The Economic Research Guardian – Vol. 2(2)2012
Semi-annual Online Journal, www.ecrg.ro
ISSN: 2247-8531, ISSN-L: 2247-8531
Econ Res Guard 2(2): 160-179

MULTIDIMENSIONAL HEALTH MODELLING: ASSOCIATION BETWEEN SOCIOECONOMIC FACTORS AND HEALTH IN LATVIA

Irina Možajeva
University of Latvia, Riga, Latvia
E-mail: irina.mozajeva@inbox.lv

Abstract

This paper proposes new approach for modelling self-assessed health. We find that the concept of health is too complicated to measure effects of health determinants using a one-dimensional econometric model. We apply two-dimensional stereotype logistic model that allows capturing nonmonotonicity in effects of factors and revealing significant effects that remain unrevealed if single dimension models, such as ordered logit or ordered probit, are used. Modelling self-assessed health using multi-dimensional stereotype logit provides higher model goodness of fit and quality measures in comparison to ordered probit model.

Multi-dimensional stereotype logit is applied to estimate association between socioeconomic factors and self-assessed health in Latvia. While research on socioeconomic determinants of health in Latvia is scarce, this paper provides new insights into the problem.

Keywords: Self-assessed health, Socioeconomic determinants, Nonmonotonicity, Stereotype logit

JEL classification: I10, I18, C52

1. Introduction

Reducing socioeconomic health inequalities is one of the main challenges within the public health sector in Europe. Nature of health inequalities varies for EU member states and Consortium of Partners for Equity in Health admits that there is no a single rule for tackling health inequalities, and country-specific data are essential to elaborate efficient policy.

Health inequalities exist not only within, but also between EU member states. There is a 10 year difference in life expectancy at birth between Switzerland, Spain and Italy (82 years) on the one hand and Latvia (72 year) on the other hand (WHO, 2011).

Poor population health indicators and significant social stratification[1] in Latvia defines the necessity of action to tackle health inequalities and to promote overall population health level. Only using country-specific information on main health determinants can ensure development of efficient national health policy, however econometric analysis of socioeconomic health determinants in Latvia is still scarce (Monden, 2004; Mackenbach, 2006; Mackenbach et.al., 2008).

This paper proposes new approach to modelling self-assessed health (SAH). We see possible problems in measuring association between socioeconomic determinants and SAH using single-dimension models. Respondents might assess their health status not just along single dimension, but rather thinking of two or more latent variables. Therefore *health might not be monotonically related to underlying variables*. If this is true, the model should be able to specify multiple equations to capture effects of these variables. Stereotype logistic model developed by Anderson (1984) provides possibility to measure effects of factors in more than one dimension. In a multinomial logistic model, the categories cannot be ranked, while in ordered logistic model the categories follow a natural ranking scheme. Stereotype logistic model can be seen as a compromise between those two models.

Stereotype logistic models are useful when researcher is not sure of the relevance of the ordering; this problem is common when SAH is used – if some two health outcomes seem similar to a respondent, he or she might be randomly picking between the two. One alternative is to combine these categories and use multinomial logistic model; however in this paper we offer a flexible alternative – stereotype logistic model. The model allows indicating whether all the categories are distinguishable and which are not.

In this paper we apply multi-dimensional stereotype logistic model to estimate association between socioeconomic factors and SAH in Latvia.

While use of self-assessed health status as a measure of health is common in empirical research, many authors admit that Likert type SAH scales should be used cautiously for the assessment of health inequalities. Some studies indicate that this type of SAH scale implies heterogeneity bias. When SAH and more 'objective' health indictors (e.g. McMaster Health Utility Index or clinical health) were used, it was found that in Canada and Britain lower income individuals were more likely to report poor level of SAH than higher income groups (Humphries and van Doorslaer, 2000; Hernandez-Quevedo et al., 2004). At the same time in Germany richer respondents for a given level of clinical health provide lower health assessment (Jürges, 2008). In France reporting heterogeneity was found for the choice between the medium labels i.e. "fair" vs. "good" and for high-income individuals (Etile and Milcent, 2006). In USA given similar diagnosed health conditions and severity levels females rate their health lower than males; divorced, widowed or separated individuals provide lower health assessment than married or never married individuals (Dodoo, 2006).

Another problem of the very good to very poor health scale is its nonstability (Crossleyand Kennedy, 2000); people often face difficulties in assessing their health in terms of good/fair or fair/poor health and therefore are randomly picking between two categories.

[1] For example, in 2010 Gini index in Latvia was the second greatest among EU member states – 36,1 (Eurostat data).

The Economic Research Guardian – Vol. 2(2)2012
Semi-annual Online Journal, www.ecrg.ro
ISSN: 2247-8531, ISSN-L: 2247-8531
Econ Res Guard 2(2): 160-179

In our research we use less subjective SAH scale which allows reducing reporting bias and respondent's perception odds therefore providing more reliable results for SAH status[2].

Some authors try to avoid mentioned SAH bias using binary logit or probit models for dichotomized multiple-category responses and compare respondents with good health to those who report their health to be "less than good" (Etile and Milcent, 2006; Mackenbach, 2006; Jusot et.al., 2007; Jürges, 2008). But it obviously results in a loss of information and requires the introduction of an arbitrary cut-off point (Wagstaff and van Doorslaer, 1994). Another popular approach is modelling health using ordered logit and probit models (van Doorslaer and Jones, 2003; Bockerman and Ilmakunnas, 2009; Bos and Bos, 2007; Ivensen, 2008). Both principles find support in the handbook for health researchers by WHO and IBRD for surveys that use SAH as dependent variable (O'Donnell et.al. 2008). In this paper we introduce another approach that uses full ordered health scale, helps to identify and cope with the above mentioned random category choice problem, and allows for nonmonotonicity in the effects of factors – multi-dimensional stereotype logistic model.

We have indicated only one study where stereotype logit was applied for modelling SAH: Abreu et al. (2009) analysed stereotype logit among other ordinal regression models. However the author didn't discuss multidimensional effects (one-dimensional stereotype logit model was used) and included into analysis three factors only – age, diabetes and skin colour.

This paper is the first where higher-dimension (two-dimensional) stereotype logit model is applied to estimate association between SAH and socioeconomic factors.

According to our best knowledge, the phenomenon of nonmonotonicity hasn't been discussed in this field before. In this paper we show that SAH is nonmonotonically related to some variables which may imply restrictions on use of ordered logit and probit models for modelling self-assessed health.

The rest of the paper is organised as follows. Section 2 describes the data source and methodology applied. The empirical results are provided in the section 3: while section 3.1. is devoted to analysis of association between SAH and socioeconomic determinants in Latvia, section 3.2. provides comparison of results and quality measures of two-dimensional stereotype logit applied in this study and ordered probit commonly used in the literature for modelling SAH. The main findings are summarised in section 4.

2. Data and methodology

This research is based on population survey that was supported by a grant from the CERGE-EI /GDN. The questionnaire employed in the survey was prepared by the author; helpful comments on the questionnaire were provided by Mihails Hazans, specialists of BISS and CERGE-EI. The survey

[2] Please see the next section.

was implemented in March-April 2008; it's representative of the Latvian population and covered residents aged 15-74; in this research we analyse adults only, i.e. respondents aged 18-74.

Data were collected in face-to-face interviews. While information is available only for one household member, the dataset has enough valid observations for our purposes. After omitting all observations with missing values for health and independent variables we obtain a sample of 921 observations.

Self-assessed health is used as a dependent variable. Respondents were asked to describe state of their health choosing one of the six possible answers: "I never ail/ ail very rarely", "I have had only minor sicknesses", "I have had serious sicknesses that are cured", "I have had serious sicknesses, injuries and I still suffer from them", "I have chronic diseases", "I am disabled[3]". We apply this type of scale in order to minimise the heterogeneity bias related to different perception of categories in various social groups as it is observed in case of Likert type health scales when SAH is measured in terms of good/poor health. Therefore we hope to provide more accurate results for SAH and further econometric analysis. We use a five point scale for our model, combining the last two categories (the last category is too small – 4.1%; furthermore according to our preliminary findings the last two groups are not statistically different).

Applying stereotype logistic model we estimate association between socioeconomic factors and self-assessed health. List of socioeconomic factors includes gender, age, labour status, marital status, income per one household member, education, place of residence and ethnicity (see Table A1 with descriptive statistics).

As it was already mentioned, stereotype logistic regression model (Anderson, 1984) applied in this research allows specifying multiple equations to capture the effects of variables. Unlike with multinomial logit, the number of equations one specifies could be less than m–1, where m is the number of categories of the dependent variable.

In the multinomial logistic model, you estimate m–1 parameter vectors β_k, $k = 1 \ldots m$–1. In the stereotype logistic model there are d parameter vectors, where d is between one and min(m–1, p), and p is the number of regressors. The relationship between the stereotype model's coefficients β_j, $j = 1, \ldots d$, and the multinomial model's coefficients is:

$$\beta_k = -\sum_{j=1}^{d} \phi_{jk} \beta_j \,. \qquad (1)$$

The ϕs are scale parameters to be estimated along with the $\beta_j s$. Given a row vector of covariates x, let $\eta_k = \theta_k - \sum_{j=1}^{d} \phi_{jk} x \beta_j$.

[3] Officially recognized.

The Economic Research Guardian – Vol. 2(2)2012
Semi-annual Online Journal, www.ecrg.ro
ISSN: 2247-8531, ISSN-L: 2247-8531
Econ Res Guard 2(2): 160-179

The probability of observing outcome k is:

$$Pr(Y_i = k) = \begin{cases} \dfrac{exp(\eta_k)}{1 + \sum_{l=1}^{m-1} exp(\eta_l)} & k < m \\ \dfrac{1}{1 + \sum_{l=1}^{m-1} exp(\eta_l)} & k = m. \end{cases} \quad (2)$$

If $d = m-1$, the stereotype logistic model is just a reparameterization of the multinomial logistic model. To identify the ϕs and the βs, at least d^2 restrictions on the parameters are essential. By default stereotype logit uses the "corner constraints" $\phi_{jj} = 1$ and $\phi_{jk} = 0$ for $j \neq k$, $k \leq d$, and $j \leq d$ (StataCorp LP, 2005).

In this paper we apply two-dimensional stereotype logistic model to estimate association between socioeconomic factors and SAH in Latvia.

Single dimension models, such as ordered logit and probit, assume that dependent variable is monotonically related to factors, i.e. that a factor can affect dependent variable in one direction only – positively or negatively; this means that if at the beginning of the health scale the sign of the effect is negative, at the end of the scale it should be positive. Multi-dimensional stereotype logit allows for nonmonotonicity, i.e. the direction of impact of a factor on health may change along the scale. In the case of a two-dimensional model, if a factor has positive effect in the first dimension (in this case probability of very good health increases), the sign of the effect also can be positive in the second dimension in the end of the scale (probability of very poor health increases). It is possible also that a factor has statistically significant effect in one dimension, but is not significant in the other one. In both cases this indicates that a factor is nonmonotonically related to the dependent variable.

This can be illustrated using vectors (Figure 1). In a two-dimensional stereotype logit model a factor that is represented by vector a affects dependent variable in different directions: positively in the first dimension, and negatively in the second dimension of the model. Vector b illustrates a situation when effect of a factor is not statistically significant in the first dimension, however is significant in the second one. Vector c represents a factor that has statistically significant effect and keeps the direction of impact (negative) in both dimensions. If the direction of impact is the same in both dimensions, a factor is monotonically related to a dependent variable; otherwise the relationship is nonmonotonic. Empirically such cases are analysed in the next section.

The Economic Research Guardian – Vol. 2(2)2012
Semi-annual Online Journal, www.ecrg.ro
ISSN: 2247-8531, ISSN-L: 2247-8531
Econ Res Guard 2(2): 160-179

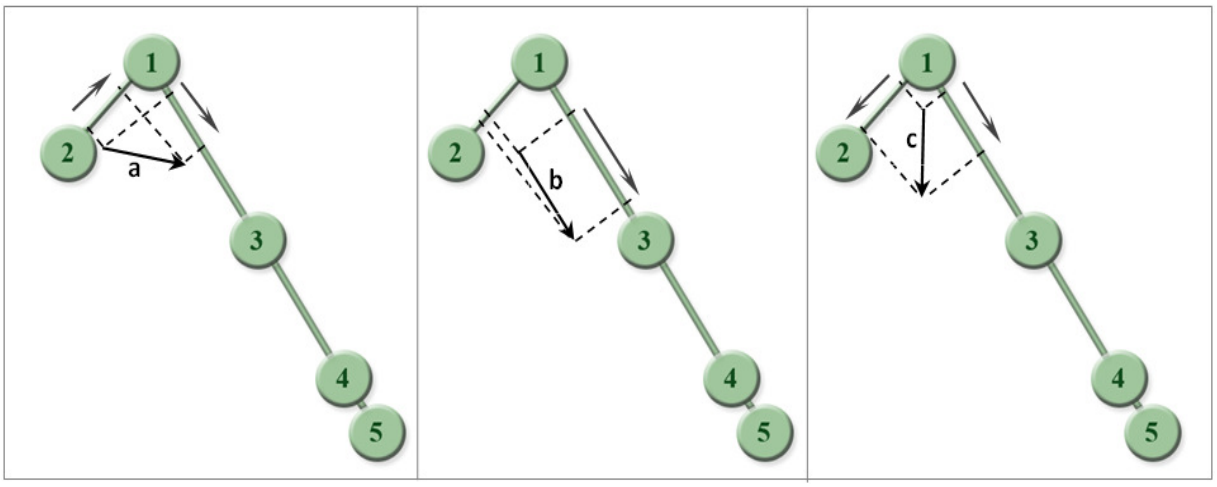

Figure 1 - Illustration of the effects of factors in the two-dimensional stereotype logit model (examples)

3. Results

As it was mentioned above, the model estimated analyses association between factors and SAH in two different dimensions. The first dimension of the model describes effects of factors when the second health outcome (Might have only minor sicknesses) is compared to the first health outcome (Never ails/ ails very rarely) (see Table A2 in the Appendix). The effects of the second dimension are measured when the third, fourth and fifth health outcomes are compared to the base outcome, i.e. the first health category.

As it was mentioned above, stereotype logit model allows indicating whether all the health categories were distinguishable for respondents. Equal coefficients for the fourth and the fifth outcomes in the second dimension state that the difference between these two health categories is not statistically significant (Table A2). This proposes that respondents with serious health problems faced difficulties when choosing one of these categories and could be randomly picking between the two. This case of two-dimensional stereotype logit model is illustrated on Figure 1.

3.1. Association between health and socioeconomic determinants

Table 1 presents results of two-dimensional stereotype logistic model designed to estimate impact of socioeconomic factors on SAH. Marginal effects show increase or decrease of probability of according health outcome for each factor after accounting for all the other factors[4]. Percent above

[4] Precise levels of significance are provided in Table 3 in the Appendix.

The Economic Research Guardian – Vol. 2(2)2012
Semi-annual Online Journal, www.ecrg.ro
ISSN: 2247-8531, ISSN-L: 2247-8531
Econ Res Guard 2(2): 160-179

each health category shows average probability of according health outcome. To be simple and to avoid too long expressions further in the text we will use definition "very good health" to describe group of respondents who never ail/ail rarely, "good health" will be used to describe those who have had only minor sicknesses etc. However please bear in mind that the original health scale used in the survey was not a Likert type scale.

Table 1 - Association between socioeconomic factors and self-assessed health in Latvia[5]

Factors		Association between each factor and health (comparison with the reference category, impact of other factors is excluded)				
Mean probabilities		29% Very good	32% Good	15% Fair	10% Poor	14% Very poor
		Never ails/ ails very rarely	Has had only minor sicknesses	Has had serious sicknesses that are cured	Has had serious sicknesses, injuries and still suffers from them	Has chronic diseases/ is disabled
		dP/dX	dP/dX	dP/dX	dP/dX	dP/dX
Female		1.6%	-1.8%	0.4%	-0.1%	-0.1%
Lives in Riga or Riga district (ref. cat.: lives outside Riga district)		0.5%	4.6%	-1.7%	-1.4%	-2.1%
Lives in Riga or Riga district, male		19.6%**	-17.2%***	2.1%	-1.8%	-2.7%
Age		7.8%**	-5.0%	0.5%	-1.3%	-2.0%
$Age^2/100$		-21.4%***	10.7%	-0.4%	4.5%*	6.6%*
$Age^3/1000$		1.6%***	-0.8%	0.0%	-0.4%*	-0.5%*
Single (ref. cat.: married or lives with a partner)		-3.9%	3.4%	-0.6%	0.4%	0.7%
Single, female		2.2%	-11.3%	3.5%	2.3%	3.4%
Divorced or widowed, female		4.8%	-21.5%***	6.5%***	4.1%*	6.1%*
Labour status (ref. cat.: employed / student)	Economically inactive	-23.3%***	-10.6%*	2.5%	12.7%***	18.7%***
	Unemployed	-2.3%	-8.9%	3.1%	3.3%	4.8%
Ethnic non-Latvian		-2.2%	4.6%	-1.3%	-0.5%	-0.7%
Education (ref. cat.: higher / incomplete higher)	Below secondary	8.1%	-19.4%***	5.4%***	2.4%	3.5%
	Secondary / vocational secondary	6.8%*	-2.5%	-0.2%	-1.7%	-2.5%

Notes: Asterisks *, **, *** indicate a statistically significant difference from the base group at 10%, 5%, 1% level respectively.
Other factors controlled: Average income per household member

Most studies addressing SAH in different countries record large gender differences with women reporting significantly worse health than men (Walters and Suhrcke, 2005). Gender health gap is also observed in Latvia with lower SAH level for females[6]. However we do not find statistically significant difference between male and female reported health when all other socioeconomic factors are controlled (see Table 1). This means that while in absolute terms gender disparities are still present in Latvia, the source of these disparities is found in unequal distribution of favourable socioeconomic factors, as well as in different impact of specific factors on male and female health. According to the obtained results, marital status is one of such factors.

[5] Author's calculations using „Health Survey 2008" data.
[6] Author's calculations using "Health Survey 2008" data.

Place of residence is associated with male and female health in a different manner as well. In this model we compare residents of Riga and Riga district to those who live outside the district. This analysis might be more interesting than urban-rural comparison since economic activity in Latvia is highly concentrated in Riga and about a third of all residents of Latvia live in this city[7].

While the difference in SAH between women living in Riga or Riga district and women living outside the district is not observed, the effect of place of residence for males is rather strong. According to the results, the variable is nonmonotonically related to health and its effect is significant only in the first dimension: male residents of Riga have greater chance to have very good health, but lower probability of good health (other parameters equal). In the second dimension of the model, the effect of place of residence is not statistically significant.

The effect of the place of residence variable might have its rise in the process of labour force migration that was rather intensive before crisis – major part of young active people living in different regions of Latvia (Kurzeme, Vidzeme, Latgale) has moved to the capital or abroad, which resulted in relative increase of amount of very healthy males in the capital and its district and reduction of amount of such males in other regions. However this still doesn't provide an explanation for the negative effect of place of residence variable for the second health outcome. Deeper analysis that is out of means of the survey data employed needs to be applied to study the source of this phenomenon.

In this model we use three age variables – linear, squared and cubed. Significance of effects for all the three variables proposes presence of two bending points in the effect of age; these points are found at about 30 and 65 years with an increasing rate of health loss after 30 years and decreasing rate after 65. The second effect might be explained by survivor bias – those who have reached age of retirement can be characterised by comparatively strong organism which reduces health risks and health loss[8].

Despite marriage is generally considered to be positively related to health, we find no statistically significant difference in health between married (or living with partner) and single. No empirical evidence for significant association between SAH and being married was found for residents of Germany and Norway as well (Ivensen, 2008; Jürges, 2008).

While the effect for divorced or widowed males is not statistically significant (this might be due to small size of the group in the sample), we find negative effect for divorced or widowed females: this status for females reduces the probability of good health by 21.5 percent points (which is impressive taking into account that mean probability of this health outcome is 31.6%). According to our preliminary findings, the third health outcome (fair health) is closer to poor health rather than to good health. Taking this into account we can see that status of divorced or widowed female is associated with increase of probability of negative health outcomes.

[7] The number of residents in the second greatest city of Latvia is 5-6 times smaller than in Riga.
[8] One should keep in mind that life expectancy in Latvia is only 71 year (WHO, 2010).

Absence of negative effect of status of divorced or widowed for the females in case of very good health can be explained as follows: very healthy women might go through negative psychological and economical effects of divorce relatively easier than less healthy women. When health is already undermined, impact of such burden may be noticeably stronger. Healthy women obviously are more confident about themselves in terms of prospects for future marriage, job opportunities etc. Due to this divorce in healthy women's life might not provide significant negative effect.

Strong association between economic activity and health has been observed in Latvia already in late 1990s (Monden, 2004). As the model results propose, status of economically inactive[9] still has a particularly strong negative effect on health. The probability of very good health for the group is 23.3 percent points lower than for employed and students which is oppressive taking into account that the mean probability of very good health is 29.3%. The effects in the second model's dimension are negative and strong as well. Association between health and economic inactivity is one of the strongest in the model developed.

The effect of status of unemployed is not found as statistically significant. This can be explained by the fact that job possibilities in early 2008 were still comparatively good, and a large part of those found in this group were frictional unemployed. Rate of unemployment in spring of 2008 was rather low (for Latvia) – about 6.3%[10] (Central Statistical Bureau of Latvia, 2010) and shift from one job to another or short term unemployment didn't provide significant negative impact on health then. However if the survey was conducted a year later, strong negative effect would be expected taking into account high unemployment rate[11] and serious economic and psychological burden experienced by unemployed in 2009.

We have also checked whether retirement has a statistically significant impact on health; when labour status with the three categories is controlled for (the category for economically inactive includes nonworking retirees), the effect of dummy for status of retired is not significant.

We find no statistically significant difference between non-Latvians and Latvians when all other socioeconomic factors are controlled. In 1990s ethnic differences were not identified for SAH in general, although some gap was found for long-standing health problems among women (MondenMonden, 2004). However in absolute terms (i.e. without control for other factors) in 2008 just as in 1990s Latvians on average reported slightly better health than non-Latvians. Probably these differences have their rise from other socioeconomic conditions.

Level of education has a significant effect on population health in Latvia. In late 1990s impact of education was less noticeable; after adjusting for income, educational differences were significant only for women (Monden, 2004
Monden). In 2008 we do not find difference in effect of education on SAH between males and females (other factors controlled).

[9] The group includes nonworking retirees, women on a maternity leave, housewives and disabled.
[10] 2nd quarter of 2008, official data.
[11] Unemployment rate was 16,7% in the 2nd quarter of 2009 (Central Statistical Bureau of Latvia, 2010).

The Economic Research Guardian – Vol. 2(2)2012
Semi-annual Online Journal, www.ecrg.ro
ISSN: 2247-8531, ISSN-L: 2247-8531
Econ Res Guard 2(2): 160-179

According to the obtained results, the difference between residents with higher or incomplete higher education and a group of residents with lower than secondary education is not significant for the extreme outcomes, but it is considerable when we analyse good and fair health: we observe strong negative effect – decrease of probability of good health and increase of probability of fair health – for residents with lower than secondary education.

Higher education doesn't seem to provide advantage in chances to maintain good health in comparison to secondary education in Latvia (other parameters controlled). Quite the contrary – despite one's expectations, the effect of higher education is even negative: those with secondary or vocational secondary education have greater probability of very good health than those with higher or incomplete higher education (other parameters equal).

Literature mostly provides support for negative effect for less educated (Jusot et. al., 2007; Jürges, 2008). In Latvia the observed negative effect for the group of residents with lower than secondary education and relative advantage of those with secondary education in comparison to the most educated residents partly can be explained by differential exposure to serious emotional problems like stress, unrest etc. (Figure 2). According to the survey data, residents with secondary or vocational secondary education are exposed to stress less often than the other two groups.

We do not find convincing empirical support for less educated to have more pronounced adverse behaviour in comparison to residents with secondary education in Latvia. Another possible explanation for the less favourable state of residents with higher education could be found in more intensive work and less time devoted for rest (see Figure 3). The data propose that residents with higher or incomplete higher education on average devote to rest less time than the other two groups; this reduces possibility for the former to maintain very good health. Further investigation that is out of means of this research would be necessary to provide explicit explanation for the absence of advantage in health of the most educated in Latvia (other socioeconomic parameters controlled).

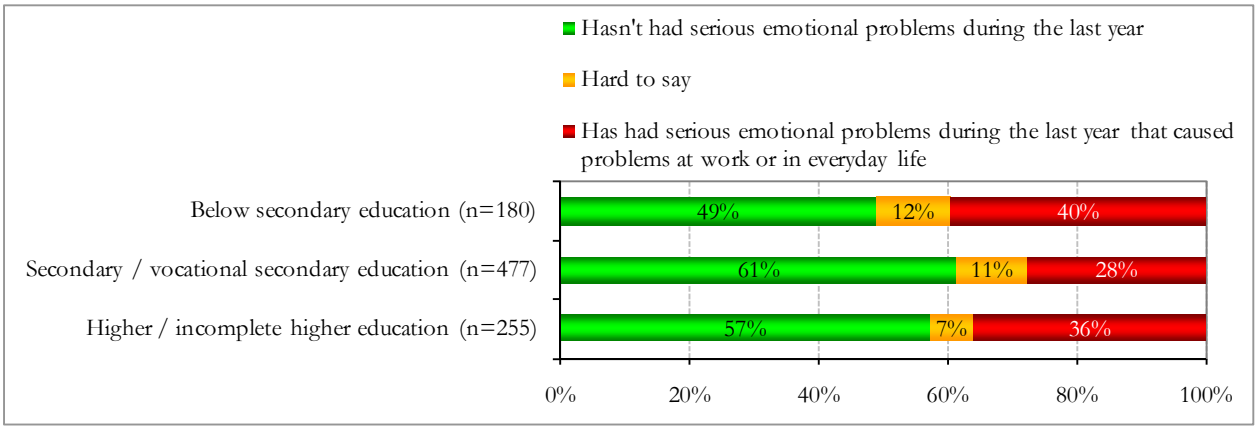

Figure 2 - Exposure to serious emotional problems in different education groups in Latvia, 2008[12]

[12] Author's calculations using „Health Survey 2008" data.

The Economic Research Guardian – Vol. 2(2)2012
Semi-annual Online Journal, www.ecrg.ro
ISSN: 2247-8531, ISSN-L: 2247-8531
Econ Res Guard 2(2): 160-179

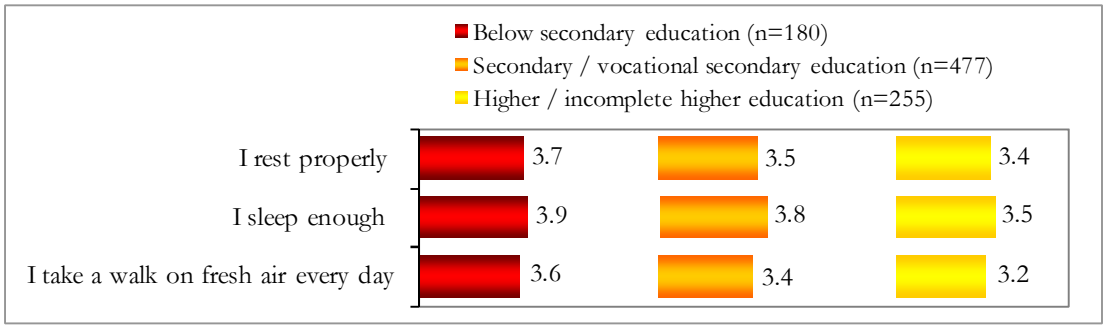

Figure 3 - Evaluation[13] of rest in different education groups in Latvia, 2008[14]

In this paper we do not examine impact of income since data on household income obtained in the survey were not persuasive – level of household income was underreported and to avoid providing misleading results we do not analyse income effect here. However average income per household member is controlled in all the models included into this paper. According to earlier research, income effect is significant in Latvia and seems to be strongly associated with access to psychosocial resources: when psychosocial factors are controlled for, income effect becomes insignificant (Zujeva[15], 2008).

3.2. Ordered Probit vs Two-Dimensional Stereotype Logit

Table 2 provides comparison of the results of the two models – ordered probit and two-dimensional stereotype logit.

Since ordered probit assumes that dependent variable is monotonically related to factors, while stereotype logit allows for nonmonotonicity in effects of variables, we find substantial difference in the results of the two models. For example, according to the stereotype model we find that effect for males who live outside the Riga district is particularly strong for the two first health outcomes (very good and good health). The model proposes that the variable is nonmonotonically related to health: the effect changes its direction – from positive effect on very good health to negative on good health, but moving further along the health scale it doesn't provide statistically significant effect. Therefore when multidimensional approach is applied, some factors may have significant effect on health in one of the two dimensions only. In the ordered probit model the effect of this variable is found as significant as well, however the model distributes the effect along the health scale proposing completely different nature of association between SAH and the factor.

[13] Evaluation on a 5 point scale (1 – very rarely/never; 5 – always).
[14] Author's calculations using „Health Survey 2008" data.
[15] The author of this paper.

The Economic Research Guardian – Vol. 2(2)2012
Semi-annual Online Journal, www.ecrg.ro
ISSN: 2247-8531, ISSN-L: 2247-8531
Econ Res Guard 2(2): 160-179

Table 2 - Association between socioeconomic factors and SAH – comparison of results of two-dimensional stereotype logit model and ordered probit model[16]

Factors			Association between each factor and health (comparison with the reference category, impact of other factors is excluded)				
Mean probabilities			29%	32%	15%	10%	14%
			Very good	Good	Fair	Poor	Very poor
			Never ails/ ails very rarely	Has had only minor sicknesses	Has had serious sicknesses that are cured	Has had serious sicknesses, injuries and still suffers from them	Has chronic diseases/ is disabled
			dP/dX	dP/dX	dP/dX	dP/dX	dP/dX
Female		slogit	1.6%	-1.8%	0.4%	-0.1%	-0.1%
		oprobit	0.7%	0.1%	-0.2%	-0.2%	-0.4%
Lives in Riga or Riga district (ref. cat.: lives outside Riga district)		slogit	0.5%	4.6%	-1.7%	-1.4%	-2.1%
		oprobit	2.9%	0.5%	-1.0%	-1.0%	-1.5%
Lives in Riga or Riga district, male		slogit	19.6%**	-17.2%***	2.1%	-1.8%	-2.7%
		oprobit	12.0%*	0.6%	-4.1%*	-3.7%**	-4.9%**
Age		slogit	7.8%**	-5.0%	0.5%	-1.3%	-2.0%
		oprobit	5.4%**	1.1%**	-1.8%**	-1.9%**	-2.8%**
Age^2/100		slogit	-21.47%***	10.7%	-0.4%	4.5%*	6.6%*
		oprobit	-14.9%***	-3.0%**	4.9%**	5.2%**	7.8%***
Age^3/1000		slogit	1.6%***	-0.8%	0.0%	-0.4%*	-0.5%*
		oprobit	1.1%***	0.2%**	-0.4%**	-0.4%**	-0.6%***
Single (ref. cat: married or lives with a partner)		slogit	-3.9%	3.4%	-0.6%	0.4%	0.7%
		oprobit	-2.0%	-0.4%	0.6%	0.7%	1.1%
Single, female		slogit	2.2%	-11.3%	3.5%	2.3%	3.4%
		oprobit	-1.6%	-0.4%	0.5%	0.6%	0.9%
Divorced or widowed, female		slogit	4.8%	-21.5%***	6.5%***	4.1%*	6.1%*
		oprobit	-5.9%	-1.7%	1.8%	2.2%	3.6%
Labour status (ref. cat: employed / student)	Economically inactive	slogit	-23.3%***	-10.6%*	2.5%	12.7%***	18.7%***
		oprobit	-21.5%***	-9.8%***	5.3%***	8.3%***	17.7%***
	Unemployed	slogit	-2.3%	-8.9%	3.1%	3.3%	4.8%
		oprobit	-5.9%	-1.8%	1.8%	2.2%	3.7%
Ethnic non-Latvian		slogit	-2.2%	4.6%	-1.3%	-0.5%	-0.7%
		oprobit	-0.2%	0.0%	0.1%	0.1%	0.1%
Education (ref. cat: higher / incomplete higher)	Below secondary	slogit	8.1%	-19.4%***	5.4%***	2.4%	3.5%
		oprobit	-1.6%	-0.3%	0.5%	0.6%	0.9%
	Secondary / vocational secondary	slogit	6.8%*	-2.5%	-0.2%	-1.7%	-2.5%
		oprobit	5.0%*	1.0%	-1.6%*	-1.7%*	-2.6%

Notes: Asterisks *, **, *** indicate a statistically significant difference from the base group at 10%, 5%, 1% level respectively.
Other factors controlled: Average income per household member

Thanks to multidimensional approach stereotype logistic regression is able to grasp significant effects of some variables that are seen as statistically insignificant if a single-dimension model is used. For example, ordered probit is not able to reveal significance of the effect of being divorced or widowed for females and the effect of below secondary education. Both factors are nonmonotonically related to health and due to this reason they are not found as significant by ordered probit model.

[16] Author's calculations using „Health Survey 2008" data.

When a factor is monotonically related to dependent variable, two-dimensional stereotype logit and ordered probit provide similar results as it is, for example, for the effect of economic inactivity. If stereotype logit doesn't find significant effect for a variable, ordered probit also doesn't find it (e.g., gender, ethnicity).

Therefore multi-dimensional approach allows revealing nonmonotonicity in effects of some variables as well as disclosing significant effects for some variables that cannot be seen when a one-dimensional model is used.

Table 3 provides measures of goodness of fit and selection criteria for the developed two-dimensional stereotype logit model and ordered probit model. Log pseudolikelihood is calculated instead of log likelihood since we use population weights and robust standard errors are estimated in the models. We also use R^2_O, an explained variation measure for ordinal response models, which is based on ordinal dispersion measure (Lacy, 2006). In this paper we use both AIC (Akaike Information Criterion) (Akaike, 1973) and BIC (Bayesian Information Criterion) (Schwarz, 1978) indicators. In practise usually when one of the criteria (AIC or BIC) improves (becomes smaller) the other one becomes greater. All the parameters analysed in Table 3 except BIC indicate that two-dimensional approach fits the data better. This proposes that two-dimensional stereotype logit in comparison to ordered probit is more appropriate methodology for estimating association between socioeconomic determinants and SAH. Comparison of the results of the two models indicates greater potential of stereotype logit for modelling SAH.

Table 3 - Model selection criteria[17]

	Two-dimensional stereotype logit	Ordered probit	Difference
Log pseudolikelihood	-1226.4	-1263.9	37.6
Lacy R^2_O	22.9%	20.7%	2.2%
AIC	2540.7	2573.8	-33.1
BIC	2752.6	2684.6	68.0
Number of statistically significant effects	11	9	2

4. Summary and conclusions

Results of the two-dimensional stereotype logistic model developed suggest that some socioeconomic factors are nonmonotonically related to SAH. This may imply restrictions on use of one-dimensional ordered models for modelling SAH. Multidimensional approach allows revealing some significant factor effects that remain unrevealed if one-dimensional models, e.g. ordered probit, are used. Analysed goodness of fit and selection criteria for the two-dimensional stereotype logit

[17] Author's calculations using „Health Survey 2008" data.

model and ordered probit model propose that multidimensional approach is more appropriate for modelling self-assessed health.

We have examined impact of economic and social determinants on population health in Latvia. In contradiction to what is mostly found in other countries, in Latvia gender health disparities were not detected (holding all other socioeconomic parameters equal). However in absolute terms we observe lower average SAH indicators for females, which may be explained by differential access to socioeconomic resources for men and women as well as by different nature of impact of some factors (e.g. marital status and place of residence) on male and female health. The three variables for age – linear, squared and cubed – are significant in the models; this proposes existence of two binding points in the effect of age and different rate of health deprivation – increasing rate of health loss after 30 years and decreasing rate after 65 years. The model reveals significant disparities between economically inactive residents and a group of employed and students with strong negative effect for the former (other parameters equal).

The stereotype logistic model uncovers strong negative effect for widowed or divorced females; the effect is nonmonotonic and can be revealed only when multidimensional approach is applied – results of ordered probit, for example, do not provide evidence of statistically significant effect for this factor. The same conclusion can be made about the difference between the group of less educated (below secondary education) and the group with higher or incomplete higher education; the disparities are revealed as statistically significant only when multidimensional model is applied.

Multidimensional stereotype logit allows obtaining more accurate estimation of association between SAH and socioeconomic determinants and revealing nature of this association more precisely than single dimension ordered models and therefore multidimensional approach can aid to development of more efficient health policies.

Acknowledgements

This research was supported by a grant from the CERGE-EI Foundation under a program of the Global Development Network. I express special gratitude to my supervisor Prof. Mihails Hazans for helpful advices in the model development and for inspiring guidance. I also express gratitude to Prof. Gil S. Epstein for helpful recommendations and suggestions. I would like to thank Viatcheslav Vinogradov and Edward Christie for comments on the paper. I thank Baltic Institute of Social Sciences for assistance. I would like to give my special thanks to GDN/ CERGE-EI for the opportunity to implement the research. All opinions expressed are those of the author and have not been endorsed by CERGE-EI or the GDN.

References

Abreu M N S, Siqueira A L, Caiaffa W T (2009). Ordinal Logistic Regression in Epidemiological Studies. *Rev. Saúde Pública*. 43(1): 183-94.

Akaike H (1973). *Information theory and an extension of the maximum likelihood principle*. In Proc. 2nd Int. Symp. Info. Theory, Petrov B N, Csaki F (Ed.), Akademia Kiado, Budapest, pp. 267-281.

Anderson J A (1984). Regression and Ordered Categorical Variables. *Journal of the Royal Statistical Society*. B.46: I-30.

Bockerman P, Ilmakunnas P (2009). Unemployment and Self Assessed Health: Evidence from Panel Data. *Health Economics*. 18: 161-179.

Bos A, Bos A (2007). *The Socio-Economic Determinants of Older People's Health in Brazil: The Importance of Marital Status and Income*. Cambridge University Press.

Crossley T, Kennedy S (2000). The Stability of Self Assessed Health Status. *Discussion Paper Nr.421,* Centre for Economic Policy Research, Australian National University.

Dodoo M (2006). Differential Reporting of Health Status by Sub-Groups of the U.S. Population. *Annual meeting of the Economics of Population Health: Inaugural Conference of the American Society of Health Economists*, TBA, Madison, WI, USA.

Etile F, Milcent C (2006). Income-Related Reporting Heterogeneity in Self-Assessed Health: Evidence from France. *Journal of Health Economics*. 15: 965-981.

Hernandez-Quevedo C, Jones A M, Rice N (2004). *Reporting Bias and Heterogeneity in Self-Assessed Health. Evidence from the British Household Panel Survey*. University of York, mimeo.

Humphries K H, van Doorslaer E (2000). Income-Related Health Inequality in Canada. *Social Science and Medicine*. 50: 663-671.

Iversen T (2008). An Exploratory Study of Associations between Social Capital and Self-Assessed Health in Norway. *Health Economics, Policy and Law*. Cambridge University Press. 3(04): 349-364.

Jürges H (2008). Self-Assessed Health, Reference Levels, and Mortality. *Applied Economics*, Taylor and Francis Journals. 40(5): 569-582.

Jusot F, Grignon M, Dourgnon P (2007). Psychosocial resources and social health inequalities in France: Exploratory findings from a general population survey. *Working paper*, IRDES. 4-6.

Lacy M G (2006). An Explained Variation Measure for Ordinal Response Models with Comparisons to Other Ordinal R^2 Measures. *Sociological Methods and Research*. 34: 469-520.

Mackenbach J P (2006). Health Inequalities: Europe in Profile. *Independent expert report commissioned by the UK Presidency of the EU*.

Mackenbach J P, Stirbu I, Roskam A J R, Schaap M M, Menvielle G, Leinsalu M, Kunst A E (2008). Socioeconomic Inequalities in Health in 22 European Countries. *The New England Journal of Medicine*. 358: 2468-81.

Monden C (2004). Socioeconomic Health Inequalities in Latvia: A Cross-Sectional Study. *Scandinavian Journal of Public Health*. 32(3): 217-23.

Možajeva (Zujeva) I (2008). *The Direct and Indirect Impact of Economic Factors on Population Health in Latvia*. Acta Universitatis Latvensis, series of Economics, University of Latvia, pp. 251-262.

O'Donnell O, van Doorslaer E, Wagstaff A, Lindelow M (2008). *Analyzing Health Equity Using Household Survey Data. A Guide to Techniques and Their Implementation*. The International Bank for Reconstruction and Development/The World Bank.

Schwarz G (1978). Estimating the Dimension of a Model. *Annals of Statistics*. 6: 461-464.

StataCorp LP (2005). *User Manual. Release 9*. College Station: Stata Press.

Van Doorslaer E, Jones A (2003). Inequalities in Self-Reported Health: Validation of a New Approach to Measurement. *Journal of Health Economics*. 22: 61-87.

Wagstaff A, van Doorslaer E (1994). Measuring Inequalities in Health in the Presence of Multiple-Category Morbidity Indicators. *Journal of Health Economics*. 3: 281-291.

Walters S, Suhrcke M (2005). Socioeconomic Inequalities in Health and Health Care Access in Central and Eastern Europe and the CIS: A Review of the Recent Literature. Venice, World Health Organization European Office for Investment for Health and Development; *Working paper 2005/1*.

WHO (2011). *World Health Statistics 2011*. Geneva, World Health Organization.

Central Statistical Bureau of Latvia (2010). Employment and Unemployment - Key Indicators Database, avialable at http://www.csb.gov.lv/en/statistikas-temas/employment-and-unemployment-key-indicators-30679.html.

Eurostat (2011). European Statistics Database, avialable at http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database.

Baltic Institute of Social Sciences (2008). Health module in Language Survey.

The Economic Research Guardian – Vol. 2(2)2012
Semi-annual Online Journal, www.ecrg.ro
ISSN: 2247-8531, ISSN-L: 2247-8531
Econ Res Guard 2(2): 160-179

Appendix

Table A1 - Descriptive statistics: socio-demographic characteristics of the sample

Characteristics	Items	N	%
Self-assessed health	I never ail	270	29.3%
	There might be only minor sicknesses	291	31.6%
	I have had more serious illnesses that have been cured	138	15.0%
	I have had serious illnesses or injuries, and I still suffer from them	91	9.9%
	I have chronic illnesses	92	10.0%
	I am disabled	38	4.1%
Gender	Male	429	46.5%
	Female	492	53.5%
Age	18-24	146	15.8%
	25-34	170	18.4%
	35-44	150	16.3%
	45-54	188	20.4%
	55-64	135	14.6%
	65-74	133	14.4%
Place of residence	Riga and Riga district	286	31.0%
	Outside Riga district	635	69.0%
Labour status	Employed / student	655	71.1%
	Economically inactive	212	23.0%
	Unemployed	54	5.9%
Marital status	Married / lives with partner	557	60.5%
	Single	208	22.5%
	Divorced / lives separately / widowed	156	17.0%
Ethnicity	Ethnic Latvian	536	58.2%
	Ethnic non-Latvian	385	41.8%
Education	Below secondary education	182	19.8%
	Secondary / vocational secondary education	484	52.5%
	Higher / incomplete higher education	243	26.4%
	Unknown	11	1.2%
Average income per household member	I quintile	159	17.3%
	II quintile	138	15.0%
	III quintile	159	17.3%
	IV quintile	148	16.1%
	V quintile	146	15.9%
	Unknown	171	18.6%

The Economic Research Guardian – Vol. 2(2)2012
Semi-annual Online Journal, www.ecrg.ro
ISSN: 2247-8531, ISSN-L: 2247-8531
Econ Res Guard 2(2): 160-179

Table A2 - Scale parameters of two-dimensional stereotype logistic model[18]

1st dimension	/phi1_1	Never ails/ ails very rarely	0	(base outcome)
	/phi1_2	Has had only minor sicknesses	1	
	/phi1_3	Has had serious sicknesses that are cured	omitted	
	/phi1_4	Has had serious sicknesses, injuries and still suffers from them	omitted	
	/phi1_5	Has chronic diseases/ is disabled	omitted	
2nd dimension	/phi2_1	Never ails/ ails very rarely	0	(base outcome)
	/phi2_2	Has had only minor sicknesses	omitted	
	/phi2_3	Has had serious sicknesses that are cured	1	
	/phi2_4	Has had serious sicknesses, injuries and still suffers from them	1.9	
	/phi2_5	Has chronic diseases/ is disabled	1.9	

[18] See formulas 1 an 2 at page 163 and 164.

The Economic Research Guardian – Vol. 2(2)2012
Semi-annual Online Journal, www.ecrg.ro
ISSN: 2247-8531, ISSN-L: 2247-8531
Econ Res Guard 2(2): 160-179

Table A3 - Association between socioeconomic factors and self-assessed health in Latvia (two-dimensional stereotype logit)

Number of observations	Wald chi2(50)	Log pseudolikelihood	df	AIC	BIC	Lacy R2O
912	114.95	-1228.012	43	2542.02	2749.10	22.85%

Factors		Association between each factor and health (comparison with the reference category, impact of other factors is excluded)									
Mean probabilities		29%		32%		15%		10%		14%	
		Very good		Good		Fair		Poor		Very poor	
		Never ails/ ails very rarely		Has had only minor sicknesses		Has had serious sicknesses that are cured		Has had serious sicknesses, injuries and still suffers from them		Has chronic diseases/ is disabled	
		dP/dX	S.E.	dP/dX	S.E.	dP/dX	S.E.	dP/dX	S.E.	dP/dX	S.E.
Female		1.6%	0.050	-1.8%	0.052	0.4%	0.016	-0.1%	0.016	-0.1%	0.024
Lives in Riga or Riga district (ref. cat.: lives outside Riga district)		0.5%	0.050	4.6%	0.059	-1.7%	0.018	-1.4%	0.017	-2.1%	0.025
Lives in Riga or Riga district, male		19.6%**	0.083	-17.2%***	0.059	2.1%	0.030	-1.8%	0.024	-2.7%	0.035
Age		7.8%**	0.033	-5.0%	0.033	0.5%	0.012	-1.3%	0.011	-2.0%	0.017
$Age^2/100$		-21.4%***	0.001	10.7%	0.001	-0.4%	0.000	4.5%*	0.000	6.6%*	0.000
$Age^3/1000$		1.6%***	0.000	-0.8%	0.000	0.0%	0.000	-0.4%*	0.000	-0.5%*	0.000
Single (ref. cat: married or lives with a partner)		-3.9%	0.061	3.4%	0.074	-0.6%	0.024	0.4%	0.027	0.7%	0.039
Single, female		2.2%	0.084	-11.3%	0.072	3.5%	0.026	2.3%	0.036	3.4%	0.053
Divorced or widowed, female		4.8%	0.071	-21.5%***	0.054	6.5%***	0.018	4.1%*	0.025	6.1%*	0.037
Labour status (ref. cat: employed / student)	Economically inactive	-23.3%***	0.044	-10.6%*	0.057	2.5%	0.027	12.7%***	0.027	18.7%***	0.040
	Unemployed	-2.3%	0.073	-8.9%	0.069	3.1%	0.022	3.3%	0.030	4.8%	0.045
Ethnic non-Latvian		-2.2%	0.036	4.6%	0.039	-1.3%	0.012	-0.5%	0.012	-0.7%	0.017
Education (ref. cat: higher / incomplete higher)	Below secondary	8.1%	0.057	-19.4%***	0.049	5.4%***	0.017	2.4%	0.021	3.5%	0.030
	Secondary / vocational secondary	6.8%*	0.038	-2.5%	0.045	-0.2%	0.015	-1.7%	0.015	-2.5%	0.022

Notes: Asterisks *, **, *** indicate a statistically significant difference from the base group at 10%, 5%, 1% level respectively.
Other factors controlled: Average income per household member

The Economic Research Guardian – Vol. 2(2)2012
Semi-annual Online Journal, www.ecrg.ro
ISSN: 2247-8531, ISSN-L: 2247-8531
Econ Res Guard 2(2): 160-179

Table A - Association between socioeconomic factors and self-assessed health in Latvia – comparison of results of two-dimensional stereotype logit model and ordered probit model

	Number of observations	Wald chi2(50)	Log pseudolikelihood	AIC	BIC	Lacy R2O
Two-dimensional stereotype logit model (slogit)	912	114.95	-1228.01	2542.02	2749.10	22.85%
Ordered probit model (oprobit)	912	248.45	-1263.92	2573.83	2684.59	20.68%

Factors			Association between each factor and health (comparison with the reference category, impact of other factors is excluded)									
		Mean probabilities	29%		32%		15%		10%		14%	
			Very good		Good		Fair		Poor		Very poor	
			Never ails/ails very rarely		Has had only minor sicknesses		Has had serious sicknesses that are cured		Has had serious sicknesses and still suffers from them		Has chronic diseases/ is disabled	
			dP/dX	S.E.	dP/dX	S.E.	dP/dX	S.E.	dP/dX	S.E.	dP/dX	S.E.
Female		slogit	1.6%	0.050	-1.8%	0.052	0.4%	0.016	-0.1%	0.016	-0.1%	0.024
		oprobit	0.7%	0.038	0.1%	0.008	-0.2%	0.012	-0.2%	0.013	-0.4%	0.020
Lives in Riga or Riga district (ref. cat.: lives outside Riga district)		slogit	0.5%	0.050	4.6%	0.059	-1.7%	0.018	-1.4%	0.017	-2.1%	0.025
		oprobit	2.9%	0.038	0.5%	0.006	-1.0%	0.013	-1.0%	0.013	-1.5%	0.019
Lives in Riga or Riga district, male		slogit	19.6%**	0.083	-17.2%***	0.059	2.1%	0.030	-1.8%	0.024	-2.7%	0.035
		oprobit	12.0%*	0.063	0.6%	0.007	-4.1%*	0.022	-3.7%**	0.017	-4.9%**	0.020
Age		slogit	7.8%**	0.033	-5.0%	0.033	0.5%	0.012	-1.3%	0.011	-2.0%	0.017
		oprobit	5.4%**	0.025	1.1%**	0.005	-1.8%**	0.008	-1.9%**	0.009	-2.8%**	0.013
$Age^2/100$		slogit	-21.47%***	0.001	10.7%	0.001	-0.4%	0.000	4.5%*	0.000	6.6%*	0.000
		oprobit	-14.9%***	0.001	-3.0%**	0.000	4.9%**	0.000	5.2%**	0.000	7.8%***	0.000
$Age^3/1000$		slogit	1.6%***	0.000	-0.8%	0.000	0.0%	0.000	-0.4%*	0.000	-0.5%*	0.000
		oprobit	1.1%***	0.000	0.2%**	0.000	-0.4%**	0.000	-0.4%**	0.000	-0.6%***	0.000
Single (ref. cat: married or lives with a partner)		slogit	-3.9%	0.061	3.4%	0.074	-0.6%	0.024	0.4%	0.027	0.7%	0.039
		oprobit	-2.0%	0.049	-0.4%	0.012	0.6%	0.016	0.7%	0.017	1.1%	0.027
Single, female		slogit	2.2%	0.084	-11.3%	0.072	3.5%	0.026	2.3%	0.036	3.4%	0.053
		oprobit	-1.6%	0.056	-0.4%	0.014	0.5%	0.018	0.6%	0.020	0.9%	0.032
Divorced or widowed, female		slogit	4.8%	0.071	-21.5%***	0.054	6.5%***	0.018	4.1%*	0.025	6.1%*	0.037
		oprobit	-5.9%	0.042	-1.7%	0.017	1.8%	0.013	2.2%	0.017	3.6%	0.030
Labour status (ref. cat: employed / student)	Economically inactive	slogit	-23.3%***	0.044	-10.6%*	0.057	2.5%	0.027	12.7%***	0.027	18.7%***	0.040
		oprobit	-21.5%***	0.032	-9.8%***	0.026	5.3%***	0.008	8.3%***	0.016	17.7%***	0.040
	Unemployed	slogit	-2.3%	0.073	-8.9%	0.069	3.1%	0.022	3.3%	0.030	4.8%	0.045
		oprobit	-5.9%	0.052	-1.8%	0.022	1.8%	0.015	2.2%	0.020	3.7%	0.038
Ethnic non-Latvian		slogit	-2.2%	0.036	4.6%	0.039	-1.3%	0.012	-0.5%	0.012	-0.7%	0.017
		oprobit	-0.2%	0.027	0.0%	0.005	0.1%	0.009	0.1%	0.009	0.1%	0.014
Education (ref. cat: higher / incomplete higher)	Below secondary	slogit	8.1%	0.057	-19.4%***	0.049	5.4%***	0.017	2.4%	0.021	3.5%	0.030
		oprobit	-1.6%	0.041	-0.3%	0.010	0.5%	0.013	0.6%	0.014	0.9%	0.023
	Secondary / vocational secondary	slogit	6.8%*	0.038	-2.5%	0.045	-0.2%	0.015	-1.7%	0.015	-2.5%	0.022
		oprobit	5.0%*	0.030	1.0%	0.007	-1.6%*	0.010	-1.7%*	0.010	-2.6%	0.016

Notes: Asterisks *, **, *** indicate a statistically significant difference from the base group at 10%, 5%, 1% level respectively.
Other factors controlled: Average income per household member.