.002	0.013	-0.112	-0.013
	(0.095)	(0.101)	(0.074)	(0.067)	(0.074)	(0.066)	(0.076)
High	-0.350**	0.250*	-0.072	-0.056	0.073	-0.019	0.025
	(0.123)	(0.116)	(0.104)	(0.082)	(0.098)	(0.059)	(0.075)
Retired	0.075	-0.171	0.096	-0.010	0.053	0.099	0.172*
	(0.100)	(0.103)	(0.086)	(0.087)	(0.049)	(0.090)	(0.077)
Married	-0.114	0.003	0.132	0.062	0.054	-0.021	-0.154*
	(0.109)	(0.105)	(0.095)	(0.071)	(0.052)	(0.061)	(0.069)
Equiv. hh income (cube root)	0.006*	-0.002	0.004	0.003	-0.000	-0.006	-0.003
	(0.003)	(0.005)	(0.007)	(0.005)	(0.003)	(0.005)	(0.002)
Risk aversion (ref.: no risk)							
Substantial	-0.310	0.160	-0.603**	0.217	-0.166	0.108	0.488
	(0.283)	(0.302)	(0.229)	(0.188)	(0.316)	(0.161)	(0.279)
Above average	-0.293	-0.001	-0.038	0.014	-0.104	-0.175**	0.033
	(0.337)	(0.282)	(0.301)	(0.190)	(0.184)	(0.059)	(0.151)
Average	-0.036	0.076	0.182	0.111	-0.137*	-0.010	0.023
	(0.108)	(0.134)	(0.128)	(0.086)	(0.069)	(0.064)	(0.057)
Constant	5.075*	1.714	-4.294	1.043	0.560	0.110	1.113
	(2.367)	(3.957)	(3.186)	(1.853)	(1.142)	(1.384)	(1.555)
Control variables wave	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	898	1,232	911	1,863	4,090	4,147	2,958
Pseudo R2	0.023	0.008	0.023	0.022	0.020	0.020	0.019
AIC	5,543	6,839	5,520	9,883	22,191	20,857	15,581
BIC	5,620	6,921	5,597	9,971	22,298	20,965	15,683
SE	cluster	cluster	cluster	cluster	cluster	cluster	cluster

Note: The coefficients are based on the sample that is objectively unimpaired, i.e. able to stand up from the chair. The dependent variable "doctor visits" is based on the annual number of doctor visits, visits to emergency rooms and outpatient clinic visits at wave $w+1$, i.e. Wave 4 or Wave 6. All explanatory variables are taken from wave w, i.e. Wave 2 or Wave 5 respectively. The estimated coefficients are based on a negative binomial regression model with mean dispersion. Standard errors are clustered at the household level and presented in parentheses. * p<0.05, ** p<0.01, *** p<0.001

Table 9: Overestimating health and annual number of doctor visits at $w+1$ by country

	(1) Austria	(2) Belgium	(3) Czechia	(4) Denmark	(5) Estonia	(6) France	(7) Germany	(8) Italy
Health perception (ref.: concordance)								
Overestimating	-0.218	-0.158	-0.160**	0.056	-0.149*	-0.140	-0.045	-0.112
	(0.113)	(0.081)	(0.062)	(0.149)	(0.071)	(0.092)	(0.106)	(0.080)
Chronic diseases	0.199***	0.120***	0.125***	0.243***	0.149***	0.067*	0.102**	0.198***
	(0.035)	(0.025)	(0.022)	(0.036)	(0.024)	(0.030)	(0.033)	(0.027)
Activity limitations	0.035	0.084***	0.015	0.118***	0.005	0.066***	0.025	0.020
	(0.018)	(0.017)	(0.013)	(0.031)	(0.016)	(0.020)	(0.017)	(0.012)
Age	0.028	0.011	0.102*	-0.161*	0.079	-0.019	-0.092	0.108*
	(0.068)	(0.053)	(0.042)	(0.081)	(0.053)	(0.054)	(0.068)	(0.043)
Age squared	-0.000	-0.000	-0.001*	0.001	-0.001	0.000	0.001	-0.001*
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Woman	-0.018	0.110	-0.049	0.121	0.035	-0.011	0.022	0.189**
	(0.122)	(0.084)	(0.063)	(0.153)	(0.069)	(0.078)	(0.109)	(0.068)
Educ. group (ref.: medium)								
Low	0.109	0.121	-0.004	-0.018	0.008	0.032	-0.155	0.214*
	(0.108)	(0.099)	(0.061)	(0.183)	(0.076)	(0.097)	(0.138)	(0.098)
High	0.291	0.012	0.040	-0.198	-0.071	0.106	-0.168	0.198
	(0.162)	(0.104)	(0.120)	(0.164)	(0.097)	(0.127)	(0.128)	(0.170)
Retired	0.073	-0.026	-0.031	-0.061	-0.121	0.257*	0.236	-0.057
	(0.122)	(0.089)	(0.105)	(0.183)	(0.111)	(0.114)	(0.145)	(0.076)
Married	-0.078	0.124	0.054	0.139	0.178**	0.060	-0.017	-0.032
	(0.103)	(0.083)	(0.063)	(0.150)	(0.069)	(0.078)	(0.117)	(0.073)
Equiv. hh income (cube root)	-0.004	-0.013***	-0.004	-0.016	0.003	0.004	0.004	-0.001
	(0.004)	(0.003)	(0.005)	(0.008)	(0.006)	(0.004)	(0.005)	(0.003)
Risk aversion (ref.: no risk)								
Substantial	-0.687***	-0.270	-0.195	-0.182	-0.625	0.293	-0.537*	-0.021
	(0.205)	(0.231)	(0.226)	(0.729)	(0.414)	(0.232)	(0.234)	(0.298)
Above average	0.658*	-0.492*	0.004	0.266	-0.054	-0.562***	-0.478	0.192
	(0.314)	(0.235)	(0.258)	(0.271)	(0.114)	(0.107)	(0.252)	(0.276)
Average	-0.116	-0.136	-0.099	0.281	0.097	-0.105	-0.219	-0.035
	(0.164)	(0.109)	(0.066)	(0.169)	(0.153)	(0.107)	(0.121)	(0.108)
Constant	1.160	1.617	-1.093	7.111*	-0.910	2.320	5.331*	-2.163
	(2.349)	(1.842)	(1.453)	(2.824)	(1.864)	(1.849)	(2.267)	(1.517)
Control variables wave	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	505	674	1,013	250	1,002	627	591	1,082
Pseudo R2	0.025	0.037	0.016	0.046	0.012	0.026	0.010	0.021
AIC	3,381	4,688	6,688	1,576	6,225	3,858	4,218	7,452
BIC	3,452	4,765	6,772	1,636	6,304	3,933	4,293	7,537
SE	cluster	cluster	cluster	cluster	cluster	cluster	cluster	cluster

Note: The coefficients are based on the sample that is objectively impaired, i.e. unable to stand up from the chair. The dependent variable "doctor visits" is based on the annual number of doctor visits, visits to emergency rooms and outpatient clinic visits at wave $w+1$, i.e. Wave 4 or Wave 6. All explanatory variables are taken from wave w, i.e. Wave 2 or Wave 5 respectively. The estimated coefficients are based on a negative binomial regression model with mean dispersion. Standard errors are clustered at the household level and presented in parentheses. * p<0.05, ** p<0.01, *** p<0.001

Table 9, continued: Overestimating health and annual number of doctor visits at $w + 1$ by country

	(1) Luxembourg	(2) Netherlands	(3) Poland	(4) Slovenia	(5) Spain	(6) Sweden	(7) Switzerland
Health perception (ref.: concordance)							
Overestimating	-0.143	-0.348	-0.108	-0.253*	-0.265***	-0.106	0.116
	(0.176)	(0.236)	(0.112)	(0.107)	(0.062)	(0.163)	(0.145)
Chronic diseases	0.062	0.342***	0.142***	0.058	0.105***	0.119**	0.143*
	(0.050)	(0.104)	(0.035)	(0.033)	(0.022)	(0.041)	(0.064)
Activity limitations	0.153***	-0.006	0.031	-0.021	0.006	0.057	0.092
	(0.035)	(0.047)	(0.029)	(0.017)	(0.010)	(0.029)	(0.047)
Age	0.192	0.510	0.017	-0.010	0.008	-0.179	0.137
	(0.122)	(0.336)	(0.137)	(0.076)	(0.036)	(0.117)	(0.104)
Age squared	-0.001	-0.004	-0.000	-0.000	-0.000	0.001	-0.001
	(0.001)	(0.003)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
Woman	-0.074	-0.228	-0.113	0.247**	-0.220**	0.020	-0.004
	(0.189)	(0.248)	(0.114)	(0.093)	(0.074)	(0.150)	(0.153)
Educ. group (ref.: medium)							
Low	-0.328	-0.722*	-0.006	0.139	0.187	-0.302	-0.071
	(0.176)	(0.312)	(0.110)	(0.103)	(0.144)	(0.212)	(0.188)
High	-0.705	-0.854*	0.190	-0.155	-0.083	-0.513*	-0.161
	(0.384)	(0.366)	(0.225)	(0.184)	(0.219)	(0.204)	(0.176)
Retired	-0.627**	-0.628	0.167	0.160	0.014	0.470*	0.125
	(0.219)	(0.355)	(0.125)	(0.117)	(0.068)	(0.238)	(0.214)
Married	-0.162	0.006	-0.069	-0.108	0.031	-0.007	-0.222
	(0.173)	(0.262)	(0.117)	(0.111)	(0.075)	(0.169)	(0.168)
Equiv. hh income (cube root)	-0.011**	0.028*	0.002	0.002	-0.000	0.005	-0.004
	(0.004)	(0.011)	(0.012)	(0.008)	(0.004)	(0.010)	(0.005)
Above average	-0.955	-1.791***		0.390**	-1.569***	-0.093	-0.354
	(0.505)	(0.390)		(0.130)	(0.260)	(0.188)	(0.347)
Average	-0.391	-0.198	-0.142	0.152	-0.095	0.283	-0.380*
	(0.225)	(0.284)	(0.164)	(0.171)	(0.150)	(0.191)	(0.155)
Risk aversion (ref.: no risk)							
Substantial		-1.654***	0.391**	0.431**	0.557	0.182	-0.048
		(0.395)	(0.151)	(0.137)	(0.399)	(0.386)	(0.398)
Constant	-2.761	-14.607	1.688	2.710	2.035	7.468	-2.932
	(4.116)	(10.495)	(4.286)	(2.530)	(1.303)	(4.503)	(3.618)
Control variables wave	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	134	97	319	405	1,388	419	274
Pseudo R2	0.060	0.044	0.015	0.018	0.016	0.027	0.021
AIC	981	655	2,025	2,497	8,535	2,452	1,757
BIC	1,024	694	2,081	2,553	8,624	2,521	1,818
SE	cluster	cluster	cluster	cluster	cluster	cluster	cluster

Note: The coefficients are based on the sample that is objectively impaired, i.e. unable to stand up from the chair. The dependent variable "doctor visits" is based on the annual number of doctor visits, visits to emergency rooms and outpatient clinic visits at wave $w + 1$, i.e. Wave 4 or Wave 6. All explanatory variables are taken from wave w, i.e. Wave 2 or Wave 5 respectively. The estimated coefficients are based on a negative binomial regression model with mean dispersion. Standard errors are clustered at the household level and presented in parentheses. * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table 10: Annual number of doctor visits at $w + 1$ by chronic diseases at $w + 1$

	(1) Unimpaired No chronic dis. at w+1	(2) Unimpaired Chronic dis. at w+1	(3) Impaired No chronic dis. at w+1	(4) Impaired Chronic dis. at w+1
Health perception (ref.: concordance)				
Underestimating	0.459***	0.254***		
	(0.043)	(0.019)		
Overestimating			-0.387***	-0.173***
			(0.061)	(0.028)
Chronic diseases	0.190***	0.102***	0.104***	0.099***
	(0.015)	(0.006)	(0.028)	(0.009)
Age	-0.035	-0.036**	-0.004	-0.007
	(0.018)	(0.012)	(0.037)	(0.018)
Age squared	0.000*	0.000***	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Woman	0.154***	0.008	0.113	-0.018
	(0.025)	(0.014)	(0.060)	(0.029)
Educ. group (ref.: medium)				
Low	0.000	-0.016	0.010	0.024
	(0.030)	(0.018)	(0.070)	(0.034)
High	0.037	-0.017	-0.009	-0.075
	(0.030)	(0.019)	(0.085)	(0.047)
Retired	0.037	0.017	0.078	-0.013
	(0.035)	(0.019)	(0.073)	(0.034)
Married	-0.032	-0.027	0.042	0.012
	(0.027)	(0.016)	(0.067)	(0.030)
Equiv. hh income (cube root)	0.002	-0.002*	0.004	-0.001
	(0.001)	(0.001)	(0.003)	(0.002)
Risk aversion (ref.: no risk)				
Substantial	-0.045	0.151	-0.198	-0.046
	(0.085)	(0.080)	(0.232)	(0.147)
Above average	-0.106	-0.130***	0.017	0.001
	(0.057)	(0.039)	(0.177)	(0.136)
Average	0.046	-0.022	0.024	-0.065
	(0.027)	(0.018)	(0.085)	(0.043)
Wave 5	-0.052	-0.064***	-0.107	0.008
	(0.028)	(0.017)	(0.073)	(0.040)
Constant	1.973**	3.076***	1.105	2.810***
	(0.613)	(0.409)	(1.296)	(0.647)
Control variables country	Yes	Yes	Yes	Yes
N	17,684	29,683	1,956	6,819
Pseudo R2	0.018	0.015	0.020	0.014
AIC	86,500	179,920	10,835	46,034
BIC	86,733	180,169	11,002	46,239
SE	cluster	cluster	cluster	cluster

Note: "Unimpaired" refers to the sample that is objectively unimpaired, i.e. able to stand up from the chair and "Impaired" refers to the sample that is objectively impaired, i.e. unable to stand up from the chair. The sample "No chronic dis. at $w + 1$" includes those that have zero chronic diseases at wave $w + 1$, whereas "Chronic dis. at $w + 1$" refers to those that have one ore more chronic diseases at wave $w + 1$. The dependent variable "doctor visits" is based on the annual number of doctor visits, visits to emergency rooms and outpatient clinic visits at wave $w + 1$, i.e. Wave 4 or Wave 6. All explanatory variables are taken from wave w, i.e. Wave 2 or Wave 5 respectively. The estimated coefficients are based on a negative binomial regression model with mean dispersion. Standard errors are clustered at the household level and presented in parentheses. * p<0.05, ** p<0.01, *** p<0.001

Table 11: Robustness analyses for annual doctor visits of the unimpaired sample

	(1) Main	(2) Income 1	(3) Income 2	(4) 'Unsafe' dropped
Health perception (ref.: concordance)				
Underestimating	0.250***	0.250***	0.250***	
	(0.018)	(0.018)	(0.018)	
Chronic diseases	0.180***	0.181***	0.181***	0.180***
	(0.005)	(0.005)	(0.005)	(0.005)
Activity limitations	0.090***	0.090***	0.090***	0.090***
	(0.008)	(0.009)	(0.008)	(0.008)
Age	-0.002	-0.002	-0.002	-0.002
	(0.010)	(0.010)	(0.010)	(0.010)
Age squared	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Woman	0.036**	0.037**	0.037**	0.036**
	(0.012)	(0.012)	(0.012)	(0.012)
Educ. group (ref.: medium)				
Low	-0.007	-0.009	-0.007	-0.007
	(0.016)	(0.016)	(0.016)	(0.016)
High	-0.008	-0.007	-0.009	-0.008
	(0.016)	(0.016)	(0.016)	(0.016)
Retired	0.031	0.032	0.031	0.031
	(0.017)	(0.017)	(0.017)	(0.017)
Married	-0.025	-0.023	-0.026	-0.025
	(0.015)	(0.015)	(0.015)	(0.015)
Equiv. hh income (cube root)	-0.001			-0.001
	(0.001)			(0.001)
Risk aversion (ref.: no risk)				
Substantial	0.067	0.068	0.067	0.067
	(0.063)	(0.063)	(0.063)	(0.063)
Above average	-0.134***	-0.132***	-0.134***	-0.134***
	(0.033)	(0.033)	(0.033)	(0.033)
Average	-0.009	-0.008	-0.009	-0.009
	(0.015)	(0.015)	(0.015)	(0.015)
Wave 5	-0.091***	-0.088***	-0.091***	-0.091***
	(0.015)	(0.015)	(0.015)	(0.015)
Equiv. hh income (cube root) 2		-0.001		
		(0.001)		
Equiv. hh income not normalised			-0.000	
			(0.000)	
Health perception (ref.: concordance)				
Underestimating				0.250***
				(0.018)
Constant	1.525***	1.518***	1.505***	1.525***
	(0.349)	(0.349)	(0.348)	(0.349)
Control variables country	Yes	Yes	Yes	Yes
N	47,377	47,377	47,377	47,377
Pseudo R2	0.024	0.024	0.024	0.024
AIC	269,248	269,247	269,249	269,248
BIC	269,520	269,519	269,521	269,520
SE	cluster	cluster	cluster	cluster

Note: The coefficients are based on the sample that is objectively unimpaired, i.e. able to stand up from the chair. The dependent variable "doctor visits" is based on the annual number of doctor visits, visits to emergency rooms and outpatient clinic visits at wave $w+1$, i.e. Wave 4 or Wave 6. All explanatory variables are taken from wave w, i.e. Wave 2 or Wave 5 respectively. The estimated coefficients are based on a negative binomial regression model with mean dispersion. Standard errors are clustered at the household level and presented in parentheses. * p<0.05, ** p<0.01, *** p<0.001

Table 12: Robustness analyses for annual doctor visits of the impaired sample

	(1) Main	(2) Income 1	(3) Income 2	(4) 'Unsafe' dropped
Health perception (ref.: concordance)				
Overestimating	-0.156***	-0.156***	-0.156***	
	(0.027)	(0.027)	(0.027)	
Chronic diseases	0.133***	0.133***	0.133***	0.186***
	(0.009)	(0.009)	(0.009)	(0.027)
Activity limitations	0.032***	0.032***	0.032***	0.051*
	(0.005)	(0.005)	(0.005)	(0.021)
Age	0.023	0.024	0.023	-0.017
	(0.017)	(0.017)	(0.017)	(0.050)
Age squared	-0.000	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Woman	0.003	0.003	0.003	-0.028
	(0.027)	(0.027)	(0.027)	(0.076)
Educ. group (ref.: medium)				
Low	0.024	0.024	0.025	0.057
	(0.032)	(0.032)	(0.032)	(0.085)
High	-0.078	-0.078	-0.079	0.001
	(0.042)	(0.042)	(0.042)	(0.126)
Retired	0.015	0.016	0.015	0.079
	(0.031)	(0.031)	(0.031)	(0.098)
Married	0.014	0.016	0.013	0.097
	(0.028)	(0.029)	(0.028)	(0.083)
Equiv. hh income (cube root)	-0.001			-0.001
	(0.001)			(0.004)
Risk aversion (ref.: no risk)				
Substantial	-0.064	-0.063	-0.065	0.316
	(0.130)	(0.131)	(0.130)	(0.267)
Above average	-0.106	-0.103	-0.106	-0.919***
	(0.108)	(0.108)	(0.108)	(0.279)
Average	-0.061	-0.061	-0.062	-0.014
	(0.040)	(0.039)	(0.040)	(0.098)
Wave 5	-0.042	-0.040	-0.042	-0.215*
	(0.037)	(0.036)	(0.037)	(0.090)
Equiv. hh income (cube root) 2		-0.001		
		(0.002)		
Equiv. hh income not normalised			-0.000	
			(0.000)	
Health perception (ref.: concordance)				
Overestimating				-0.104
				(0.077)
Constant	1.331*	1.315*	1.312*	2.365
	(0.592)	(0.592)	(0.589)	(1.759)
Control variables country	Yes	Yes	Yes	Yes
N	8,780	8,780	8,780	958
Pseudo R2	0.019	0.019	0.019	0.030
AIC	57,293	57,293	57,293	5,947
BIC	57,512	57,513	57,512	6,097
SE	cluster	cluster	cluster	cluster

Note: The coefficients are based on the sample that is objectively impaired, i.e. unable to stand up from the chair. The dependent variable "doctor visits" is based on the annual number of doctor visits, visits to emergency rooms and outpatient clinic visits at wave $w+1$, i.e. Wave 4 or Wave 6. All explanatory variables are taken from wave w, i.e. Wave 2 or Wave 5 respectively. The estimated coefficients are based on a negative binomial regression model with mean dispersion. Standard errors are clustered at the household level and presented in parentheses. * p<0.05, ** p<0.01, *** p<0.001

Table 13: Robustness analysis: Annual number of doctor visits at $w+1$ by survey wave

	(1) Unimpaired Wave 2	(2) Impaired Wave 2	(3) Unimpaired Wave 5	(4) Impaired Wave 5
Health perception (ref.: concordance)				
Underestimating	0.195***		0.266***	
	(0.036)		(0.021)	
Overestimating		-0.146*		-0.155***
		(0.062)		(0.030)
Chronic diseases	0.199***	0.156***	0.175***	0.126***
	(0.011)	(0.020)	(0.006)	(0.010)
Activity limitations	0.094***	0.045**	0.090***	0.030***
	(0.018)	(0.014)	(0.009)	(0.005)
Age	-0.069*	0.046	0.001	0.025
	(0.035)	(0.079)	(0.012)	(0.018)
Age squared	0.001*	-0.000	0.000	-0.000
	(0.000)	(0.001)	(0.000)	(0.000)
Woman	0.047*	0.124*	0.033*	-0.032
	(0.022)	(0.053)	(0.014)	(0.029)
Educ. group (ref.: medium)				
Low	0.046	0.018	-0.024	0.019
	(0.028)	(0.068)	(0.018)	(0.034)
High	0.015	-0.074	-0.013	-0.083
	(0.031)	(0.089)	(0.018)	(0.046)
Retired	0.033	0.043	0.027	-0.001
	(0.029)	(0.058)	(0.020)	(0.035)
Married	-0.010	-0.071	-0.026	0.030
	(0.028)	(0.066)	(0.016)	(0.030)
Equiv. hh income (cube root)	-0.002	0.000	-0.001	-0.002
	(0.001)	(0.003)	(0.001)	(0.002)
Risk aversion (ref.: no risk)				
Substantial	-0.252**	-0.237	0.161*	0.024
	(0.096)	(0.261)	(0.074)	(0.139)
Above average	-0.107*	-0.165	-0.078	0.074
	(0.049)	(0.188)	(0.042)	(0.136)
Average	0.009	-0.205**	-0.029	-0.013
	(0.032)	(0.066)	(0.017)	(0.046)
Constant	3.339**	0.597	1.378***	1.278*
	(1.102)	(2.520)	(0.389)	(0.644)
Control variables country	Yes	Yes	Yes	Yes
N	12,318	2,101	35,059	6,679
Pseudo R2	0.028	0.024	0.024	0.020
AIC	70,099	13,610	198,857	43,607
BIC	70,299	13,763	199,094	43,798
SE	cluster	cluster	cluster	cluster

Note: "Unimpaired" refers to the sample that is objectively unimpaired, i.e. able to stand up from the chair and "Impaired" refers to the sample that is objectively impaired, i.e. unable to stand up from the chair. The dependent variable "doctor visits" is based on the annual number of doctor visits, visits to emergency rooms and outpatient clinic visits at wave $w+1$, i.e. Wave 4 or Wave 6. All explanatory variables are taken from wave w, i.e. Wave 2 or Wave 5 respectively. The estimated coefficients are based on a negative binomial regression model with mean dispersion. Standard errors are clustered at the household level and presented in parentheses. * p<0.05, ** p<0.01, *** p<0.001

Table 14: Health perception based on cognition

	(1) Unimpaired 3 words	(2) Impaired 3 words	(3) Unimpaired 2 words	(4) Impaired 2 words
Health perception (ref.: concordance)				
Underestimating	0.076***			
	(0.014)			
Overestimating		-0.078**		
		(0.028)		
Underestimating			0.077***	
			(0.013)	
Overestimating				-0.139**
				(0.042)
Chronic diseases	0.194***	0.133***	0.188***	0.137***
	(0.005)	(0.010)	(0.004)	(0.013)
Activity limitations	0.102***	0.043***	0.093***	0.036***
	(0.005)	(0.005)	(0.005)	(0.006)
Age	-0.004	0.042*	0.001	0.051*
	(0.009)	(0.019)	(0.009)	(0.026)
Age squared	0.000	-0.000*	0.000	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)
Woman	0.069***	0.003	0.064***	-0.046
	(0.012)	(0.029)	(0.012)	(0.044)
Educ. group (ref.: medium)				
Low	0.003	-0.065	-0.002	-0.001
	(0.015)	(0.038)	(0.014)	(0.055)
High	0.018	-0.077	0.007	-0.031
	(0.016)	(0.061)	(0.015)	(0.094)
Retired	0.046**	-0.010	0.036*	0.028
	(0.017)	(0.036)	(0.016)	(0.050)
Married	-0.007	-0.030	-0.011	-0.032
	(0.014)	(0.032)	(0.013)	(0.045)
Equiv. hh income (cube root)	-0.001	-0.001	-0.001	-0.001
	(0.001)	(0.002)	(0.001)	(0.002)
Risk aversion (ref.: no risk)				
Substantial	0.144*	-0.228	0.118*	-0.229
	(0.058)	(0.141)	(0.056)	(0.191)
Above average	-0.093**	-0.129	-0.090*	-0.239
	(0.034)	(0.139)	(0.035)	(0.129)
Average	-0.054***	-0.064	-0.053***	-0.225**
	(0.014)	(0.051)	(0.014)	(0.069)
Wave 5	-0.010	-0.051	-0.012	-0.092*
	(0.012)	(0.031)	(0.012)	(0.046)
Constant	1.549***	0.454	1.388***	0.430
	(0.315)	(0.685)	(0.299)	(0.908)
Control variables country	Yes	Yes	Yes	Yes
N	54,948	8,550	59,790	3,708
Pseudo R2	0.024	0.018	0.023	0.023
AIC	317,264	53,027	347,485	22,924
BIC	317,532	53,238	347,755	23,111
SE	cluster	cluster	cluster	cluster

Note: In columns 1 and 2, individuals are considered objectively impaired if they recall 3 words or less ("3 words"), while in columns 3 and 4 the cutoff is at 2 words or less ("2 words"). The dependent variable "doctor visits" is based on the annual number of doctor visits, visits to emergency rooms and outpatient clinic visits at wave $w + 1$, i.e. Wave 5 or Wave 6. All explanatory variables are taken from wave w, i.e. Wave 4 or Wave 5 respectively. The estimated coefficients are based on a negative binomial regression model with mean dispersion. Standard errors are clustered at the household level and presented in parentheses. * p<0.05, ** p<0.01, *** p<0.001

Table 15: Health perception based on walking ability

	(1) Unimpaired	(2) Impaired
Health perception (ref.: concordance)		
Underestimating	0.277**	
	(0.085)	
Overestimating		-0.245*
		(0.123)
Chronic diseases	0.118***	-0.037
	(0.020)	(0.048)
Age	0.186***	0.104
	(0.051)	(0.095)
Age squared	-0.001***	-0.001
	(0.000)	(0.001)
Woman	-0.208***	-0.115
	(0.059)	(0.135)
Educ. group (ref.: medium)		
Low	-0.236**	0.137
	(0.075)	(0.177)
High	-0.205*	-0.067
	(0.097)	(0.247)
Retired	-0.116	-0.235
	(0.089)	(0.146)
Married	-0.055	0.073
	(0.061)	(0.112)
Equiv. hh income (cube root) 2	-0.003	-0.005
	(0.004)	(0.011)
Wave 2	-0.258**	0.104
	(0.096)	(0.211)
Constant	-4.310*	-0.439
	(1.897)	(3.470)
Control variables country	Yes	Yes
N	1,771	295
Pseudo R2	0.030	0.040
AIC	10,925	1,965
BIC	11,062	2,057
SE	cluster	cluster

Note: The dependent variable "doctor visits" is based on the annual number of doctor visits, visits to emergency rooms and outpatient clinic visits at wave $w+1$, i.e. Wave 2 or Wave 4. All explanatory variables are taken from wave w, i.e. Wave 1 or Wave 2 respectively. The estimated coefficients are based on a negative binomial regression model with mean dispersion. Standard errors are clustered at the household level and presented in parentheses. * p<0.05, ** p<0.01, *** p<0.001