



Article

# Does Self-Assessed Health Reflect the True Health State?

Pavitra Paul <sup>1,2,\*</sup> , Ulrich Nguemdjo <sup>3,4</sup>, Natalia Kovtun <sup>5</sup> and Bruno Ventelou <sup>3</sup>

<sup>1</sup> Centre de Sciences Humaines (CSH, UMIFRE n°20), a Network of Research Units of the National Centre for Scientific Research (CNRS) and the French Ministry of European and Foreign Affairs, New Delhi 110011, India

<sup>2</sup> Department of Public Health, Yerevan State Medical University, Yerevan 0025, Armenia

<sup>3</sup> CNRS, AMSE, Aix-Marseille Université, 13001 Marseille, France;

ulrich-boris.nguemdjo-kamguem@univ-amu.fr (U.N.); bruno.ventelou@univ-amu.fr (B.V.)

<sup>4</sup> Laboratoire Population Environnement Développement, Aix-Marseille Université, 13001 Marseille, France

<sup>5</sup> Department of Statistics and Demography, Taras Shevchenko National University of Kyiv, 03022 Kiev, Ukraine; kovtun\_natali@ukr.net

\* Correspondence: pavitra.paul@mail.fr

**Abstract:** Self-assessed health (SAH) is a widely used tool to estimate population health. However, the debate continues as to what exactly this ubiquitous measure of social science research means for policy conclusions. This study is aimed at understanding the tenability of the construct of SAH by simultaneously modelling SAH and clinical morbidity. Using data from 17 waves (2001–2017) of the Russian Longitudinal Monitoring Survey, which captures repeated response for SAH and frequently updates information on clinical morbidity, we operationalise a recursive semi-ordered probit model. Our approach allows for the estimation of the distributional effect of clinical morbidity on perceived health. This study establishes the superiority of inferences from the recursive model. We illustrated the model use for examining the endogeneity problem of perceived health for SAH, contributing to population health research and public policy development, in particular, towards the organisation of health systems.

**Keywords:** clinical morbidity; endogeneity; perceived health; recursive; semi-ordered; Russia



**Citation:** Paul, P.; Nguemdjo, U.; Kovtun, N.; Ventelou, B. Does Self-Assessed Health Reflect the True Health State? *Int. J. Environ. Res. Public Health* **2021**, *18*, 11153. <https://doi.org/10.3390/ijerph182111153>

Academic Editor: Ceri J. Phillips

Received: 24 July 2021

Accepted: 19 October 2021

Published: 23 October 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The self-assessed health (SAH) measure results from an “ordered” response of the individual on a qualitative (subjective) perception of her/his own health status (perceived health), although the subjective nature of SAH remains questionable in the literature [1–5]. If all individuals have a similar perception of health status, SAH would reflect ‘true’ health and could therefore serve as a valid indicator. However, problems occur when individuals vary in their reporting behaviour and, as a consequence, SAH deviates from the underlying ‘true’ health. For instance, the same clinical health condition is acknowledged differently according to individual characteristics and is determined by the cultural and historical context, social position, and health experiences of the individual [1,6]. A potential explanation for such varying assessments lies in different people’s varying ability to adapt to ill health [2,7,8]. People who have been ill for a long period of time may report levels of health that are much better than those of individuals suffering from the same illness for the first time, which then leads to the differential reporting of SAH among those with an objective health status. We attempt to examine this further.

An abundance of literature seeks to explain the variations in SAH [9,10]. Confronted by a given measurement of ‘true’ health, SAH seems to be associated with sociodemographic, socioeconomic, behavioural, psychosocial, and chronic health conditions [11–16]. Tay et al. [17] established a relationship between SAH and sociodemographic factors, wherein gender and material circumstances consistently and independently predict SAH. Hernández-Quevedo et al. [18] found that reporting heterogeneity is associated with socioeconomic strata; this differs from the concept of social class, which refers to social groups arising from interdependent economic relationships among people [19]. Au and

Johnston [9] found associations between health dimensions (including history of illness) and SAH for an average level of perceived health. Lindeboom and van Doorslaer [20] found evidence of age and gender-related reporting bias in SAH. Krause and Jay [21] found that people of different age groups tend to think of health differently when making evaluations: older respondents were more likely to use specific health problems (hypertension) as a reference point for their health, while younger people focused on their own physical functioning such as mobility issues and acute conditions. Lastly, and probably the most troubling result, some chronic medical morbidity was found to be positively associated with SAH for late middle-aged and elderly non-institutionalised people living in the northern part of the Netherlands [22]. A qualitative study [23] concluded that subgroup differences in SAH could be attributed to experiences with ill health.

Systematic differential health status reporting by subgroups of the population presents a serious problem to the validity of subgroup comparisons of health levels and thus to measure degrees of inequality. In particular, the possible sensitivity of SAH to the information previously delivered to the respondents by the health system is an acknowledged concern for the use of SAH metrics in health inequality measurements; this implies that the response of the individual is not independent from the individuals' (or groups') healthcare services utilisation [24], although such independency is precisely expected when SAH is used for evidence-informed policy development. Experience of interaction with the health system is not independent of socioeconomic strata and thus the thresholds of reported health status is often influenced by the frequency and quality of encounters for each specific individual with the respective health system [25,26].

Distinguishing the underlying cause of differences in reporting behaviour becomes a necessity to unveil the bias in the measurement of socioeconomic inequality in health. Although several studies have investigated factors that are likely associated with SAH and socioeconomic biases, we focus on finding the answer to the question of effect regarding clinical morbidity (diagnosed chronic disease) on different levels of the perceived health status. We also aim at examining the effect of duration of suffering from chronic disease on the perceived health status. However, considering that having clinical morbidity as a predictor of perceived health may raise an endogeneity problem when using a naive regression model, we introduce the application of the recursive semi-ordered probit model to control for endogeneity, as well as to examine the "causal" effect of the clinical morbidity on the perceived health in the direction of such relationship. The strength of this study concerns understanding the tenability of the construct of SAH by simultaneously modelling SAH and clinical morbidity, while operationalising a simultaneous binary and ordered probit-semi ordered approach [27] in healthcare and health policy research. Furthermore, our approach also contributes to the estimation of the distributional effect of clinical morbidity on different levels of perceived health. We describe the data and analysis strategy in Section 2. Section 3 presents the descriptive statistics and results from the application of the econometric models. We discuss the empirical evidence that emerged in Section 4 and conclude with the contributions of this study in Section 5.

## 2. Materials and Methods

Our data derives from 17 waves (2001–2017) of the Russian Longitudinal Monitoring Survey (RLMS) that captures repeated response on SAH and updates information on diagnosed chronic illness (considered as clinical morbidity in our study). SAH is the subjective response elicited from the structured question "How would you evaluate your health?" in an ordered scale (1 = very good (excellent), 2 = good, 3 = average, 4 = bad, and 5 = very bad), while diagnosed chronic illness indicates the diagnosis established by a physician. The multivariate distribution of the sample by gender, age, and urban/rural location compared quite well with the corresponding multivariate distribution of the Soviet/Russian census: there is usually a difference of only one percentage point or less between the two distributions—<http://www.cpc.unc.edu/rlms> (accessed on 21 February 2021). The RLMS (<http://www.cpc.unc.edu/rlms>, accessed on 21 February 2021) is an

ongoing longitudinal household survey of the Russian Federation and the survey captures both the perceived health status and specific (objective) health status at the individual level (Table 1). The RLMS applies a multi-stage sampling method with precomputed cross-sectional post-stratification weights. These weights adjust not only for design factors but also for deviations from the census characteristics. Due to the decline in response rate in large cities, sample repair was done in Round 15 (2006; <http://www.cpc.unc.edu/projects/rlms-hse/data>, accessed on 21 February 2021).

**Table 1.** Clinical morbidity and levels of perceived health status.

Clinical Morbidity	Duration (Mean) of Suffering in Years	Perceived Health (%)				
		Very Bad	Bad	Average	Good	Excellent
Heart disease ( <i>N</i> = 6795)	14.96 (0.17)	7.49	37.44	51.48	3.49	0.10
Gastrointestinal disease ( <i>N</i> = 6822)	15.72 (0.16)	4.63	25.49	61.24	8.46	0.18
Spinal diseases ( <i>N</i> = 6495)	15.69 (0.16)	4.46	26.82	60.63	7.90	0.18
Other chronic disease ( <i>N</i> = 9127)	14.42 (0.13)	5.23	27.93	57.97	8.63	0.24
Assigned disability ( <i>N</i> = 6761)	9.89 (0.12)	9.50	41.15	44.56	4.61	0.18

The test of independence ( $\chi^2$ ) for all values (presence of clinical morbidity in the distribution of perceived health status) is at the *p*-value = 0.000. Figures in parentheses indicate standard error.

Our data is a mix of cross-sectional and panel data (in which a segment of the households were followed over time, although only 313 individuals were present in all 17 waves), although the panel is not a balanced one. After omitting respondents with missing information on the outcome variables or covariates, the estimation sample included 145,239 observations of respondents aged 18 years and above from 79,795 households (Table 2).

**Table 2.** Data used.

Survey Year	Respondents (Aged 18 Years and above)	Year-on-Year Attrition (%)
2001	4773	16.47
2002	4874	10.65
2003	4768	10.63
2004	9627	9.21
2005	9393	5.75
2006	11,460	24.15
2007	11,294	9.00
2008	6011	8.27
2009	5916	6.37
2010	8993	37.95
2011	10,074	17.22
2012	10,942	15.40
2013	10,468	10.79
2014	9097	6.41
2015	9091	6.84
2016	9136	5.81
2017	9322	5.13
Total	145,239	12.44

Our analysis included adult individuals with clinical morbidity, i.e., heart disease, gastrointestinal disease, spinal disease, and other chronic disease. Furthermore, we also included individuals with assigned disabilities. A large number of missing information on the year of diagnosis for other diseases did not allow us to include other morbid conditions captured in the survey. Therefore, our variable of clinical morbidity was any one (and/or in combination) of the four specified diseases and the duration (the maximum duration of suffering with any one (and/or in combination) of these four diseases and the maximum duration of an assigned disability; Table 1). Table 1 also indicates the presence of clinical morbidity at different levels of perceived health.

To control for response heterogeneity, we included in our analysis a vector of time-varying characteristics, i.e., age, educational attainment, marital status, and working status. We additionally controlled individual-level fixed effects, which represent time-invariant

characteristics, i.e., settlement of residence (village, town, and city). We also included the response for income and overall life satisfaction in our study.

Our variable of interest (perceived health status) is an ordered categorical variable and the simple ordered probit model (naive model; details in Appendix A) could appear as a natural choice for assessing the effect of clinical morbidity on perceived health status. However, clinical morbidity, when considered as a determinant of perceived health status, raises an endogeneity issue concerning the affliction with clinical morbidity and perceived health status that is likely to be dependent and perhaps often influenced by the same unobservable confounders.

The standard instrumental variable (IV) approach is likely to produce inconsistent estimates when endogeneity stems from discrete regressors [27]. The classical ordered probit model assumes that errors in the latent regression equations for the selection mechanism and outcome variable follow a bivariate Gaussian distribution. The maximum likelihood (ML) estimator of an ordered probit model is known to be inconsistent if the unobservable factors affecting the outcome of interest are correlated with the unobservable factors affecting the selection mechanism. So, we applied a recursive semi-ordered probit model [27] under the assumption of normally distributed unobserved errors; we attributed a distribution to  $\beta$  with a mean,  $b_0 = E(\beta)$  and a dispersion matrix,  $\sigma_2 P_0 = D(\beta)$ ; this is, in effect, a Bayesian prior. The model is a recursive, simultaneous equation model. Using the sample of 145,239 adults, we simultaneously modelled perceived health and clinical morbidity. The model is recursive because  $y_{1i}$ , the observed binary realisation of the latent variable  $y_{2i}^*$ , appears on the right-hand side of Equation (2).

$$y_{1i}^* = \sum_{j=1}^7 x'_{1ij} \beta_{1j} + \varepsilon_{1i} \quad (1)$$

where  $y_{1i}$  is a vector of clinical morbidity (1 = yes; 0 = no) and  $x'_{1ij}$  are vectors for each  $j$  covariate for list of variables:

1. Age in years;
2. Gender (1 = female; 0 = male);
3. Marital status (1 = yes; 0 = otherwise);
4. Level of education (1 = incomplete secondary, 2 = secondary level completed with or without vocational training, and 3 = higher education);
5. Settlement of residence (1 = village, 2 = urban/small town, and 3 = city);
6. Working status (1 = employed; 0 = otherwise); and
7. Had an episode of acute illness during the last 12 months (1 = yes; 0 = otherwise).

$\beta_{1j}$  are vectors of unknown parameters and  $\varepsilon_{1i}$  is error term.

$$y_{2i}^* = \sum_{k=1}^7 x'_{2ik} \beta_{2k} + \gamma y_{1i} + \varepsilon_{2i} \quad (2)$$

where  $y_{2i}$  is a vector of perceived health status (SAH: 5 = excellent perceived health, . . . , 1 = very bad perceived health) and  $x'_{2ik}$  are vectors for each  $k$  covariate for a list of variables:

1. Duration in years of suffering with chronic disease;
2. Duration in years of disability;
3. Age in years;
4. Gender (1 = female; 0 = male);
5. Marital status (1 = yes; 0 = no);
6. Socioeconomic position (SEP; 1 = above median income; 0 = otherwise); and
7. Overall life satisfaction (1 = satisfied; 0 = otherwise), which is a control variable for possible systematic bias in the self-reported responses.

$\beta_{2k}$  and  $\gamma$  are vectors of unknown parameters.  $\gamma$  represents the effect that  $y_{1i}^*$  has on  $y_{2i}^*$  and  $\varepsilon_{2i}$  is the error term.

$$\text{Cov} [\varepsilon_{1i}, \varepsilon_{2i} | x_{1ij}, x_{2ik}] = \rho \quad (3)$$

Two error terms, namely  $\varepsilon_{1i}$  and  $\varepsilon_{2i}$ , are assumed to be jointly normal with correlation coefficient  $\rho$ .

Clinical morbidity is denoted by  $y_{1i}$  and takes on a unitary value if diagnosed with any or a combination of the chronic diseases. SAH is an ordered discrete variable,  $y_{2i}$ , which takes on values from 1 to 5 (5 = excellent perceived health, 4 = good, 3 = average, 2 = bad, and 1 = very bad perceived health). Let  $y_{1i}^*$  be the latent variable of clinical morbidity and  $y_{2i}^*$  be the latent variable of SAH, which depend on the exogenous variables  $x_{1i}$  and  $x_{2i}$ , respectively. Endogeneity is considered in Equation (2) which employs parameter  $\gamma$ . Testing the recursivity of the model is done by testing  $\gamma = 0$ , whereas testing the endogeneity of  $y_{1i}$  in Equation (2) requires testing the hypothesis that  $\rho = 0$  using a likelihood ratio.

The probability of taking certain discrete values depends on the cut-off points estimated. For clinical morbidity, there is only one cut-off point,  $c_{1,1}$ , whereas for SAH, there are four cut-off points:  $c_{2,1}$ ,  $c_{2,2}$ ,  $c_{2,3}$ , and  $c_{2,4}$ .

$$y_{1i} = \begin{cases} 0 & \text{if } y_{1i}^* \leq c_{1,1} \\ 1 & \text{otherwise} \end{cases} \quad (4)$$

$$y_{2i} = \begin{cases} 1 & \text{if } y_{2i}^* \leq c_{2,1} \\ 2 & \text{if } c_{2,1} < y_{2i}^* \leq c_{2,2} \\ 3 & \text{if } c_{2,2} < y_{2i}^* \leq c_{2,3} \\ 4 & \text{if } c_{2,3} < y_{2i}^* \leq c_{2,4} \\ 5 & \text{if } c_{2,4} < y_{2i}^* \end{cases} \quad (5)$$

We employed instruments to confirm the recursivity of the model and to correct for its endogeneity. Our instruments are ‘settlement of residence’ and ‘working status’. These instruments were independent of unmeasured confounders but likely induce substantial variation in the endogenous covariate, i.e., in the presence of clinical morbidity. We chose such variables that have a higher probability of an effective interaction with the health system and these variables are settlement of residence and working status but are unrelated to SAH (exclusion restriction). Such an approach identified the model parameters in a better way and thus established the robustness of the model results [28].

The model was estimated by the full information maximum likelihood function (an asymptotically efficient estimator for simultaneous models with normally distributed errors), all estimated parameters being structural.

We also fully derived the conditional probabilities and partial effects on differences in the conditional probabilities within the recursive semi-ordered probit model (Appendix B).

The mean variance inflation factor (VIF) of 2.2 suggested the stability of the coefficients and the tolerance (1/VIF) values of not lower than 0.5 for none of the predictor variables implied a substantial large contribution of each of the predictor variables used in the models.

### 3. Results

Testing the exogeneity of the clinical morbidity in the recursive semi-ordered probit model concerned testing significance of Equation (3).  $\rho$ , the ‘atanhrho’ statistic, measured Fisher’s Z transformation of the correlation between error terms from both models [29] and was strongly significant (+0.23,  $p$ -value = 0.000). Such strong evidence of the unobserved heterogeneity conjointly affecting the morbidity (diagnosed chronic disease) and perceived health established the endogeneity effect of clinical morbidity (as a predictor) on perceived health. Furthermore, a strong and significant likelihood ratio (LR) test (91.77,  $p$ -value = 0.000) established a parsimoniousness of the models used.

Table 3 (column 1) presents that the probability of having clinical morbidity was influenced by age, gender, education, marital status, settlement of residence, working status, and episode of acute illness. The effect of age was found to be significant, in tandem with the expected effects from other variables (gender, economic position, and so on).

Table 3. The model results.

	Recursive Model	Naive Model
Dependent Variable = Clinical Morbidity	Coeff.	Coeff.
Age	0.018 *** [(0.02)–(0.02)]	
Gender (comparison: female)		
Male	−0.202 *** [(−0.22)–(−0.18)]	
Education (comparison: secondary education incomplete)		
Secondary with/without vocational training	−0.072 *** [(−0.10)–(−0.05)]	
Higher education	−0.135 *** [(−0.16)–(−0.11)]	
Marital status (comparison: otherwise)		
Married	−0.035 ** [(−0.06)–(−0.01)]	
Settlement of residence (comparison: village)		
Town	0.104 *** [(0.08)–(0.13)]	
City	0.202 *** [(0.18)–(0.22)]	
Working status (comparison: otherwise)		
Employed	−0.116 *** [(−0.14)–(−0.10)]	
Acute illness episode (comparison: no heart attack/no stroke)		
Heart attack	1.654 *** [(1.57)–(1.74)]	
Stroke	0.695 *** [(0.63)–(0.76)]	
<b>Dependent Variable = Perceived Health Status</b>		
Clinical morbidity	−1.238 *** [(−1.32)–(−1.16)]	−0.817 *** [(−0.85)–(−0.79)]
Duration of clinical morbidity		0.003 *** [(0.00)–(0.00)]
Interaction effect (duration of clinical morbidity)	0.004 *** [(0.00)–(0.01)]	
Assigned with disability	−0.947 *** [(−0.99)–(−0.90)]	−1.003 *** [(−1.05)–(−0.96)]
Duration of life with disability	0.009 *** [(0.01)–(0.01)]	0.010 *** [(0.01)–(0.01)]
Age	−0.028 *** [(−0.03)–(−0.03)]	−0.030 *** [(−0.03)–(−0.03)]
Gender (comparison: female)		
Male	0.218 *** [(0.20)–(0.23)]	0.236 *** [(0.22)–(0.25)]
Marital status (comparison: otherwise)		
Married	−0.063 *** [(−0.08)–(−0.05)]	−0.052 *** [(−0.07)–(−0.03)]
Economic position (comparison: below median income)		
Income above median level	0.083 *** [(0.07)–(0.10)]	0.086 *** [(0.07)–(0.10)]
Satisfied with Life (comparison: otherwise)		
Satisfied with Life	0.424 *** [(0.41)–(0.44)]	0.428 *** [(0.41)–(0.44)]
Year (comparison: 2001)		
2002	−0.042 [(−0.09)–(0.00)]	−0.043 [(−0.09)–(0.01)]
2003	0.066 ** [(0.02)–(0.11)]	0.066 ** [(0.02)–(0.12)]
2004	0.033 [(−0.01)–(0.07)]	0.035 [(−0.01)–(0.08)]
2005	0.089 *** [(0.05)–(0.13)]	0.092 *** [(0.05)–(0.14)]
2006	0.058 ** [(0.02)–(0.10)]	0.060 ** [(0.02)–(0.10)]
2007	0.092 *** [(0.05)–(0.13)]	0.095 *** [(0.05)–(0.14)]
2008	0.130 *** [(0.09)–(0.17)]	0.132 *** [(0.09)–(0.18)]

Table 3. Cont.

Dependent Variable = Perceived Health Status	Recursive Model	Naive Model
	Coeff.	Coeff.
2009	0.138 *** [(0.09)–(0.18)]	0.141 *** [(0.09)–(0.19)]
2010	0.175 *** [(0.13)–(0.22)]	0.178 *** [(0.14)–(0.22)]
2011	0.206 *** [(0.17)–(0.25)]	0.210 *** [(0.17)–(0.25)]
2012	0.132 *** [(0.09)–(0.17)]	0.136 *** [(0.09)–(0.18)]
2013	0.119 *** [(0.08)–(0.16)]	0.123 *** [(0.08)–(0.16)]
2014	0.161 *** [(0.12)–(0.20)]	0.166 *** [(0.12)–(0.21)]
2015	0.166 *** [(0.12)–(0.21)]	0.171 *** [(0.13)–(0.21)]
2016	0.223 *** [(0.18)–(0.26)]	0.229 *** [(0.19)–(0.27)]
2017	0.222 *** [(0.18)–(0.26)]	0.229 *** [(0.19)–(0.27)]
N	145,239	145,239
AIC	–13.943	–12.674
BIC	–13.941	–12.623
Log-likelihood	–170,215.08	–120,338.33

Legends: \*\*,  $p$ -value  $\leq 0.01$ ; and \*\*\*,  $p$ -value  $\leq 0.001$ . Figures in parentheses indicate confidence interval at 95%.

The second part of Table 3 provides the result for SAH with the two instruments (settlement of residence and working status) used in the first step. We examined the endogeneity effect of clinical morbidity (as a predictor) on perceived health. Our variable of interest, i.e., clinical morbidity, was found to have a significant influence on the perceived health status. The recursive model displayed a much stronger negative and significant impact on perceived health (coefficient value,  $-1.24$ ) after treatment of the endogeneity effect compared to the naive model (coefficient value,  $-0.82$ ) implying an estimated coefficient biased with endogeneity. Most strikingly, a positive sign of our interaction variable, as captured in the recursive model (could not be captured in the naive model because being diagnosed with any one of the chronic diseases was more significant than the duration of suffering, thus the number of the levels of the interaction effect was too big compared to the diagnosed levels with any one of the chronic diseases (i.e., 0 or 1)), indicated that the duration of suffering also had a significant influence on the perceived health status.

A strong year effect on perceived health status, *ceteris paribus*, was evident consistently from 2004 and onwards compared to 2001, suggesting an overall influence of country environment (possibly the living conditions) on perceived health status. The year effect on SAH exhibited an increasing trend with the reference year as 2001. The minimised AIC and BIC scores favoured the recursive model (Table 3).

Estimating the recursive semi-ordered probit model allowed us to compute conditional probabilities and the differences in conditional probabilities at each given level of perceived health (Table 4). Although statistically strongly significant, conditional probabilities (condition of having or not having clinical morbidity) were almost similar in both situations for very bad perceived health status, while the differences between the conditional probabilities were substantial (even more than 20%) for the higher levels of perceived health status. Furthermore, the impact of clinical morbidity on the higher levels of perceived health status were continuously increasing from 2014 compared to the preceding year. The limited number of respondents with excellent perceived health status in our data did not allow the model to capture any effect of clinical morbidity on the probabilities of excellent perceived health status across the years. However, interestingly, the condition of having clinical morbidity did not affect the probability of good perceived health, while not having clinical morbidity increased the probability of good perceived health.

**Table 4.** Conditional probabilities and difference in conditional probabilities for different levels of perceived health.

Probability	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
PrCond(clinical morbidity = 1)	0.016 (0.001)	0.017 (0.001)	0.015 (0.001)	0.016 (0.001)	0.014 (0.001)	0.015 (0.001)	0.014 (0.001)	0.014 (0.001)	0.013 (0.001)	0.013 (0.001)	0.012 (0.001)	0.013 (0.001)	0.014 (0.001)	0.013 (0.001)	0.013 (0.001)	0.012 (0.001)	0.012 (0.001)
PrCond(clinical morbidity = 0)	0.015 (0.001)	0.016 (0.001)	0.014 (0.001)	0.014 (0.001)	0.013 (0.001)	0.014 (0.001)	0.013 (0.001)	0.012 (0.001)	0.012 (0.001)	0.012 (0.001)	0.011 (0.001)	0.012 (0.001)	0.013 (0.001)	0.012 (0.001)	0.012 (0.001)	0.011 (0.001)	0.011 (0.001)
Difference at perceived health = very bad	0.001	0.001	0.001	0.002	0.001	0.001	0.001	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Prcond(clinical morbidity = 1)	0.108 (0.003)	0.112 (0.002)	0.102 (0.002)	0.105 (0.002)	0.100 (0.002)	0.103 (0.002)	0.100 (0.002)	0.097 (0.002)	0.096 (0.002)	0.093 (0.002)	0.090 (0.002)	0.096 (0.002)	0.098 (0.002)	0.094 (0.002)	0.094 (0.002)	0.089 (0.002)	0.089 (0.002)
Prcond(clinical morbidity = 0)	0.076 (0.002)	0.079 (0.002)	0.071 (0.002)	0.073 (0.001)	0.069 (0.001)	0.071 (0.001)	0.069 (0.001)	0.066 (0.001)	0.066 (0.001)	0.063 (0.001)	0.061 (0.001)	0.066 (0.001)	0.067 (0.001)	0.064 (0.001)	0.064 (0.001)	0.060 (0.001)	0.060 (0.001)
Difference at perceived health = bad	0.032	0.033	0.031	0.032	0.031	0.032	0.031	0.031	0.030	0.030	0.029	0.030	0.031	0.030	0.030	0.029	0.029
PrCond(clinical morbidity = 1)	0.875 (0.004)	0.870 (0.003)	0.883 (0.003)	0.879 (0.002)	0.885 (0.002)	0.882 (0.002)	0.886 (0.002)	0.890 (0.003)	0.891 (0.003)	0.894 (0.002)	0.898 (0.002)	0.890 (0.002)	0.889 (0.002)	0.893 (0.002)	0.894 (0.002)	0.899 (0.002)	0.899 (0.002)
PrCond(clinical morbidity = 0)	0.634 (0.006)	0.638 (0.005)	0.627 (0.005)	0.630 (0.005)	0.624 (0.005)	0.628 (0.005)	0.624 (0.005)	0.619 (0.005)	0.618 (0.005)	0.614 (0.005)	0.610 (0.005)	0.619 (0.005)	0.621 (0.005)	0.615 (0.005)	0.615 (0.005)	0.607 (0.006)	0.607 (0.006)
Difference at perceived health = average	0.241	0.232	0.256	0.249	0.261	0.254	0.262	0.271	0.273	0.28	0.288	0.271	0.268	0.278	0.279	0.292	0.292
Prcond(clinical morbidity = 1)	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	0.000
Prcond(clinical morbidity = 0)	0.275 (0.008)	0.266 (0.006)	0.289 (0.006)	0.282 (0.005)	0.294 (0.006)	0.287 (0.005)	0.294 (0.005)	0.302 (0.006)	0.304 (0.006)	0.312 (0.006)	0.318 (0.006)	0.303 (0.006)	0.300 (0.006)	0.309 (0.006)	0.310 (0.006)	0.322 (0.006)	0.322 (0.006)
Difference at perceived health = good	−0.275	−0.266	−0.289	−0.282	−0.294	−0.287	−0.294	−0.302	−0.304	−0.312	−0.318	−0.303	−0.300	−0.309	−0.310	−0.322	−0.322

Figures in parentheses indicate standard errors. All values are significant at a  $p$ -value of  $\leq 0.001$ .



Figures 1–5 present the year-on-year probability of the average effect of clinical morbidity on the different levels of perceived health status when the reference year was 2001. Overall, the effect of clinical morbidity on the different levels of perceived health status was not consistent across the years. With clinical morbidity, the probability of higher levels (above average) of perceived health decreased. In general, the probability of the average effect of clinical morbidity on the probability of “average level” of perceived health status (Figure 3) compared to such effect (s) on “very bad” (Figure 1) and “bad” (Figure 2) perceived health statuses was substantially higher from 2007 and onwards.

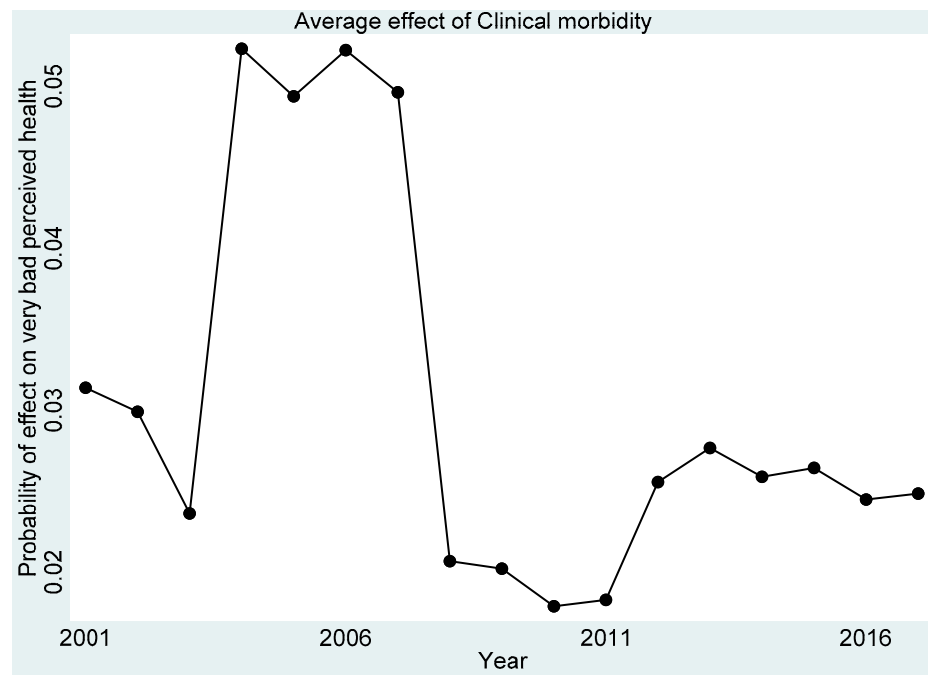


Figure 1. Probability of effects on “very bad” perceived health from the average effect of clinical morbidity.

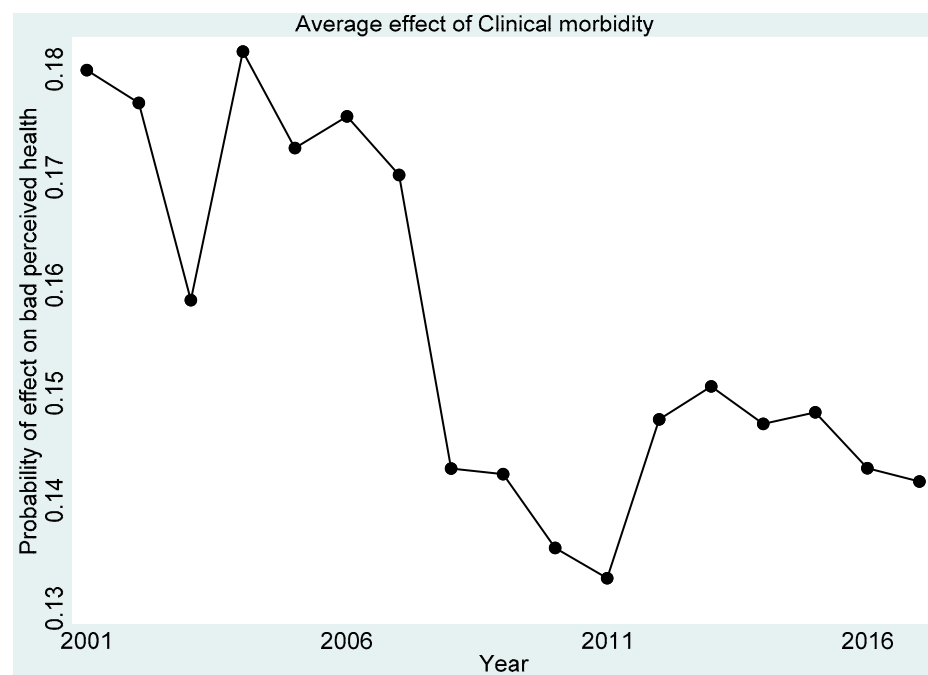


Figure 2. Probability of effects on “bad” perceived health from the average effect of clinical morbidity.

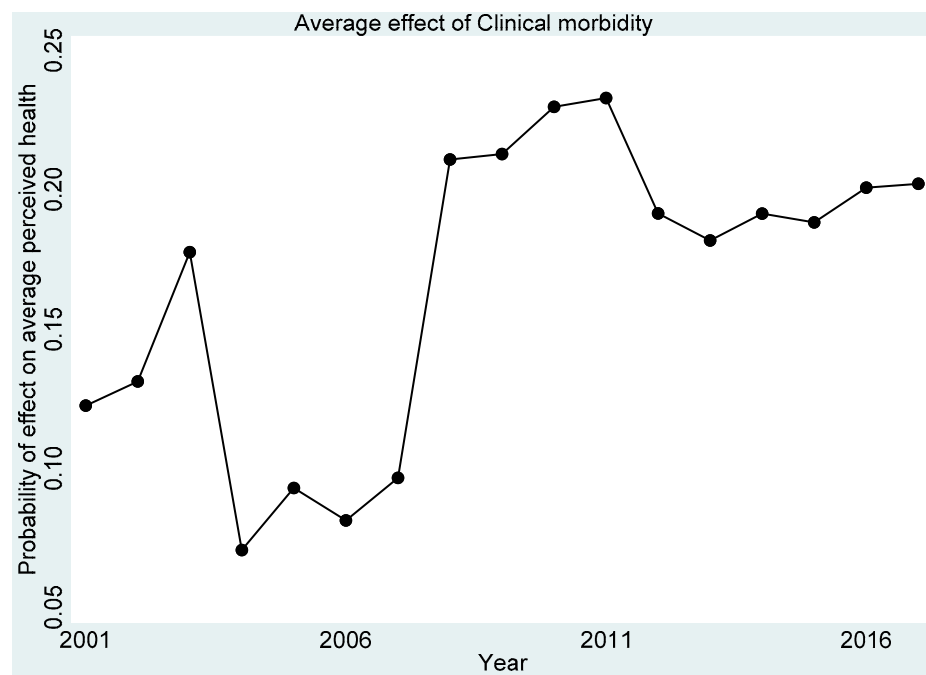


Figure 3. Probability of effects on “average” perceived health from the average effect of clinical morbidity.

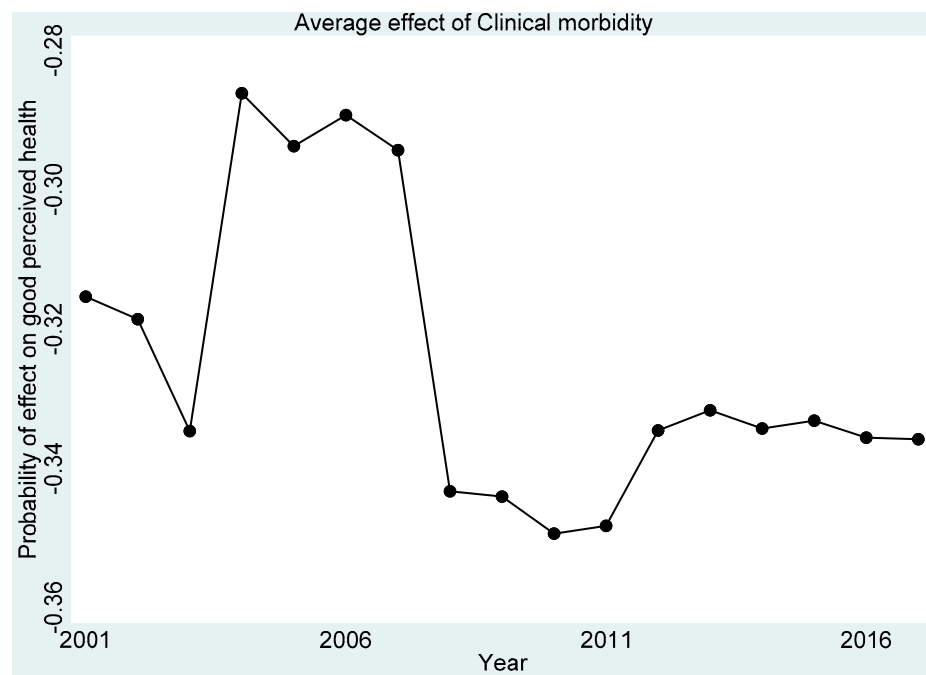
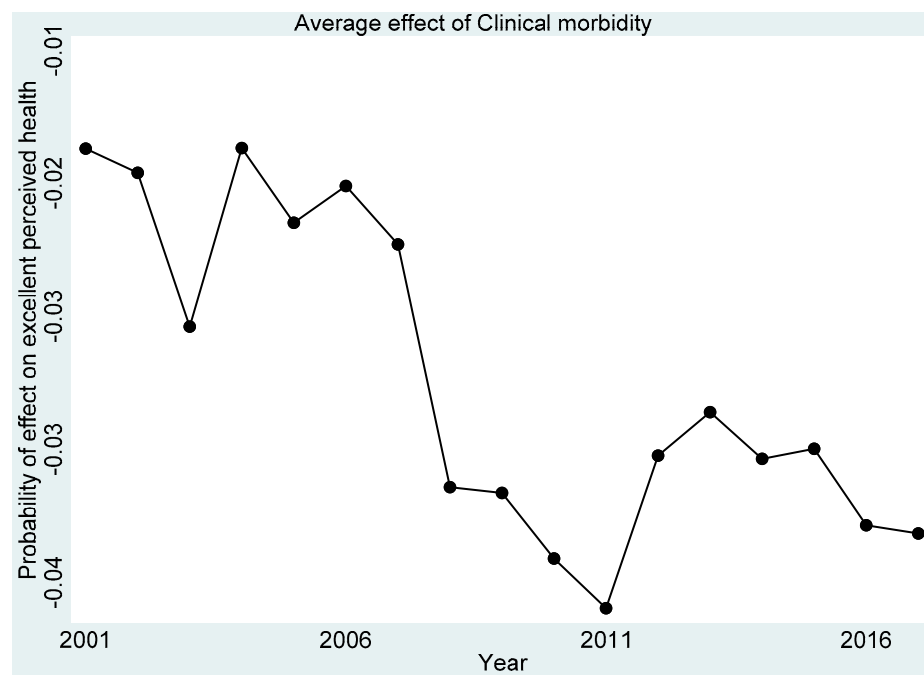


Figure 4. Probability of effects on “good” perceived health from the average effect of clinical morbidity.



**Figure 5.** Probability of effects on “excellent” perceived health from the average effect of clinical morbidity.

#### 4. Discussion

Reporting heterogeneity in SAH is a potentially serious problem for decision making and for achieving distributional efficiency in population health development. Thus, differential reporting by a subgroup of the population deserves an objective assessment. The concern for the reporting heterogeneity often correlated with clinical morbidity limits the unbiased use of SAH for measuring the performance of the resources allocated to the health system. Our data provided us a unique opportunity to examine the disentangled effect of a system-biased health measure, i.e., clinical morbidity (diagnosed chronic disease), on SAH. We introduced the application of a recursive model to examine reporting heterogeneity. Furthermore, we empirically validated the competitiveness of the recursive model over the classical approach.

A reasonably strong and significant correlation with a negative sign between clinical morbidity and perceived health status suggested that clinical morbidity was an endogenous regressor for SAH. The positive value of the duration of suffering interacted with the corresponding clinical morbidity, plausibly reflecting adaptation to the suffering with the chronic disease over time. This revelation confirmed a probable psychological adaptation of patients to the chronic disease they had been suffering from [8]. Such an interaction effect was possible to capture only in the recursive model without any computational stress.

Furthermore, the computation of the conditional probabilities in the framework of the recursive semi-ordered probit model established the distributional effect of clinical morbidity at different levels of perceived health. Predictions and estimations are not the same: the recursive model allowed us to estimate the conditional probabilities. Our findings confirmed that inferences from conditional probabilities were greatly superior forms of evidence for decision making.

Our study contributes further to the literature that argues SAH alone is not enough for policy conclusions, with a particular concern for the perceived health status reported as “average”, confirming the result of Au and Johnston [9]. Therefore, having a control variable tagged to the SAH question is a necessity in the survey questionnaire for effectively addressing the problem of reporting heterogeneity by subgroups of the population.

We used pooled data of 17 waves from a nationally representative sample. However, repeated observations on the same individuals were not systematic, thus it is possible that this study could not well-capture time-varying phenomena concerning the period of suffering in the construct of perceived health for individual adaptability. Our results also support the positive contributions of the instruments used, although we could not eliminate the possible recall bias in the date of diagnosis of the chronic disease. A positive and statistically strongly significant effect of the interaction variable, i.e., duration of suffering, with the clinical morbidity variable, confirmed the appropriateness of the IVs used. Moreover, the sample size used in the study was large enough to verify the stability of the IV-result across time and across different classes of SAH (levels of perceived health), which was quite reassuring.

## 5. Conclusions

To conclude, this study has introduced a new tool in the methodological approach for examining the endogeneity effect on the reporting behaviour for SAH. The methodological approach that we demonstrated can also be used for guiding distributional dimensions of the variable of interest regarding public policy performance. Controlling for endogeneity to establish a “causal” effect necessitates operationalising a simultaneous binary and ordered probit-semi ordered approach. Hence, this empirical validation of the recursive semi-ordered probit model contributes to the literature with the possibilities for replications of model use, where the aim is to estimate the distribution of the endogeneity effect on the variable of interest instead of just examining a point-estimate of the bias. Notwithstanding sample-specific results, we also highlight the policy dimension of heterogeneity in the impact of clinical morbidity on SAH. Perceived health status determines the demand for healthcare service consumption and hence the findings and approach from this study provide insights into public policy development for healthcare service production and the organisation of health systems.

**Author Contributions:** P.P. concept development, literature review, study design, data acquisition and organisation, data analysis, interpretation of results from the analysis, and writing of the manuscript. U.N.: software syntax development. N.K.: construction of mathematical expressions and validation of the models used. B.V.: provided critical comments and revised with intellectual considerations. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research study was funded by the French National Research Agency under grant number ANR-17-EURE-0020.

**Institutional Review Board Statement:** “Not applicable”. This study uses survey data.

**Informed Consent Statement:** “Not applicable”. This study uses survey data that is anonymised.

**Data Availability Statement:** Data used in this study comes from <http://www.cpc.unc.edu/rims>, accessed on 21 February 2021.

**Acknowledgments:** The authors express sincere gratitude to Olivier BEAUMAIS of the University of Rouen Normandy, France, for motivating use of the methods applied in this study. The authors are grateful for the funding support received from the French National Research Agency (grant number: ANR-17-EURE-0020).

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

### Naive model-ordered probit

The dependent variable, perceived health, is ordered in nature with the assumption of an underlying latent (true) health status of an individual,  $i$ . Thus, the ordered nature of the dependent variable directs us to use an ordered probit (OP) model. The observed scale scores on  $y$  are assumed to be discretised measurements on an otherwise continuous but latent response variable  $y^*$ .

$$y_i^* = \alpha + x_i\beta + u_i \quad (\text{A1})$$

$y_i^*$  = the latent index of reported health. The response variable in the model is unobserved;  $y^*$  is connected to  $y$  through a series of cut-points.

$$y = 1 \equiv y^* \leq \alpha_1 \quad (\text{A2})$$

$$y = 2 \equiv \alpha_1 < y^* \leq \alpha_2 \quad (\text{A3})$$

$$y = 3 \equiv \alpha_2 < y^* \leq \alpha_3 \quad (\text{A4})$$

$$y = 4 \equiv \alpha_3 < y^* \leq \alpha_4 \quad (\text{A5})$$

$$y = 5 \equiv y^* > \alpha_4 \quad (\text{A6})$$

The cut point  $\alpha_j$  partitions the latent variable for  $j$  categories of  $y$ . Hence, for a  $j$ -category response variable,  $j - 1$  cut-points fully partition  $y^*$ . Thus, the latent index measures an individual's scale of health. Once  $y_i^*$  crosses a certain value (threshold), perceived health is reported as 'very bad', 'bad', 'average', 'good', and 'excellent'.

$$y_i = j \Leftrightarrow k_{j-1} \leq y_i^* \leq k_j, \quad j = 1, \dots, 5. \quad (\text{A7})$$

$$E[u_i] = 0 \quad (\text{A8})$$

$$\text{Var}[u_i] = 0 \quad (\text{A9})$$

Here,  $k_j$  are the unknown threshold parameters and  $\beta$ s are the unknown coefficients. Thus, the threshold divides the real line ( $y^*$ ) into  $j$  categories.

Observable and unobservable factors influence the latent variable ( $y_i^*$  is a function of observed and unobserved variables). The OP model follows that the coefficient vector  $\beta$  is the same for all categories of  $j$ , implying that the independent variable can shift the cumulative distribution either to the left or right, but the slope remains constant: the probability curve changes, leaving the slope unchanged.

Table A1 of Appendix A reports the marginal effect of clinical morbidity evaluated at each level of perceived health. The negative effect of clinical morbidity on the probability of "good" perceived health was strongest across all the years.

**Table A1.** Marginal effects of clinical morbidity on perceived health of the naive model.

Marginal Effect of Clinical Morbidity	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Pr(perceived health status ≤ very bad), perceived health status = very bad	0.015 (0.000)	0.016 (0.000)	0.014 (0.000)	0.014 (0.000)	0.013 (0.000)	0.014 (0.000)	0.013 (0.000)	0.013 (0.000)	0.012 (0.000)	0.012 (0.000)	0.011 (0.000)	0.012 (0.000)	0.013 (0.000)	0.012 (0.000)	0.012 (0.000)	0.011 (0.000)	0.011 (0.000)
Pr(very bad < perceived health status ≤ bad), perceived health status = bad	0.067 (0.002)	0.070 (0.002)	0.063 (0.002)	0.065 (0.001)	0.061 (0.001)	0.063 (0.001)	0.061 (0.001)	0.059 (0.002)	0.058 (0.002)	0.056 (0.001)	0.054 (0.001)	0.058 (0.001)	0.059 (0.001)	0.057 (0.001)	0.056 (0.001)	0.053 (0.001)	0.053 (0.001)
Pr(bad < perceived health status ≤ average), perceived health status = average	0.178 (0.004)	0.172 (0.004)	0.185 (0.004)	0.182 (0.004)	0.188 (0.004)	0.185 (0.004)	0.188 (0.004)	0.192 (0.004)	0.193 (0.004)	0.196 (0.004)	0.199 (0.004)	0.192 (0.004)	0.192 (0.004)	0.195 (0.004)	0.196 (0.004)	0.200 (0.004)	0.200 (0.004)
Pr(average < perceived health status ≤ good), perceived health status = good	-0.227 (0.004)	-0.227 (0.004)	-0.225 (0.004)	-0.226 (0.004)	-0.224 (0.004)	-0.225 (0.004)	-0.223 (0.004)	-0.222 (0.004)	-0.221 (0.004)	-0.219 (0.004)	-0.216 (0.004)	-0.221 (0.004)	-0.222 (0.004)	-0.219 (0.004)	-0.219 (0.004)	-0.215 (0.004)	-0.215 (0.004)
Pr(good < perceived health status ≤ excellent), perceived health status = excellent	-0.033 (0.001)	-0.030 (0.001)	-0.037 (0.001)	-0.035 (0.001)	-0.039 (0.001)	-0.037 (0.001)	-0.039 (0.001)	-0.042 (0.001)	-0.042 (0.001)	-0.045 (0.001)	-0.048 (0.001)	-0.042 (0.001)	-0.041 (0.001)	-0.044 (0.001)	-0.045 (0.001)	-0.049 (0.001)	-0.049 (0.001)

Figures in parentheses indicate standard errors. All values are significant at a  $p$ -value of  $\leq 0.001$ .

## Appendix B

Computation of partial effects on conditional probability:

$$Pr(y_{2i} = k | y_{1i} = j) = \frac{Pr(y_{1i} = j, y_{2i} = k)}{Pr(y_{1i} = j)}, \quad (A10)$$

where the marginal univariate probability is

$$Pr(y_{1i} = j) = \Phi(A_{Uj}) - \Phi(A_{Lj}). \quad (A11)$$

In our model specification,  $y_{1i}$  takes on two values (0,1). Thus, we compute the difference in conditional probabilities, i.e.,

$$Dif_{ki} = Pr(y_{2i} = k | y_{1i} = 1) - Pr(y_{2i} = k | y_{1i} = 0) \quad (A12)$$

to estimate the effect of  $y_{1i}$  on  $y_{2i}$ .

Computing differences in conditional probabilities allows for the assessment of whether the differences are significant and allows the calculations concerning the partial effect on the differences. Therefore,

$$Dif_{ki} = \frac{Pr(y_{1i} = 1, y_{2i} = k)}{Pr(y_{1i} = 1)} - \frac{Pr(y_{1i} = 0, y_{2i} = k)}{Pr(y_{1i} = 0)} \quad (A13)$$

The average difference is obtained by averaging the individual differences over all observations.

Computation of the partial effect, when  $\rho = 0$ ,

$$\frac{\delta Dif_{ki}}{\delta x_1} = \left( \frac{\frac{\delta Pr(y_{1i}=1, y_{2i}=k)}{\delta x_1}}{Pr(y_{1i}=1)} - Pr(y_{2i} = k | y_{1i} = 1) \frac{\Phi(A_{U1}) - \Phi(A_{L1})}{Pr(y_{1i}=1)} (-\beta_1) \right) - \left( \frac{\frac{\delta Pr(y_{1i}=0, y_{2i}=k)}{\delta x_1}}{Pr(y_{1i}=0)} + Pr(y_{2i} = k | y_{1i} = 0) \frac{\Phi(A_{U0}) - \Phi(A_{L0})}{Pr(y_{1i}=0)} (-\beta_1) \right) \text{ and} \quad (A14)$$

$$\frac{\delta Dif_{ki}}{\delta x_2} = \left( \frac{\frac{\delta Pr(y_{1i}=1, y_{2i}=k)}{\delta x_2}}{Pr(y_{1i} = 1)} - \frac{\frac{\delta Pr(y_{1i}=0, y_{2i}=k)}{\delta x_2}}{Pr(y_{1i} = 0)} \right) \quad (A15)$$

Thus, the marginal effect of  $y_{1i}$  on  $y_{2i}$  is

$$DI f_{ki} = \frac{Pr(y_{1i}=1, y_{2i}=k)}{Pr(y_{1i}=1)} - \frac{Pr(y_{1i}=0, y_{2i}=k)}{Pr(y_{1i}=0)} = \frac{Pr(y_{1i}=1) \times Pr(y_{2i}=k | y_{1i}=1)}{Pr(y_{1i}=1)} - \frac{Pr(y_{1i}=0) \times Pr(y_{2i}=k | y_{1i}=0)}{Pr(y_{1i}=0)} \quad (A16)$$

$$= Pr(y_{2i} = k | y_{1i} = 1) - Pr(y_{2i} = k | y_{1i} = 0). \quad (A17)$$

## References

1. Kerkhofs, M.; Lindeboom, M. Subjective health measures and state dependent reporting errors. *Health Econ.* **1995**, *4*, 221–235. [[CrossRef](#)] [[PubMed](#)]
2. Groot, W. Adaptation and scale of reference bias in self-assessments of quality of life. *J. Health Econ.* **2000**, *19*, 403–420. [[CrossRef](#)]
3. Sadana, R.; Mathers, C.D.; Lopez, A.D.; Murray, C.J.L.; Iburg, K. *Comparative Analysis of More than 50 Household Surveys on Health Status*; GPE Discussion Paper No. 15. EIP/GPE/EBD; World Health Organisation: Geneva, Switzerland, 2000.
4. Murray, C.J.L.; Tandon, A.; Salomon, J.; Mathers, C.D.; Sadana, R. *Cross-Population Comparability of Evidence for Health Policy*; GPE Discussion Paper No. 46. WHO/EIP; World Health Organisation: Geneva, Switzerland, 2001.
5. Shmueli, A. Reporting Heterogeneity in the Measurement of Health and Health-Related Quality of Life. *Pharmacoeconomics* **2002**, *20*, 405–412. [[CrossRef](#)] [[PubMed](#)]
6. Baron-Epel, O.; Kaplan, G.; Haviv-Messika, A.; Tarabeia, J.; Green, M.S.; Kaluski, D.N. Self-reported health as a cultural health determinant in Arab and Jewish Israelis: MABAT—National Health and Nutrition Survey 1999–2001. *Soc. Sci. Med.* **2005**, *61*, 1256–1266. [[CrossRef](#)]

7. Bailis, D.S.; Segall, A.; Chipperfield, J. Two views of self-rated general health status. *Soc. Sci. Med.* **2003**, *56*, 203–217. [[CrossRef](#)]
8. Goldberg, P.; Guéguen, A.; Schmaus, A.; Nakache, J.-P. Longitudinal study of associations between perceived health status and self reported diseases in the French Gazel cohort. *J. Epidemiol. Community Health* **2001**, *55*, 233–238. [[CrossRef](#)]
9. Au, N.; Johnston, D.W. Self-assessed health: What does it mean and what does it hide? *Soc. Sci. Med.* **2014**, *121*, 21–28. [[CrossRef](#)]
10. Layes, A.; Asada, Y.; Kephart, G. Whiners and deniers—What does self-rated health measure? *Soc. Sci. Med.* **2012**, *75*, 1–9. [[CrossRef](#)] [[PubMed](#)]
11. López-Casasnovas, G.; Saez, M. Saved by Wealth? Income, Wealth, and Self-Perceived Health in Spain during the Financial Crisis. *Int. J. Environ. Res. Public Health* **2020**, *17*, 7018. [[CrossRef](#)]
12. Kaneva, M.; Gerry, C.J.; Baidin, V. The effect of chronic conditions and multi-morbidity on self-assessed health in Russia. *Scand. J. Public Health* **2018**, *46*, 886–896. [[CrossRef](#)] [[PubMed](#)]
13. Darviri, C.; Artemiadis, A.K.; Tigani, X.; Alexopoulos, E.C. Lifestyle and self-rated health: A cross-sectional study of 3,601 citizens of Athens, Greece. *BMC Public Health* **2011**, *11*, 619. [[CrossRef](#)] [[PubMed](#)]
14. Demirchyan, A.; E Thompson, M. Determinants of self-rated health in women: A population-based study in Armavir Marz, Armenia, 2001 & 2004. *Int. J. Equity Health* **2008**, *7*, 25. [[CrossRef](#)]
15. Asfar, T.; Ahmad, B.; Rastam, S.; Mulloli, T.P.; Ward, K.D.; Maziak, W. Self-rated health and its determinants among adults in Syria: A model from the Middle East. *BMC Public Health* **2007**, *7*, 177. [[CrossRef](#)]
16. Kasmel, A.; Helasoja, V.; Lipand, A.; Prättälä, R.; Klumbiene, J.; Pudule, I. Association between health behaviour and self-reported health in Estonia, Finland, Latvia and Lithuania. *Eur. J. Public Health* **2004**, *14*, 32–36. [[CrossRef](#)]
17. Tay, J.B.; Kelleher, C.C.; Hope, A.; Barry, M.; Nic Gabhainn, S.; Sixsmith, J. Influence of sociodemographic and neighbourhood factors on self rated health and quality of life in rural communities: Findings from the Agriproject in the Republic of Ireland. *J. Epidemiol. Community Health* **2004**, *58*, 904–911. [[CrossRef](#)]
18. Hernández-Quevedo, C.; Jones, A.M.; Rice, N. *Reporting Bias and Heterogeneity in Self-Assessed Health. Evidence from the British Household Panel Survey. Health, Econometrics and Data Group (HEDG); Working Papers No. 05/04; HEDG, Department of Economics, University of York: Heslington, UK, 2005.*
19. Krieger, N. A glossary for social epidemiology. *J. Epidemiol. Community Health* **2001**, *55*, 693–700. [[CrossRef](#)]
20. Lindeboom, M.; van Doorslaer, E. Cut-point shift and index shift in self-reported health. *J. Health Econ.* **2004**, *23*, 1083–1099. [[CrossRef](#)]
21. Krause, N.M.; Jay, G.M. What Do Global Self-Rated Health Items Measure? *Med. Care* **1994**, *32*, 930–942. [[CrossRef](#)] [[PubMed](#)]
22. Kempen, G.I.; Miedema, I.; Bos, G.A.V.D.; Ormel, J. Relationship of Domain-Specific Measures of Health to Perceived Overall Health among Older Subjects. *J. Clin. Epidemiol.* **1998**, *51*, 11–18. [[CrossRef](#)]
23. Simon, J.; De Boer, J.; A Joung, I.M.; Bosma, H.; MacKenbach, J.P. How is your health in general? A qualitative study on self-assessed health. *Eur. J. Public Health* **2005**, *15*, 200–208. [[CrossRef](#)] [[PubMed](#)]
24. Zhang, T.; Chaojie, L.; Ziling, N. Association of Access to Healthcare with Self-Assessed Health and Quality of Life among Old Adults with Chronic Disease in China: Urban Versus Rural Populations. *Int. J. Environ. Res. Public Health* **2019**, *16*, 2592. [[CrossRef](#)] [[PubMed](#)]
25. Park, J.M. Health status and health services utilization in elderly Koreans. *Int. J. Equity Health* **2014**, *13*, 1–7. [[CrossRef](#)] [[PubMed](#)]
26. Ellis, R.P.; Fiebig, D.G.; Johar, M.; Jones, G.; Savage, E. Explaining Health Care Expenditure Variation: Large-Sample Evidence Using Linked Survey and Health Administrative Data. *Health Econ.* **2013**, *22*, 1093–1110. [[CrossRef](#)] [[PubMed](#)]
27. Greene, W.H.; Hensher, D.A. *Modeling Ordered Choices*; Cambridge University Press (CUP): Cambridge, UK, 2010.
28. Jones, A.M. Identification of treatment effects in Health Economics. *Health Econ.* **2007**, *16*, 1127–1131. [[CrossRef](#)] [[PubMed](#)]
29. Kawatsu, H.; Largey, A.G. EM Algorithms for Ordered Probit Models with Endogenous Regressors. *Econom. J.* **2009**, *12*, 164–186. [[CrossRef](#)]