Katherine Keenan, Else Foverskov, Emily Grundy

Data sources on the older population in Europe: comparison of the generations and gender survey (GGS) and the survey of health, ageing and retirement in Europe (SHARE)

Article (Accepted version) (Refereed)

Original citation:
Keenan, Katherine, Foverskov, Else and Grundy, Emily (2015) *Data sources on the older population in Europe: comparison of the generations and gender survey (GGS) and the survey of health, ageing and retirement in Europe (SHARE).* Population: Édition Française. ISSN 0032-4663

Title: Data sources on the older population in Europe: Comparison of the Generations and Gender Survey (GGS) and the Survey of Health, Ageing and Retirement in Europe (SHARE).

Keywords: Generations and Gender Survey (GGS); Survey of Health, Ageing and Retirement in Europe (SHARE); data quality; health; cross-survey comparison

Authors:

Dr. Katherine Keenan

Ms. Else Foverskov

Professor Emily Grundy

All affiliated to:

Department of Social Policy,
London School of Economics,
Houghton Street, London,
WC2A 2AE.

Abstract

The Survey of Health, Ageing and Retirement in Europe (SHARE) and the Generations and Gender Survey (GGS) are two widely used European longitudinal surveys with data on socio-demographic and health topics, but their comparability has not been systematically investigated. We compared SHARE and GGS data for 50-80 year olds in seven European countries (Belgium, Estonia, France, Germany, Hungary, the Netherlands and Poland) to assess data quality and the potential for joint analyses. The results showed that information on, and distributions by, age, gender, marriage and fertility patterns were broadly similar in both sources. For some countries distributions by educational level varied between the two sources even though both reported using the International Standard Classification of Education, which may reflect variations in the timings of surveys. The wording of health questions and their placement in the questionnaire sometimes differed between the surveys. This may account to some extent for differences between them in estimates of the prevalence of poor health. We investigated what effect these variations might have on analyses of health inequalities by undertaking multivariable analysis of associations between education and marital status and two health indicators.

Introduction

Understanding age associated changes in socio-demographic circumstances, health, resources and activity patterns is a key priority in Europe given substantial past and projected future increases in the representation of older people in the population (United Nations 2013). High quality, representative longitudinal data are required as a basis for developing this understanding. To this end considerable resources have been devoted to establishing comparable large scale cross-national longitudinal data sets notably the Survey of Health, Ageing and Retirement in Europe (SHARE) (Börsch-Supan et al. 2013), and the Generations and Gender Survey (GGS) (Vikat et al. 2007), both of which are freely available and widely used. While they both collect data on older people, the surveys have different objectives. The GGS was primarily developed to underpin the study of family and intergenerational processes from young adulthood to old age (Vikat et al. 2007), while the focus of SHARE is on ageing (Börsch-Supan et al. 2013). Both studies were initiated in the early 2000s and while the quality of the data collected in each has been investigated through comparison with other sources, including national population data and European population surveys such as European Union Statistics on Income and Living Conditions (EU-SILC) (Fokkema, Kveder, and Liefbroer 2014, Börsch-Supan et al. 2005, Croezen, Burdorf, and Lenthe 2013, Vergauwen et al. 2015), they have not to date been systematically compared with each other. We compare data from SHARE and GGS where they cover the same countries and age ranges.

The aims of this paper are to present sources of differences in the SHARE and GGS surveys and to investigate whether common health measures drawn from the two surveys provide comparable information, either in terms of prevalence or their patterning by socio-demographic characteristics. We start by examining the GGS and SHARE survey methodologies,

then compare data on common indicators including age, gender, education, fertility, marriage and health. We go on to assess the comparability of results from multivariate regression modelling of associations between two socio-demographic indicators, education and marital status, with measures of health.

Methodology

Comparison of SHARE and GGS surveys

Survey design

Baseline SHARE and GGS surveys have been carried out in ten common countries: Austria, Belgium, Estonia, France, Germany, Hungary, Italy, the Netherlands, Poland and Sweden. We compared measures from seven of these countries for the population aged 50-80 years. We excluded Italy and Austria because neither included respondents aged 65 and over. At the time of writing the Swedish GGS data was not available so was initially excluded. We also note that the Swedish GGS and SHARE surveys were conducted more years apart than in any other country[1] which might threaten their comparability.

Table 1 shows the main characteristics of the sampling procedures and fieldwork for the surveys included in this study: GGS surveys from wave 1, and SHARE surveys from the baseline wave (Estonia, Hungary and Poland joined the SHARE survey at wave 2 or later). All countries except Estonia and the Netherlands (GGS) used multi-stage sampling strategies with the most common sampling frames being the census or population register. In most countries the

[1] The Swedish SHARE wave 1 was conducted in 2004; the Swedish GGS wave 1 was conducted in 2012-13.

Table 1: Main characteristics of sampling and fieldwork for GGS Wave 1 and SHARE baseline data (common countries except Austria, Italy and Sweden)

Country	Survey	Fieldwork period	Number sample stages	Sampling method[a]	Frame	Geographic coverage	Sampling unit	Institutionalised people included	Age range sampled	Overall response rate
Belgium	**GGS W1**	Feb 08- May 10[b]	2	PPS + SRS	Population register	Full coverage	Individuals	No	18-79	43.8[c]
	SHARE W1	Jan 05-Jul 05	2/3	PPS+SRS	Telephone number list	Not German-speaking areas (<1% population)	Individuals	No	50+[d]	39.2
Estonia	**GGS W1**	Sep 04- Dec 05	1	SRS	Census and update new dwellings	Full coverage	Individuals	Yes	21-80	70.2[c]
	SHARE W4	Feb-Oct 11	1	SRS	Population register	Full coverage	Individuals	No	50+[d]	61.0
France	**GGS W1**	Sep 05 – Dec 05	2	PPS and SRS	Census and update new dwellings	Full coverage	Dwellings/ addresses	No	18-79	71.7[c]
	SHARE W1	Oct-Nov 04	2 (3 in urban)	PPS + SysRS	Census and update new dwellings	6 regions (covering approx.50% of population)[e]	Dwellings/ addresses	No	50+[d]	73.6
Germany	**GGS W1**	Feb 05 – May 05	2	PPS and SRS	List of households (ADM-Master-Sample)	Full coverage	Addresses	No	18-79	55.4[c]
	SHARE W1	May 04 – Oct 04	2	SysRS + SRS	Population register	Full coverage	Individuals	Yes	50+[d]	63.4
Hungary	**GGS W1**[f]	Oct 04 – May 05	2	PPS+ SRS	Population register	Full coverage	Individuals	na	21-78	83.2[c]
	SHARE W4	Mar-Oct 11	2	PPS+SRS	Population register	Full coverage	Individuals	Yes	50+[d]	63.0
Netherlands	**GGS W1**	Sep 02 – Mar 04	1	SRS	Address list from mail company	Full coverage	Addresses	No	18-79	44.7[c]
	SHARE W1	May-Oct 04	2	SysRS + SRS	Population register	Full coverage	Households	Yes	50+[d]	61.6
Poland	**GGS W1**	Nov 10 – Feb 11	2	SRS+SRS	Address list	Full coverage	Addresses	No	18-79	na
	SHARE W2	Nov 06-Aug 07	na	na	na	na	na	na	50+[d]	na

na: information is currently unavailable.[a] PPS = Probability Proportional to Size, SRS = Simple random sampling, RR = random route, SysR = Systematic sampling with a random start. [b] Statistics Belgium conducted fieldwork from Feb 08-Apr 09, and TNS Dimarso from Nov 09-May 10. [c] 'Average rate' calculated by Fokkema, Kveder, and Liefbroer (2014). [d] and their partners/spouses of any age. [e] Aquitaine, Île-de-France, Languedoc-Roussillon, Nord-Pas-de-Calais, Pays de la Loire, Rhône-Alpes. [f] Refers to the sampling characteristic of wave 1 of "Turning points of the Life-Course program".
Sources : Börsch-Supan et al. (2013), Börsch-Supan and Jürges (2005), Malter and Börsch-Supan (2013), Fokkema, Kveder, and Liefbroer (2014), Kapitány (2003), GGS webpage http : //www.ggp-i.org/data/data-documentation.html and SHARE webpage : http://www.share-project.org/data-access-documentation/sample.html

sampling frame was designed to provide coverage of the population living in private households, although some countries and surveys also included those living in institutions[2]. Reflecting the different objectives of the surveys a major difference between SHARE and GGS is the target population. In SHARE it was defined as all households with at least one member aged 50 years or older, and within these households, all individuals aged 50 years or older (respondents' spouses/partners were also included, regardless of age) (Börsch-Supan and Jürges 2005). By contrast, in the GGS the target population was individuals aged between 18-79[3] years and only one individual from each household was interviewed. Response rates for both the GGS and SHARE surveys were lowest in Belgium (43.8% and 39.2% respectively). The Netherlands GGS survey also had a rather low response rate of 44.7%. However in most other countries response rates for both surveys were over 60%.

Analysis sample

We selected men and women aged 50-80 years at the time of survey based on their reported dates of birth (50-79 in the Hungarian GGS data)[3]. Partners outside the SHARE age range (less than 50 years) were excluded from the analyses because they are not a representative sample. This resulted in samples of between 2,255-10,447 for each of the surveys and countries included (Table 2). The proportion of respondents with missing values on variables considered ranged between 0.1% and 4.1%. The different timings of the GGS and SHARE surveys, especially for Belgium, Estonia, Hungary, and Poland, mean that the

[2] The Estonian GGS, and the German, Hungarian, Netherlands SHARE stated that institutionalised people were included in the sampling frame. The proportion of institutionalised individuals aged 50-80 years in our analysis was negligible: in the Estonia GGS there were 7 people (<0.001%), and in SHARE they were not identifiable.

[3] In the GGS countries which sampled to age 79 the data includes some people aged 80, who had a birthday between sampling and fieldwork. The exceptions to this were GGS Estonia which deliberately sampled 80 year olds and GGS Hungary which sampled only to age 78.

surveys include survivors of different birth cohorts (see Table 2), who could be subject to different time trends in some variables. To try to assess whether these different timings affected comparability we made additional comparisons using a subset including only the equivalent birth cohorts. However we acknowledge that this means we are comparing cohorts who have survived to different ages.

Table 2: Timing of surveys, sample size and birth cohorts

	GGS			**SHARE**			
	Survey year	Birth cohorts	Sample 50-80 years	Survey year	Birth cohorts	Sample 50-80 years	Equivalent birth cohorts
Belgium	2008-10	1928-60	3,151	2005	1924-55	3,383	1928-55
Estonia	2004-05	1924-56	3,696	2011	1930-61	6,087	1930-56
France	2005	1926-56	4,542	2004	1923-55	2,742	1926-55
Germany	2005	1925-55	4,373	2004	1923-54	2,758	1925-54
Hungary	2004-05	1926-55	6,251	2011	1931-61	2,744	1931-55
Netherlands	2002-04	1923-54	3,460	2004	1923-54	2,651	1923-54
Poland	2010-11	1930-61	10,447	2006-07	1926-57	2,255	1930-57

Sources: GGS webpage http: //www.ggp-i.org/data/data-documentation.html and SHARE webpage: http://www.share-project.org/data-access-documentation/sample.html, authors calculations.

Socio-demographic measures

We compared distributions by gender, age, education, marital status, and number of children. Age was categorised into 5-year groups. Both SHARE and GGS used the International Standard Classification of Education (ISCED) 1997 framework[4] to classify educational level and we distinguish between a low level of education (ISCED 0-2, including no education, primary and lower secondary); versus medium and higher (ISCED 3-6, including upper secondary, and non-tertiary post-secondary, and tertiary education)

We used a 'de jure' measure of marital status which was grouped into four categories: never-married, married, divorced, and widowed, and for the main summary and regression analyses dichotomised into married versus non-married. In the GGS marital

[4] In GGS many questionnaires used their own country-specific classifications and these were post-coded into ISCED.

status was commonly derived from answers to a series of questions on marriage and partnership history, and the resulting variable had the four categories as above. In SHARE we used a single question which asked the respondent if they were 1) married and living together with their spouse, 2) in a registered partnership, 3) married and living separated from their spouse, 4) never-married, 5) divorced or 6) widowed. We grouped together the first three of these categories in SHARE as 'married'. In SHARE or GGS if respondents reported (in other questions) that they were unmarried but living with their partners they were coded according to their reported marital status (never married, divorced or widowed).

As a summary indicator of fertility, we grouped number of children into five categories: 0, 1, 2, 3 and 4+, and for the main summary dichotomised it into childless versus non-childless. Number of children in SHARE was measured by asking "How many children do you have that are still alive? Please count all natural children, fostered, adopted and stepchildren, including those of your husband/your wife/your partner". The SHARE wave 1 questionnaire also asked specifically whether children were biological or non-biological, but only for the first four children reported. A similar count of number of biological, adopted, foster and stepchildren still alive was obtained in GGS by combining information from the household roster and questions on non-resident children. However, lack of information on non-resident stepchildren in the Polish and Estonian surveys and on deaths of stepchildren in the French survey may result in slight estimation differences for those countries.

Health measures

We concentrate on two health measures: self-rated health (SRH) and presence of long-standing illness or chronic conditions. Differences in question wording mean that SRH was directly comparable across both surveys only in Belgium, France, Germany, and the Netherlands. GGS wave 1 used the European variant of SRH (Robine, Jagger, and Romieu 2002) which has an ordered scale of 1-5 ranging from 'very good' to 'very poor'. The SHARE baseline surveys in Belgium, France, Germany, and the Netherlands used *both* the European variant of SRH and the US variant, which has an ordered scale of 1-5 ranging from 'excellent' to 'poor'. The SHARE surveys from Estonia, Hungary and Poland collected SRH using *only* the US version. For the European variant, we dichotomised responses into 'good' (those reporting very good or good health) or 'poor' (fair, bad, or very bad health). For the US scale, respondents reporting excellent, very good or good health were considered to have good SRH while respondents reporting fair or poor health were considered to have poor SRH. Previous studies have shown that the US and European versions of the SRH question are not exactly comparable (Jürges, Avendano, and Mackenbach 2008), and that responses to the US scale usually produce estimates of better health than responses to the European version. Where available, we show distributions of SRH using both European and US variants, which allowed us to see the effect of wording differences.

Long-standing illness (LSI) was measured in the GGS by asking "Do you have any long-standing illness or chronic conditions?" and in SHARE by asking "Do you have any long-term health problems, illness, disability or infirmity?" with response options yes or no. In addition the Estonian GGS specified an illness "lasting 3 months or more".

Analytical methods and post-stratification weights

First, to assess representativeness we compared the weighted age and gender distributions in the GGS and SHARE to each other and to national population data as reported in EUROSTAT for the year the survey took place[5]. Next we undertook descriptive analyses to compare the weighted proportions with particular characteristics in SHARE and GGS by country and gender. Comparisons of the non-dichotomised versions of education, marital status, and number of children are presented in supplementary Figures S1-S3. We used the post-stratification weights provided with the GGS and SHARE surveys for the comparisons. In the GGS the weighting factors vary between countries, and for some countries such as Hungary no information was available (Fokkema, Kveder, and Liefbroer 2014). We provide summary information on GGS weighting factors in Appendix Table S1. For most countries the weights aimed to adjust the sample so that it was nationally representative on important aspects such as age, gender, urbanisation, region and household size. The GGS weights for Belgium, Estonia and Sweden did not take account of household size, so in those countries we recalculated the weight to include it, using population data on household size distributions from the year of sampling downloaded from EUROSTAT (EUROSTAT 2015). This adjustment is important because only one person per household was interviewed in the GSS (resulting in higher sampling probabilities for people living alone) while in SHARE all individuals 50+ were sampled. The Poland GGS did not include any weights, so we weighted the data for household size only (but not for other factors such as gender or age). For the SHARE surveys we used the 'calibrated cross-sectional weights' which adjust for unequal

[5] Population data from EUROSTAT reports age distributions on January 1st, whereas in the survey data we calculated age at the time of survey using dates of birth.

sampling probabilities[6], and to the known proportions of gender and age in the general population (Mannheim Research Institute for the Economics of Aging (MEA) 2013).

Despite variations in the proportions with particular characteristics, we would expect the socio-demographic gradient in health to be consistent across the surveys. To assess whether this was the case we fitted multivariable logistic regression models for associations between two socio-demographic indicators (education and marital status) and two health indicators (SRH and long-standing illness). We did not use weights for the multivariable models, but adjusted them for age (continuous measure) and marital status/education respectively. To test whether the multivariable estimates were significantly different in SHARE and the GGS, we combined the datasets, and in pooled models introduced an interaction term for data source.

Representativeness

In previous studies, both surveys have been compared separately with national population estimates to assess their representativeness (Fokkema, Kveder, and Liefbroer 2014, Börsch-Supan and Jürges 2005). For the GGS the use of post-stratification weights reduced (but did not eradicate) deviation from whole population sources for age, gender and region, but did not correct biases by marital status or education (Fokkema, Kveder, and Liefbroer 2014). Other studies have found that fertility and marriage rates for older cohorts of the GGS are underestimated when compared with population data even when the sampling weights are applied, although period rates from 1970 broadly approximate those in population data (Vergauwen et al. 2015, Kreyenfeld, Hornung, and Kubisch 2013). In SHARE, use of the post-

[6] We chose the SHARE 'calibrated weights' in preference to the 'design weights' which adjust only for unequal sampling probabilities.

stratification 'calibrated weights' produce estimates that are very similar to the target populations (Börsch-Supan and Jürges 2005). As a further check we compared weighted distributions by age group and gender for the SHARE and GGS samples with distributions from national population data for the respective survey years (Figure 1). Overall the SHARE and GGS samples had very similar age structures, with the exception of Hungary which in SHARE had a markedly higher proportion of men and women aged 55-59 years than the GGS. This is partly due to a 5-6 year difference in the timing of the surveys which meant that the later survey included a larger proportion of people born during the post-war 'baby boom'. The largest disparities between national and survey data were seen for women in the Belgian and German GGS samples, men in the French and Polish GGS samples and the Hungarian SHARE samples. The weights in the Polish GGS did not adjust for age so it was not surprising to see more differences. It was common in the GGS (Estonia, France, and Poland) for weighted survey estimates to underrepresent men and women aged 50-54 years, which could be related to the fact that a different age categorisation was used to construct the GGS weights (45-64 years) whereas in SHARE the weights adjust from age 50. Figure 2 shows the weighted proportion of females in SHARE and GGS compared with national population data. Unsurprisingly the largest disparity was in the Poland GGS where the weights did not adjust for gender. The survey gender distributions were most similar for Germany and Hungary, and in no country was the difference larger than 2 percentage points.

Results

Distributions by socio-demographic characteristics and health

Tables 3 and 4 show the weighted (but unadjusted) distributions of socio-demographic and health indicators in SHARE and GGS, for women and men respectively. The differences

between SHARE and GGS in the proportion married and childless were relatively small (results were similar using the non-dichotomised measures of marital status and number of children). In Belgium, the GGS sample had proportionally more childless respondents than the SHARE sample. Overall the largest differences between the surveys in the distributions were seen for education and long-standing illness, with similar differentials for men and women.

In Estonia and Hungary, the GGS samples reported higher proportions of people reporting a low educational level, whereas in the Netherlands and Poland, SHARE included higher proportions of respondents reporting a low educational level. In the Polish GGS, 30% of women had low education, compared with 51% in SHARE. The proportion of Polish women with higher education was also twice as high in the GGS as in SHARE (12% vs. 5%) (see supplementary Figure S1). These differences might reflect inadequate weighting in the Polish GGS resulting in different age structures, which suggests the importance of adjustment for basic demographic variables such as age and sex when comparing cross-country estimates. When the analysis was repeated using equivalent birth cohorts instead of age groups to assess if any differences were introduced by the different timing of the surveys, the differences in education completely disappeared for Estonia (where the SHARE and GGS surveys were conducted 5-7 years apart). In Hungary and Poland (where surveys were conducted 6-7 and 4-5 years apart respectively), the differences were attenuated but substantial differences remained.

When using the European variant of SRH in both surveys there were few substantial differences. The US variant of SRH produced a lower prevalence of poor health than when using the European variant, which is consistent with previous studies (Jürges, Avendano,

and Mackenbach 2008). The proportions with poor SRH, however it was measured, were highest in the Eastern countries, and were particularly high in Estonia.

Table 3: Weighted distributions by socio-demographic variables in GGS and SHARE, women aged 50-80 years in common countries.

Characteristic/country	GGS % (95% CI)	SHARE % (95% CI)	Absolute difference in proportions (GGS-SHARE)
Low education			
Belgium	48.0 (45.5-50.5)	52.1 (49.5-54.6)	-4.1
Estonia	33.1 (31.1-35.0)	25.0 (23.5-26.5)	8.1
France	55.9 (53.7-57.9)	56.5 (53.7-59.1)	-0.6
Germany	26.8 (24.6-28.9)	27.3 (24.7-29.8)	-0.5
Hungary	48.3 (46.6-49.9)	39.9 (34.8-45.2)	8.4
Netherlands	58.6 (56.1-61.0)	64.0 (61.1-66.6)	-5.4
Poland	30.4 (29.0-31.7)	51.1 (48.0-54.1)	-20.7
Married			
Belgium	68.1 (65.7-70.4)	67.3 (67.3-69.7)	0.8
Estonia	51.7 (49.6-53.7)	44.7 (42.9-46.3)	7.0
France	62.1 (60.1-64.1))	61.5 (58.8-64.1)	0.6
Germany	60.1 (57.8-62.3)	59.6 (56.6-62.4)	0.5
Hungary	51.6 (49.9-53.2)	50.4 (44.9-55.8)	1.2
Netherlands	68.0 (65.8-70.0)	66.6 (63.6-69.4)	1.4
Poland	64.7 (63.4-66.1)	58.3 (55.0-61.4)	6.4
Childless			
Belgium	21.6 (19.6-23.7)	12.1 (10.0-14.5)	9.5
Estonia	10.2 (9.0-11.4)	10.0 (8.8-11.3)	0.2
France	10.5 (9.3-11.7)	11.5 (9.6-13.6)	-1.0
Germany	17.4 (15.6-19.2)	15.0 (15.0-17.7)	2.4
Hungary	10.6 (9.5-11.6)	8.6 (5.8-12.4)	2.0
Netherlands	11.1 (9.8-12.5)	12.3 (10.0-14.9)	-1.2
Poland	7.2 (6.6-7.8)	6.5 (4.8-8.6)	0.7
Less than good self-rated health (EURO version)			
Belgium	34.1 (31.7-36.5)	31.3 (29.0-33.6)	2.8
Estonia	74.7 (72.8-76.3)	n/a	n/a
France	42.1 (40.0-44.2)	35.5 (33.0-38.1)	6.6
Germany	43.8 (41.5-46.1)	46.3 (43.5-49.1)	-2.5
Hungary	65.3 (63.6-66.8)	n/a	n/a
Netherlands	30.2 (27.9-32.4)	31.3 (28.5-33.9)	-1.1
Poland	68.9 (67.5-70.1)	n/a	n/a
Less than good self-rated health (US version)			
Belgium	n/a	25.6 (23.4-27.7)	n/a
Estonia	n/a	68.7 (67.0-70.3)	n/a
France	n/a	32.0 (29.5-34.5)	n/a
Germany	n/a	38.7 (35.9-41.5)	n/a
Hungary	n/a	62.6 (57.4-67.5)	n/a
Netherlands	n/a	27.2 (24.6-29.8)	n/a
Poland	n/a	61.1 (58.0-64.0)	n/a

Long standing illness[a]			
Belgium	34.6 (32.2-37.0)	46.2 (43.6-48.6)	-11.6
Estonia	42.7 (40.6-44.8)	73.3 (71.7-74.8)	-30.6
France	41.3 (39.2-43.4)	49.2 (46.5-51.9)	-8.0
Germany	31.9 (31.2-35.6)	59.6 (56.9-58.1)	-27.7
Hungary	58.4 (56.9-60.1)	75.1 (70.9-78.8)	-16.7
Netherlands	42.6 (40.1-45.0)	44.1 (41.2-46.8)	-1.5
Poland	57.5 (55.9-58.7)	66.7 (63.7-69.5)	-9.2

Source: SHARE baseline surveys and GGS wave 1. [a] Questions were worded differently in GGS and SHARE.

Table 4: Weighted distributions by socio-demographic variables in GGS and SHARE, men aged 50-80 years in common countries.

Characteristic/country	GGS % (95% CI)	SHARE % (95% CI)	Absolute difference in proportions (GGS-SHARE)
Low education			
Belgium	42.6 (40.1-45.1)	45.5 (42.9-48.1)	-2.9
Estonia	39.9 (37.0-42.7)	29.7 (27.8-31.6)	10.2
France	44.1 (41.7-46.5)	45.8 (42.8-48.7)	-1.7
Germany	9.7 (8.2-11.3)	6.8 (5.4-8.4)	2.9
Hungary	27.5 (25.7-29.3)	16.0 (13.1-19.3)	11.5
Netherlands	38.4 (35.7-41.0)	48.8 (45.8-51.8)	-10.4
Poland	21.1 (19.7-22.4)	34.5 (31.4-37.7)	-13.4
Married			
Belgium	78.2 (76.0-80.1)	82.0 (79.7-83.9)	-3.8
Estonia	74.3 (71.7-76.5)	69.6 (67.4-71.7)	4.7
France	77.8 (75.8-79.5)	80.6 (78.1-82.8)	-2.8
Germany	76.7 (74.6-78.6)	75.6 (72.6-78.2)	1.1
Hungary	78.3 (76.5-80.0)	76.4 (71.3-80.8)	1.9
Netherlands	80.2 (78.2-82.0)	82.4 (79.6-84.9)	-2.2
Poland	84.3 (83.2-85.3)	79.1 (75.8-82.0)	5.2
Childless			
Belgium	24.7 (22.5-26.9)	13.7 (11.7-16.0)	11.0
Estonia	12.2 (10.3-14.2)	13.2 (11.2-15.3)	-1
France	10.9 (9.5-12.3)	13.1 (10.9-15.6)	-2.2
Germany	19.6 (17.7-21.5)	23.6 (20.6-26.9)	-4
Hungary	10.8 (9.4-12.1)	10.0 (7.7-12.8)	0.8
Netherlands	11.1 (9.7-12.7)	14.0 (11.4-16.9)	-2.9
Poland	9.6 (8.7-10.4)	12.8 (10.1-16.0)	-3.2
Less than good self-rated health (EURO version)			
Belgium	30.3 (28.0-32.6)	27.7 (25.4-30.1)	2.6
Estonia	71.6 (68.9-74.1)	n/a	n/a
France	37.6 (35.2-39.9)	36.1 (33.3-38.9)	2.8
Germany	45.0 (42.6- 47.4)	42.2 (39.4-45.1)	-4.7
Hungary	58.8 (56.8-60.7)	n/a	n/a
Netherlands	23.1 (20.8-25.4)	29.3 (26.7-32.1)	-0.3
Poland	62.1 (60.4-63.6)	n/a	n/a
Less than good self-rated health (US version)			
Belgium	n/a	23.2 (21.0-25.5)	n/a
Estonia	n/a	66.3 (64.1-68.2)	n/a
France	n/a	30.2 (27.5-32.9)	n/a
Germany	n/a	36.2 (33.4-39.0)	n/a
Hungary	n/a	56.0 (50.3-61.5)	n/a
Netherlands	n/a	24.9 (22.4-27.6)	n/a

Poland	n/a	57.5 (54.1-60.9)	n/a
Long standing illness [a]			
Belgium	30.6 (28.3-32.9)	42.8 (40.2-45.4)	-12.2
Estonia	37.5 (34.5-40.5)	67.9 (65.8-69.8)	-30.4
France	42.2 (39.7-44.5)	50.4 (47.4-53.3)	-8.2
Germany	35.2 (32.8-37.5)	55.3 (52.3-58.1)	-20.1
Hungary	50.8 (48.0-52.0)	66.7 (61.2-71.7)	-15.9
Netherlands	35.7 (33.1-38.3)	39.5 (36.3-42.4)	-3.8
Poland	48.5 (46.7-50.0)	58.9 (58.9-62.2)	-10.4

Source: SHARE baseline surveys and GGS wave 1. [a] Questions were worded differently in GGS and SHARE.

As expected from the different question wording, the proportions reporting LSI were consistently different between the surveys (higher in SHARE than in GGS). All countries except the Netherlands and France had differences in excess of 10 percentage points, and in Germany and Estonia this approached a difference of 30 percentage points. The proportional differences between the surveys were not correlated with the overall prevalence of reported poor health in each country.

Figures 3 and 4 show results from several multivariable logistic regression models fitted to explore the associations between socio-demographic factors and poor health in the different countries and surveys. Figure 3 shows (unweighted) associations between low education (ISCED score 1-2) and two binary outcomes – reporting poor SRH, and reporting having an LSI, adjusted for age and marital status. All odds ratios in Figure 3 are above one meaning that low education is associated with higher odds of reporting poor SRH or long-standing illness. For some countries (Hungary and Poland particularly), the GGS sample was larger than the SHARE sample, and standard errors accordingly lower, which may affect differences in significance levels. Models fitted to data from each survey pooled across countries showed that odds of poor SRH and long-standing illness associated with low education were slightly larger in the GGS than SHARE although the confidence intervals are overlapping. For SRH the same pattern

of effect was seen for men and women, but there was some gender variation in results from the models for LSI. Some of the country-specific coefficients show substantial differences, most notably in Hungarian men and women where the SHARE analysis showed higher odds of poor health by low education than the GGS results. This cannot be explained by large differences in the prevalence of poor health in general. In very few countries were the odds from both surveys similar (possibly France, the Netherlands and Germany were most consistent). As described in the methods section, we combined the SHARE and GGS data, and tested for significant differences in the associations by survey. In most cases the differences were non-significant, but notable exceptions were SRH and LSI in Hungary, and SRH in Poland (see appendix Table S2).

Figure 4 shows associations between being unmarried (never married, divorced and widowed) and the two health outcomes, adjusted for education and age. Results from analyses of samples pooled across countries showed that being unmarried was associated with poor health in both men and women. However country-specific analyses showed that the association between being unmarried and poor SRH only reached conventional levels of statistical significance in Belgium, Germany and, for women only, the Netherlands. There was more similarity between the surveys in the country-specific odds ratios than appeared the case on the analyses of differentials by low education, particularly for Belgium and Germany and, to a lesser extent, France and Hungary. Estimates for Poland appeared inconsistent being higher in the GGS than SHARE. We found evidence that the estimates were different between SHARE and GGS for Polish women and the outcome of LSI.

Discussion and conclusion

This paper makes a unique contribution by comparing the socio-demographic and health data of the population aged 50-80 years from the common SHARE and GGS surveys, and complements previous studies which have compared SHARE or GGS data to national population data or other surveys such as EU-SILC (Kreyenfeld, Hornung, and Kubisch 2013, Sauer, Ruckdeschel, and Naderi 2012, Fokkema, Kveder, and Liefbroer 2014, Croezen, Burdorf, and Lenthe 2013, Vergauwen et al. 2015). In addition, this paper extends previous work by comparing the socio-demographic gradient in health variables in the two surveys.

We assumed that estimates of basic demographic variables from both surveys would be similar after applying post-stratification weights but this was not always the case. Comparisons of weighted estimates of age from the two surveys with national data (Figure 1) show that the post-stratification weights provided, particularly in the GGS, were not always successful at adjusting to the target population. The factors used in the weights were inconsistent between the GGS surveys, and between the GGS and SHARE so differences seen could partly be a consequence of this. Tables 3 and 4 shows that the surveys are relatively comparable in terms of distributions by age, gender, marital status, proportions childless, and poor SRH (provided the same variant of SRH was used). Differences in education in Estonia, Hungary and Poland could be related to the different timing of surveys. In these countries the GGS and SHARE surveys were conducted 4-7 years apart, and changes in compulsory schooling occurring during the Soviet period could have resulted in changed distributions of education for older cohorts (Róbert 1991, Szebenyi 1992, Pennar, Bakalo, and Beredy 1971). However, it is worth noting that even when comparing the equivalent birth cohorts some differences remained for Hungary and Poland. There were also some differences for the Netherlands despite the two surveys

being conducted during a similar time frame. These differences may suggest inconsistencies in matching responses to ISCED codes.

The differences according to survey in the prevalence of LSI highlight how sensitive health reporting is to question wording and question order (Jette 1994, Picavet and van den Bos 1996, Dubuc et al. 2004, Freedman et al. 2004). In SHARE the question on LSI mentioned 'disability and infirmity' implying permanence which would usually lead to lower reporting of health problems (Picavet and van den Bos 1996). However survey content and the ordering of questions are also known to be important (Freedman et al. 2004, Bowling and Windsor 2008), with some studies suggesting that question order may have a stronger effect on older people's health assessments compared with younger people's (Crossley and Kennedy 2002). As commonly recommended, the question on SRH in both surveys was asked at the start of the health module, but this may produce biases in different cultural contexts (Lee and Grant 2009). The GGS collected a rather limited range of information on health whereas in SHARE it constituted a major part of the survey and this is likely to have meant respondents were more focussed on considering their health than in the GGS. Our finding that SHARE respondents report better SRH is consistent with a previous study comparing SHARE with other surveys namely the Health Interview Surveys (HIS), the European Social Survey (ESS), and the EU Statistics on Income and Living Conditions (EU-SILC) survey (Croezen, Burdorf, and Lenthe 2013).

The differences seen when comparing multivariable regression coefficients (Figures 3 and 4) could in many cases be explained by sampling variability. The main exception was Hungary, and to some extent Poland, where large significant differences were seen (Figure 3)

which might reflect differences between the surveys in distributions by education. There were also more differences between the surveys for multivariable associations with education compared to marital status, where the underlying distributions of the independent variable were more unequal. Overall this points to the importance of examining the distribution of underlying variables in detail when interpreting multivariable results.

The comparison has highlighted important differences between the surveys in their objectives and target populations which should guide researchers when choosing which survey to use. For studying the population over 50, the weighted estimates in SHARE usually better approximate age and sex distributions in the target population, however this could be resolved by more effective weights being developed for the GGS surveys. Measurement consistency is crucial for cross national comparisons, and some respects SHARE is more straightforward because the same instrument was used in all countries, rather than being harmonised post-hoc as in the GGS surveys. The SHARE survey asked a larger variety of health questions whereas the GGS has greater breadth in other areas such as attitudes. Our comparisons of health variables show that (contingent on similar wording) distributions and multivariable associations are relatively similar, but that researchers using the surveys jointly should pay attention to differences in question wording and representativeness when analysing the data and interpreting results.

Acknowledgements

This research was funded by a European Research Council Advanced Grant (FAMHEALTH) awarded to Professor Emily Grundy, reference number 324055.

GGS data were obtained from the Generations and Gender Programme Data Archive and were created by the organizations and individuals listed under each respective country name at http://www.ggp-i.org/online-data-analysis.html.

This paper uses data from SHARE Wave 4 release 1.1.1, as of March 28th 2013 (DOI: 10.6103/SHARE.w4.111) and SHARE Waves 1 and 2 release 2.6.0, as of November 29th 2013 (DOI: 10.6103/SHARE.w1.260 and 10.6103/SHARE.w2.260). The SHARE data collection has been primarily funded by the European Commission through the 5th Framework Programme (project QLK6-CT-2001-00360 in the thematic programme Quality of Life), through the 6th Framework Programme (projects SHARE-I3, RII-CT-2006-062193, COMPARE, CIT5- CT-2005-028857, and SHARELIFE, CIT4-CT-2006-028812) and through the 7th Framework Programme (SHARE-PREP, N° 211909, SHARE-LEAP, N° 227822 and SHARE M4, N° 261982). Additional funding from the U.S. National Institute on Aging (U01 AG09740-13S2, P01 AG005842, P01 AG08291, P30 AG12815, R21 AG025169, Y1-AG-4553-01, IAG BSR06-11 and OGHA 04-064) and the German Ministry of Education and Research as well as from various national sources is gratefully acknowledged (see www.share-project.org for a full list of funding institutions)."

Appendix/ supplementary material

Table S1: Factors used to develop post-stratification weights in GGS surveys

	Country specific weight factors wave 1
Belgium	age, sex, region
Estonia	age, sex
France	age, sex, citizenship, social and occupational status, type of household, number of household members, urbanization
Germany	age, sex, region, education
Hungary	Unknown
Netherlands	age, sex, region, urbanization, household type
Poland	No weights provided

Sources: Fokkema, Kveder, and Liefbroer (2014), GGS webpage http : //www.ggp-i.org/data/data-documentation.html.

Table S2: Results of testing for significant differences in odds ratios estimated from SHARE and GGS for the associations shown in Figures 3 and 4

	P values for significant difference in odds ratios estimated using SHARE and GGS							
	Outcome: Poor SRH-Predictor: low education		**Outcome: Long-standing illness Predictor: low education**		**Outcome: Poor SRH-Predictor: unmarried**		**Outcome: Long-standing illness-Predictor: unmarried**	
Country	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**
Belgium	ns	ns	0.049	ns	ns	ns	ns	ns
Estonia	0.036	ns	ns	ns	ns	ns	ns	ns
France	ns	ns	ns	ns	ns	ns	ns	ns
Germany	0.002	ns	ns	0.022	ns	ns	ns	ns
Hungary	0.003	<0.001	0.019	<0.001	ns	ns	ns	ns
Netherlands	ns	ns	ns	ns	ns	ns	ns	ns
Poland	0.020	0.001	ns	ns	ns	ns	ns	0.004
All pooled	ns	0.039	0.018	0.018	ns	ns	ns	ns

ns=non-significant (P<0.05). Source: SHARE baseline surveys and GGS wave 1

References

BÖRSCH-SUPAN Axel , BRUGIAVINI Agar, JÜRGES Hendrik, MACKENBACH Johan , SIEGRIST Johannes , and WEBER Guglielmo, eds. 2005. Health, ageing and retirement in Europe – First results from the Survey of Health, Ageing and Retirement in Europe. Mannheim, Mannheim Research Institute for the Economics of Aging (MEA).

BÖRSCH-SUPAN Axel, BRANDT Martina, HUNKLER Christian, KNEIP Thorsten, KORBMACHER Julie, MALTER Frederic, SCHAAN Barbara, STUCK Stephanie, and ZUBER Sabrina. 2013. "Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE)." International Journal of Epidemiology no. 42 (4):992-1001. doi: 10.1093/ije/dyt088.

BÖRSCH-SUPAN Axel, and JÜRGES Hendrik eds. 2005. The Survey of Health, Ageing and Retirement in Europe – Methodology. Mannheim, Mannheim Research Institute for the Economics of Aging (MEA).

BOWLING A, and WINDSOR J. 2008. "The effects of question order and response-choice on self-rated health status in the English Longitudinal Study of Ageing (ELSA)." Journal of Epidemiology and Community Health no. 62 (1):81-85. doi: 10.1136/jech.2006.058214.

CROEZEN Simone, BURDORF Alex, and LENTHE Frank J. van. 2013. Agreement and disagreement in prevalence estimates of health between SHARE and other European population studies. Department of Public Health, Erasmus MC, Rotterdam, The Netherlands.

CROSSLEY Thomas F., and KENNEDY Steven. 2002. "The reliability of self-assessed health status." Journal of Health Economics no. 21 (4):643-658. doi: http://dx.doi.org/10.1016/S0167-6296(02)00007-3.

DUBUC Nicole, HALEY Stephen, KOOYOOMJIAN Jill, and JETTE Alan. 2004. "Assessing disability in older adults: the effects of asking questions with and without health attribution." Journal of Rehabilitation Medicine no. 36 (5):226-231.

EUROSTAT. 2015. Database. Luxembourg, European Commission.

FOKKEMA Tineke, KVEDER Andrej, and LIEFBROER Aart C. 2014. Report and recommendations for sample and data adjustment procedures. Netherlands Interdisciplinary Demographic Institute GGP 212749 – Work Package 5 Panel methodology.

FREEDMAN Vicki A, CRIMMINS Eileen, SCHOENI Robert F, SPILLMAN Brenda C, AYKAN Hakan, KRAMAROW Ellen, LAND Kenneth, LUBITZ James, MANTON Kenneth, and MARTIN Linda G. 2004. "Resolving inconsistencies in trends in old-age disability: report from a technical working group." Demography no. 41 (3):417-441.

JETTE Alan M. 1994. "How measurement techniques influence estimates of disability in older populations." Social Science & Medicine no. 38 (7):937-942. doi: http://dx.doi.org/10.1016/0277-9536(94)90426-X.

JÜRGES Hendrik, AVENDANO Mauricio, and MACKENBACH JohanP. 2008. "Are different measures of self-rated health comparable? An assessment in five European countries." European Journal of Epidemiology no. 23 (12):773-781. doi: 10.1007/s10654-008-9287-6.

KAPITÁNY Balázs 2003. Turning points of the life-course. Sampling, reliability of raw data. http://www.demografia.hu/letoltes/dpa/panel1_minta_a.pdf.

KREYENFELD Michaela, HORNUNG Anne, and KUBISCH Karolin. 2013. "The German Generations and Gender Survey: Some Critical Reflections on the Validity of Fertility Histories." Comparative Population Studies – Zeitschrift für Bevölkerungswissenschaft no. 38 (1):3-28.

LEE Sunghee, and GRANT David. 2009. "The Effect of Question Order on Self-rated General Health Status in a Multilingual Survey Context." American Journal of Epidemiology no. 169 (12):1525-1530. doi: 10.1093/aje/kwp070.

MALTER Frederic, and BÖRSCH-SUPAN Axel. 2013. "SHARE Wave 4: innovations & methodology." Munich: MEA, Max Planck Institute for Social Law and Social Policy.

MANNHEIM RESEARCH INSTITUTE FOR THE ECONOMICS OF AGING (MEA). 2013. SHARE Release Guide 2.6.0 Waves 1 & 2 Mannheim Mannheim Research Institute for the Economics of Aging (MEA).

PENNAR Jaan, BAKALO Ivan I., and BEREDY George Z. F. 1971. Modernization and Diversity in Soviet Education. New York, Praeger.

PICAVET Susan, and VAN DEN BOS Geertrudis. 1996. "Comparing survey data on functional disability: the impact of some methodological differences." Journal of Epidemiology and Community Health no. 50 (1):86-93. doi: 10.1136/jech.50.1.86.

RÓBERT Péter. 1991. "Educational Transition in Hungary from the Post-War Period to the End of the 1980s." European Sociological Review no. 7 (3):213-236. doi: 10.2307/522691.

ROBINE JM, JAGGER C , and ROMIEU I. 2002. Selection of a coherent set of health indicators for the European Union. Phase II: final report. Montpellier, Euro-REVES.

SAUER Lenore, RUCKDESCHEL Kerstin, and NADERI Robert. 2012. "Reliability of retrospective event histories within the German Generations and Gender Survey." BIB Working Paper no. 1:2012.

SZEBENYI Péter. 1992. "Change in the Systems of Public Education in East Central Europe." Comparative Education no. 28 (1):19-31. doi: 10.2307/3099339.

UNITED NATIONS. 2013. World Population Ageing 2013. In ST/ESA/SER.A/348. New York, United Nations, Department of Economic and Social Affairs, Population Division.

VERGAUWEN Jorik, WOOD Jonas, DE WACHTER David, and NEELS Karel. 2015. "Quality of demographic data in GGS Wave 1." Demographic Research no. 32 (24):723-774.

VIKAT Andres, SPÉDER Zsolt, BEETS Gijs, BILLARI Francesco C, BÜHLER Christoph, DÉSESQUELLES Aline, FOKKEMA Tineke, HOEM Jan M, MACDONALD Alphonse, and NEYER Gerda. 2007. "Generations and Gender Survey (GGS): Towards a better understanding of relationships and processes in the life course." Demographic Research no. 17 (14):389-436.