

# Proxy responses to subjective questions: the influence on the results of the health expectancy indicator

Darja Lavtar<sup>1</sup>, Metka Zaletel<sup>2</sup>, Ajda Rogelj<sup>3</sup>

<sup>1</sup> *National Institute of Public Health, Ljubljana, Slovenia; darja.lavtar@nijz.si*

<sup>2</sup> *National Institute of Public Health, Ljubljana, Slovenia; metka.zaletel@nijz.si*

<sup>3</sup> *National Institute of Public Health, Ljubljana, Slovenia; ajda.rogelj@nijz.si*

## Abstract

Healthy life years (HLY) is a health expectancy indicator that is on yearly basis calculated from life tables (mortality component) and prevalence of limitation because of health problems in the activities people usually do (disability component). The disability prevalence data are provided by the GALI (Global Activity Limitation Instrument) question from Statistics on Income and Living Conditions (EU-SILC). The HLY indicator is calculated using the Sullivan's method.

Comparability is mainly pre-determined by the comparability of data sources. For the GALI question, the use of proxy respondents should be limited as much as possible. Even though proxy respondents are beneficial in case of sample units that are unable to answer on their own behalf (especially old and disabled interviewees), there is however a question of quality of proxy response. Studies report fairly good agreement between proxy respondent and target respondent in assessments of functioning, physical health, and cognitive status, but this agreement depends on many circumstances, where survey mode and socio-demographic characteristics of respondents should not be neglected. The paper will address the differences of proxy and non-proxy respondents (self-response) and try to identify the characteristics of respondents for whom most proxy responses are given.

The case study will be presented: from the Slovenian EU-SILC survey data from 2008 until 2014 GALI question will be analysed in the view of different survey modes used and proxy/self-response prevalences. Estimates of HLY based only on self-reported limitation, prevalences do not differ from those taking into account proxy and non-proxy responses. There are about one fifth proxy responses to the questions on health. Women tend to be proxies more often than men are. Male proxy partners seem to under-report to higher extent the limitations of their female partners than the opposite. Conclusions will show the robustness of the results to the under-reported prevalences.

**Keywords:** proxy respondents, subjective questions, health expectancies.

## **1. Introduction**

Healthy Life Years (HLY) is a structural indicator that is on yearly basis calculated from life tables and prevalence of limitation because of health problems in the activities people usually do. The source of data is Statistics on Income and Living Conditions (SILC). The standard wording of the GALI question is: *For at least the past 6 months, to what extent have you been limited because of a health problem in activities people usually do? Would you say you have been... 1. Severely limited 2. Limited but not severely, or 3. Not limited at all?*

## **2. Background**

In literature, there is a methodological assumption that the best method of collecting information about a target individual is from that individual directly (Mathiowetz and Groves 1985). Hendershot (2003), however, stated that a review of the literature on proxy reporting found little evidence that self-reports were inherently superior to data provided by a proxy (Moore 1988). Both Hendershot (2003) and Moore (1988) warn that well designed studies of the self/proxy issue are rare, and the range of topics limited. Thus, in this report we will try to understand the effect of proxy respondents (versus self-response) on GALI question. In short, we will address the differences of proxy and non-proxy respondents and try to identify if proxy responses (in comparison to self-report) on GALI question can be interpreted as bias.

Based on the agreement between the National Institute of Public Health (NIJZ) and Statistical Office of the Republic of Slovenia (SURS), the NIJZ yearly receives microdata from Slovenian SILC (health variables, some socio-demographic variables and some methodological variables). We will analyse the SILC data for years 2008-2014, with a focus on a year 2011 (this year was chosen due to some practical reasons).

## **3. Use of proxy respondents**

Mathiowetz (2010) described the use of proxy respondents in a following way: “Proxy reports are used in surveys as an alternative to interviewing all individual sampled unit members, when attempting to obtain individual reports is neither reasonable nor practical”. Thus, proxies play a critical role as sources of health information not only for older persons with cognitive

impairment and other chronic debilitating conditions (Neumann et al. 2010), but also other persons that are not available at the given moment. Even though proxy respondents are absolutely beneficial in case of sample units that are unable to answer on their own behalf (especially old and disabled units), there is however a question of the quality of proxy response. Proxy responses can indeed in many ways not be as accurate as the response obtained from the non-proxy respondent or primary target.

There are few assumptions present in survey design community about the nature of proxy and non-proxy responses. As Swamy (1994) writes, the first assumption is based on the notion that proxy respondents lack knowledge and that the scope of shortage is related to the relationship between proxy respondents and target respondents. Another assumption is based on the idea that health related activities create a role within the household, which one household member must fill and thus this proxy respondent is better suited to accurately report health events than the target respondent (for instance, mothers possess accurate knowledge about children's health). This assumption can be attributed to the reporting of "embarrassing" or stigmatizing health conditions; for instance Berk et al. (1986) stated that in respect to reporting physical stigmatizing conditions, proxies are preferable to self-respondents. Last assumption is based on the idea that there is a good communication between family members and that proxy respondents can indeed provide accurate information about target respondents' physical and emotional well-being. According to this assumption, there is no difference between proxy and non-proxy responses regarding their answers' quality and validity (Swamy 1994).

Proxy response validity is related to the type of the question asked. Many so called objective questions (for instance, number of cars in household) can be obtained from proxy respondents with the same accuracy as from non-proxy respondents. However, subjective, sensitive and personal questions can hardly be accurately obtained from proxy respondents. As GALI question can be understood as more or less subjective and personal question, we wondered if the proxy responses are in accordance with self-report (non-proxy) responses.

Topics that we applied to investigate the impact of proxy respondents were following:

- What kind of proxy respondents (their socio-demographic characteristics) under-report or over-report activity limitation prevalence?
- Does the mode of interview matter in reporting activity limitation prevalence?

Neumann et al. (2000) wrote that in general, studies report fairly good agreement between proxy respondent and target respondent in assessments of functioning, physical health, and cognitive status, and fair-to-poor agreement in assessments of psychological well-being. Hendershot (2003), however, stated that there is evidence that proxies do respond to disability-related questions differently from self-respondents. Hendershot indicated that Todorov and Kirchner (2000) found out that proxy respondents, compared to self-respondents, were less likely to report an activity limitation for sample persons of working age. According to Hendershot (2003, p.9), this statements can also be recognized in the earlier literature cited by Todorov: Bassett et al.1990; Epstein et al. 1989; Kovar and Wright 1973; Mathiowetz and Groves 1985; and Rothman et al. 1991. Based on Slovenian data, we can also state that there are some differences between proxy and non-proxy respondents regarding GALI results and that it is evident that proxy respondents under-report activity limitations.

Hendershot (2003) continues with the following finding: “If proxy reporting is more prevalent for sample persons with severe disabilities, and if proxies tend to underreport disability, survey response patterns may be biased toward underreporting of severe disability”.

#### **4. Mode of interview**

In Slovenian SILC, CAPI interviews are done in the first interview – that is in the first wave, when the person is selected in the sample. Consecutive interviews in the following years are mostly done by CATI (if the telephone number of the respondent is available).

#### **5. Limitations of analysis**

One of the limitations of our research is that we cannot investigate some of the proxy characteristics that would help us clarify the difference between proxy and non-proxy respondents on GALI question. One of them is the effect of family relationship between proxy

respondent and target respondent, as we cannot identify their relationship from the existing database<sup>1</sup>.

There is also a limitation regarding mode of data collection, as we can only investigate the effect of data collection mode for non-proxy respondent (CAPI vs. CATI), while we cannot identify what was the mode used for proxy respondents<sup>2</sup>.

Another characteristics, that are really demanding to recognize from the existing database are the income of the proxy respondent (e.g. income based on OECD equivalence scale) and age difference between proxy and target respondent (person selected in the sample).

Some analyses were done by the SURS (as NIJZ did not receive all the necessary variables for such methodological analysis at the time of the analysis) so that we have more information on the proxy respondents.

Still, the database is organized in a way that we have the information about the person selected in the sample. The information about other members of the household (proxy person is one of them) is limited. That means, that in the case a proxy provides the answers, we have the information about the selected person, not about the proxy. In order to get the characteristics of the proxy, very advanced analysis should be done.

## **6. Results**

Results will be presented in the form of figures and tables with short comments.

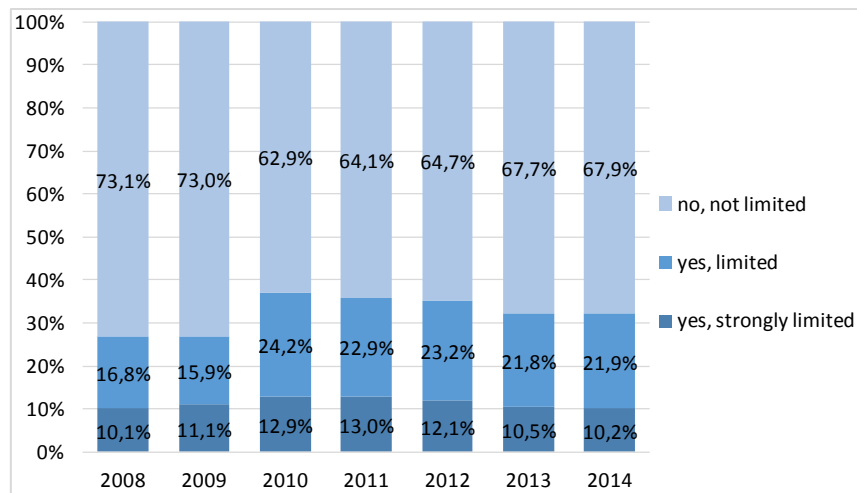
---

<sup>1</sup> The SURS database includes that information, so for the purpose of preparing this paper we asked them do prepare some tables. We have already started arrangements to add this variable to the agreement, so we can receive it in the future.

<sup>2</sup> The SURS database includes that information, so for the purpose of preparing this paper we asked them do prepare some tables. We have already started arrangements to add this variable to the agreement, so we can receive it in the future.

### 6.1. Limitation over the years

Until 2009, the prevalence of limitation was around 27% (limited and severely limited). From 2010 on, the prevalence was always over 30% with its peak in 2010. The GALI question was slightly changed in 2010 in order to be harmonized with the GALI question in European Health Interview Survey (EHIS). Until 2009, the wording of the question was referring to a limitation that lasted for a long time in the past 6 months (“Has ‘selected person’ been in the last 6 months limited for longer period of time in usual activities because of health problems?”). In 2010, the wording of the question changed and is now referring to the limitation that has lasted for 6 months or more. There was a break in series in 2010, but from 2010 on the prevalence of limitation is slowly lowering. From 2010 on, the GALI question remained the same, also its position in the questionnaire. However, in the recent years more attention has been given to GALI question at interviewers’ trainings.



**Figure 1:** Level of limitations, unweighted data (SILC 2011)

### 6.2. Socio-demographics and limitations (SILC 2011)

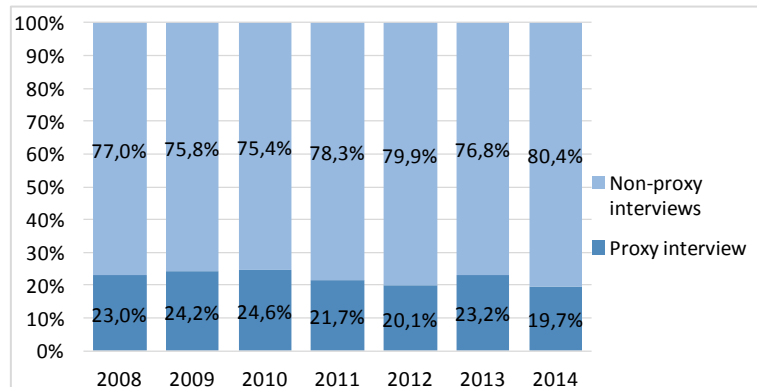
The share of reported limitations in activities people usually do is higher among women, older age groups, people with below upper-secondary education, retired, people who are in risk of poverty, people who answered the survey questions by themselves (either CAPI or CATI). Proxy respondents report less activity limitations (30.2%) compared to self-reports (37.4% for CAPI and 37.5% for CATI).

**Table 1:** Socio-demographic characteristics of limited and not limited persons (SILC 2011)

Limitation in activities because of health problems		Limited or strongly limited	Not limited
		<i>Row N %</i>	<i>Row N %</i>
Total		35.9	64.1
<b>Sex</b> <i>Chi-square=34.618, sig=0.0</i>	Male	33.1	66.9
	Female	38.6	61.4
<b>Age</b> (10-year age groups)  <i>Chi-square=1337.006, sig=0,0</i>	16-24	16.4	83.6
	25-34	18.6	81.4
	35-44	24.3	75.7
	45-54	36.9	63.1
	55-64	47.9	52.1
	65-74	58.1	41.9
	75+	72.1	27.9
<b>Highest ISCED level attained</b> (3 groups) <i>Chi-square=327.346, sig=0.0</i>	Below upper-secondary education	50.6	49.4
	Upper-secondary education	33.3	66.7
	Tertiary education	26.7	73.3
<b>Self-defined economic status</b> (4 groups)  <i>Chi-square=1117.655, sig=0.0</i>	Employed	23.3	76.7
	Unemployed	39.4	60.6
	Retired	58.9	41.1
	Inactive	27.8	72.2
<b>Risk of poverty</b> <i>Chi-square: NA</i>	In risk of poverty	47.1	52.9
	Not in risk of poverty	34.2	65.8
<b>Type of interview</b>  <i>Chi-square=34.182, sig=0.0</i>	Face-to-face interview (CAPI)	37.4	62.6
	Telephone interview (CATI)	37.5	62.5
	Proxy respondent	30.2	69.8

### 6.2. Proxy interviews over the years

The number of proxy interviews among all interviews was relatively stable across the years between 2008 and 2014. The absolute difference between 2012 when the share of proxy interviews was the lowest (20.1%) and between 2010 when the share of proxy interviews was the highest (24.6%) equals 4.5 percentage points.



**Figure 2:** Proxy in interviews over the years (SILC 2011)

### 6.3. Mode of interview and proxy/non-proxy answers (SILC 2011)

In 2011, among the respondents there were 78.5% self-reports (non-proxy responses) and 21.5% proxy responses<sup>3</sup>. There were two times more CATI responses than CAPI in both categories (proxy/non-proxy responses).

**Table 2:** Mode of interview and proxy/non-proxy answers (SILC 2011)

Non-proxy CAPI	28.2%
Non-proxy CATI	50.3%
Proxy CAPI	7.0%
Proxy CATI	14.5%
Total	100.0%

The reason for the use of proxy was only investigated in 2013. The results show that 91% of proxy responses were due to absence of the persons selected in the sample (5% did not want to take part in the survey, 4% were incapable to participate).

### 6.4. Socio-demographic characteristics of the 'proxy' respondents (SILC 2011)

As already stated in the 'limitations of analysis' part, we have the information about the person selected in the sample but not about the proxy respondent. That means that in all analyses the 'socio-demographic characteristic of proxy' is indeed the socio-demographic characteristic of the person selected in the sample. More proxy answers are given for: men, younger age groups, people who attained upper-secondary education, employed, partners and children.

**Table 3:** Socio-demographic characteristics of the 'proxy' respondents (SILC 2011)

		Column N %			Column N %
<b>Sex</b>	Male	66.1	<b>Highest ISCED level attained</b> (3 groups)	Below upper-secondary education	24.9
	Female	33.9		Upper-secondary education	61.5
<b>Age</b> (10-year age groups)	16-24	27.6		<b>Self defined economic status</b> (4 groups)	Tertiary education
	25-34	21.8	Employed		47.6
	35-44	14.6	Unemployed		9.3
	45-54	14.9	Retired		16.4
	55-64	11.3	<b>Relationship of proxy to the selected person</b>	Inactive	26.7
	65-74	5.6		Father	11.5
	75+	4.2		Mother	33.2
			Partner	41.4	

<sup>3</sup> There is a small difference in the share of non-proxy respondents (78.5%) compared to the above figure (78.3%) – both referring to SILC 2011 due to missing values and different tabulation.



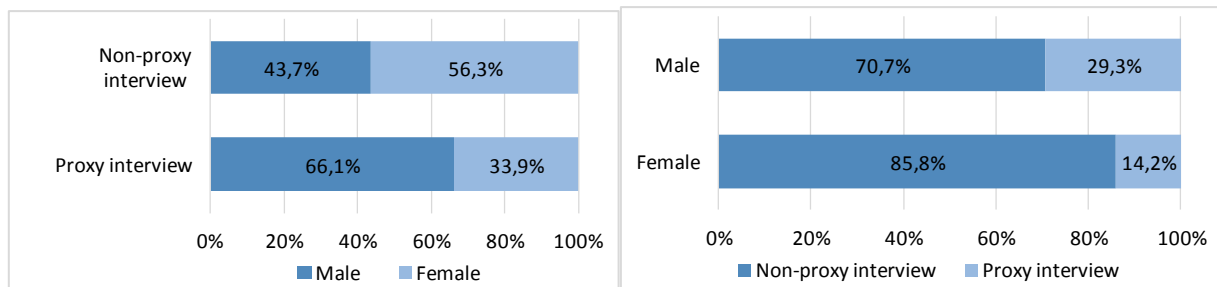
	Other	13.9
--	-------	------

As partners provide answers in 41.4% of all proxies, we took a look in the gender and age structure of persons for whose answers were provided by partner proxies: most proxy reports are for men (75.2%) and age groups over 36 years of age.

### 6.5. Proxy/non-proxy answers according to the gender of selected persons (SILC 2011)

Self-reported (non-proxy) limitations in activities people usually do are reported by 43.7% of men and 56.3% of women. On the other hand, proxy interviews are done for two thirds (66.1%) of men and one third (33.9%) of women (Figure 3, left). There are more proxy interviews for male respondents (29.3%) than for female respondents (14.2%) (Figure 3, right).

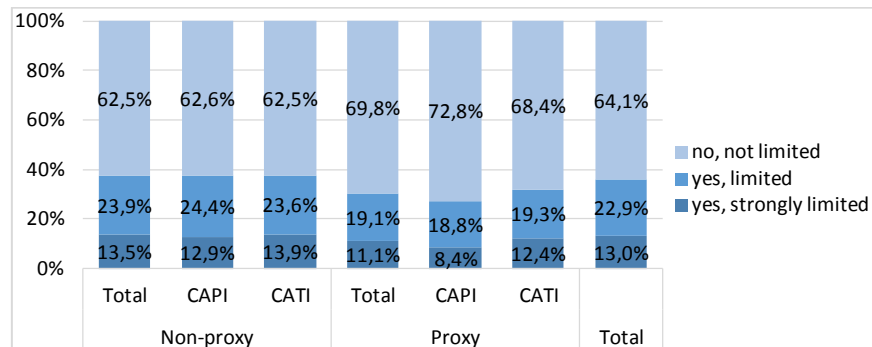
**Figure 3:** Gender of non-proxy and proxy respondents (left) and proxy/non-proxy interviews by gender (right) (SILC 2011)



### 6.6. Activity limitations by mode of interview and by proxy/non-proxy answers

Proxies in general report less limitations than selected persons do. Among proxies, fewer limitations are reported in CAPI interviews than in CATI interviews. Among non-proxy respondents, there are fewer differences between CAPI and CATI responses – slightly more strong limitations are reported in CATI interviews.

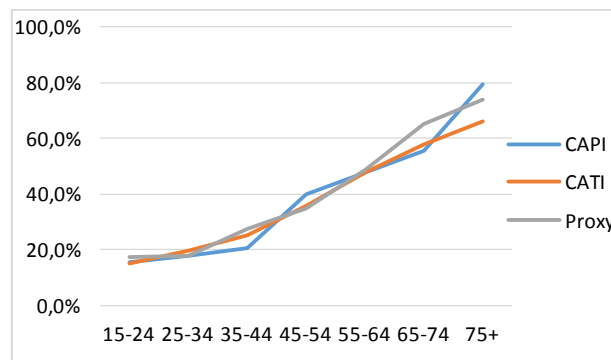
**Figure 4:** Activity limitations by mode of interview and by proxy/non-proxy answers (SILC 2011)



### 6.7. Activity limitations by mode of interview and by age categories and by gender

Reporting of activity limitations by mode of interview is stable in almost all age categories; there are some differences in older age categories.

**Figure 5:** Activity limitations by mode of interview and by age categories (SILC 2011)

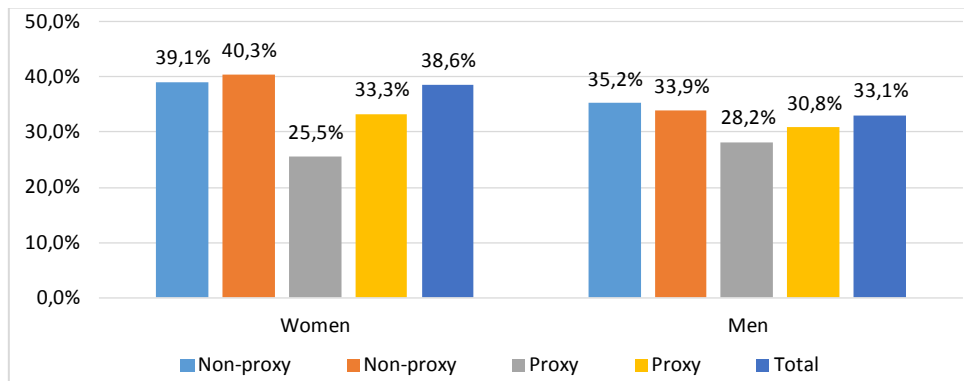


Proxies answering the survey in the name of older women (65+ years) report a larger share of activity limitations, but on the other hand, the proxies answering the survey in the name of middle-aged women (55-64 years) report a smaller share of activity limitation. These inconsistencies might be the consequence of small number of cases in different table cells.

### 6.8. Activity limitations by gender by mode of interview and by proxy/non-proxy answers

Most women self-reported activity limitations in CATI interviews (40.3%), the least limitations because of health problems in activities people usually do were reported in CAPI interviews when proxy was answering the questions (25.5%). The differences among women are bigger than among men where there is quite stable reporting of limitations according to mode of interview and proxy/non-proxy respondents. As we saw that partners are proxies in most cases we could say that female partners produce more accurate limitations than male partners do.

**Figure 6:** Activity limitations by gender by mode of interview and by proxy/non-proxy answers (SILC 2011)



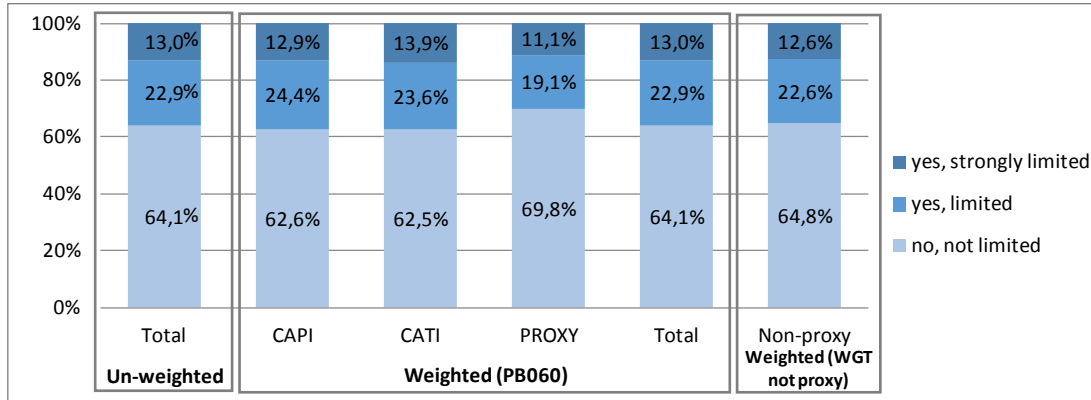
*6.9. If only non-proxy responses counted*

When trying to find an answer to the question if proxy answers count when we investigate activity limitations among adults, we decided to calculate limitations only among self-reports (non-proxy respondents). For this purpose, SURS calculated special weights that only take into account the non-proxy reports. That means that the proxy responses are not considered. The results show that prevalences of limitation does not differ (weighted total vs non-proxy weighted). So, the impact of proxies on total prevalence of activity limitation is not that big because of:

- Not too many proxies
- There are more proxies participating in the survey and answering in the name of men (for whom women answer) and their answers are more consistent with the men's.

That means that, although proxies under-report limitations, their answers do not produce less activity limitations. We can only assume that this happens because the most answers come from self-reporting. Also, there are more proxies for male respondents that do not tend to under-report the limitations.

**Figure 7:** Activity limitations (un-weighted, weighted by PB060 and weighted by special weight for non-proxies) (SILC 2011)



6.10. Healthy life years (HLY) at 0, 50 and 65 years of age (female) by mode of interview and proxy answers

Healthy life years for females were calculated by mode of interview and proxy answers at ages 0, 50 and 65. The results show that HLY at the age of 0 and 50 are quite stable and that the mode of interview is not an important characteristic. Only proxies tend to under-report activity limitations and that can also be seen in the HLY results. Again, if we produce results based only on self-reported limitation prevalences the HLY results are not different from those taking in account proxy and non-proxy responses.

**Table 3:** Healthy Life Years and % Healthy Life Years of Total Life Expectancy (TLE) at the age of 0, 50 and 65 by mode of interview and use of different weights (SILC 2011)

	Age	HLY (DFLE)	%DFLE/TLE
CAPI (weighted)	0	56	66.9
	50	13.7	39.2
	65	6.5	29.9
CATI (weighted)	0	55.4	66.2
	50	14.5	41.5
	65	7.4	34.3
PROXY (weighted)	0	53.4	63.8
	50	12.9	36.9
	65	4.5	20.8
TOTAL (weighted)	0	55.5	66.3
	50	14.2	40.5
	65	6.9	32.0
NOT PROXY (weighted by special weight for non-proxy responses)	0	55.4	66.2
	50	14.2	40.5
	65	7.0	32.4

7. Results

Results from the Slovenian SILC show that there are about one fifth proxy responses to the questions on health. Women tend to be proxies more often than men are. Male proxy partners seem to under-report to higher extent the limitations of their female partners than the opposite. In general, the results are so robust that those under-reported prevalences do not count – the prevalences are similar if we take or do not take proxies into calculations.

## 8. References

- Berk ML, Horgan CM, and Meyers SM. (1986), The Reporting of Stigmatizing Health Conditions: A. Comparison of Proxy and Self-Reporting, *JESM* 14(3):197-205.
- Hendershot GE. (2003), The Effects of Survey Nonresponse and Proxy Response on Measures of Employment for Persons with Disabilities, Available at: <http://dri.uiuc.edu/research/dissemination/Hendershot.pdf> (5.5.2015)
- Mathiowetz N and Groves R. (1985), The Effects of Respondent Rules on Health Survey Reports, *Am J Public Health* (75). 633-639.
- Mathiowetz NA. (2010), Self and Proxy Reporting In the Consumer Expenditure Survey Program, Paper Prepared for the Consumer Expenditure Methods Workshop Bureau of Labor Statistics, Milwaukee, USA.
- Moore JC. (1988), Self/Proxy Response Status and Survey Response Quality: A Review of Literature, *JOS* 4(2): 155-172.
- Swamy M. (1994), Analysis of self and proxy respondents in the assessment of health status, Proceedings of the Survey Research Methods Section, ASA.
- Todorov A and Kirchner C. (2000), Bias in proxies' reports of disability: Data from the National Health Interview Survey on disability, *Am J Public Health* 90(8):1248-1253.