Later life health in Europe: how important are country level influences?

George B. Ploubidis · Caroline Dale ·
Emily Grundy

Published online: 24 January 2012
© Springer-Verlag 2012

Abstract In this article, we examine the extent and pattern of country level differences in later life health in Europe and compare five competing explanations for this variation. We used data from 14 European countries, drawn from Northern (Denmark and Sweden), Western (Austria, France, Ireland, Germany Belgium, the Netherlands and Switzerland), Mediterranean (Spain, Italy and Greece) and Eastern (Poland and Czechia) regions of Europe, $N = 33,528$. Our results suggest that about a quarter (24%) of the overall variation in later life health in Europe appears to be due to country level differences. The Scandinavian countries along with Germany, the Netherlands and Switzerland appear to have the best health, whereas Spain, Italy and Poland had the lowest health score. Country level influences on health were largely associated with differences in the level of egalitarianism of each country as measured by the Gini coefficient, with more inequality being associated with poorer health. Differences in health-related lifestyle, as approximated by the prevalence of obesity in each country, also had a substantial macrolevel influence on later life health, with a lower national prevalence of obesity being associated with better health. Our results indicate the presence of systematic macrolevel health variation in Europe and suggest that policies to reduce income inequality as well as population interventions to promote healthier lifestyles and decrease the prevalence of obesity have the potential to improve

population health and potentially offset some of the challenges posed by population ageing in Europe.

Keywords Health inequalities · Ageing · Gini coefficient · Country level variation · Multilevel model · Income distribution

Introduction

Nearly, all European countries experienced major demographic changes in the twentieth century including declines in fertility, substantial increases in survival to older ages and, more recently, large reductions in late age mortality (Grundy 2006). Consequent large increases in the numbers and proportions of older people have focussed attention on the public health and policy importance of health status in later life. National differences in indicators of health at older ages may be valuable in providing insights into macrolevel determinants, including structural and policy influences (McKee 1998) and if alleviated have the potential to improve population health.

Findings from a recent study suggest that older people in continental Europe have better health compared with the USA and the UK (Avendano et al. 2009). However, differences between European countries in the health of their older populations have also been widely reported and seem consistent regardless of the health outcome analysed (Eikemo et al. 2008; Karlsson et al. 2010; Mackenbach et al. 2008; Minicuci et al. 2004; Olsen and Dahl 2007; Verropoulou 2009). In general, these results show poorer health outcomes for older people in Eastern and Southern compared to Northern and Western regions (Minicuci et al. 2004; Ploubidis and Grundy 2009; van den Brink et al. 2003; Minicuci et al. 2003).

Responsible Editor: H. Litwin (Guest Editor), D. J. H. Deeg.

G. B. Ploubidis (✉) · C. Dale · E. Grundy
Department of Population Studies, Faculty of Epidemiology and
Population Health, London School of Hygiene and Tropical
Medicine, University of London, London, UK
e-mail: George.Ploubidis@LSHTM.ac.uk

Several explanations for this variation have been suggested. For example, it has been proposed that characteristics of welfare states, such as social benefits and generally high social expenditure are related to better population health (Navarro and Shi 2001; Stuckler et al. 2010). Another explanation revolves around variations in economic growth and development which are viewed as important predictors of population health. In accordance with this hypothesis, GDP per capita has been consistently linked with population health (Beckfield 2004), with wealthier countries generally having better population health, although above a certain threshold of national income this association is much weaker (Bloom and Canning 2007). In addition, differences in the extent of income inequality are thought to influence country level variation, with several studies reporting an association between the Gini coefficient and population health (Wilkinson and Pickett 2006). A further explanation revolves around differences between European populations in health-related lifestyle (de Groot et al. 2004). Finally, from a psychosocial perspective, social capital and its dimensions, such as social trust and social participation have been suggested as important predictors of country level variation in health (Carlson 1998).

These possible explanations for differences between European countries in the health status of the older population are not mutually exclusive and indeed may be interrelated. For example, it has been argued that income inequalities may adversely influence population health because of their negative effects on social capital (Wilkinson 1997; Wilkinson and Pickett 2006). Furthermore, it could be argued that the effect of different welfare regimes on health is mediated by social inequalities, health-related lifestyle, social capital and a country's economic development. For example, welfare regime characteristics may reflect particular policies introduced to promote economic growth and development, reduce inequalities, enhance social capital or influence health related lifestyle.

A critical component of comparative investigations is valid, reliable and comparable measures of health. All cause mortality has been employed in country level comparisons (Coburn 2004) but for European countries which are all past the epidemiologic transition (Omran 1971) mortality is no longer adequate as the sole indicator of population health. Previous studies including measures of morbidity or health have often relied on self-reported indicators (Espelt et al. 2008) which are known to suffer from the influence of response bias, such as social desirability (Hebert et al. 2001). Evidence shows that, for example, there may be differences in the way which individuals of high and low socio-economic position assess their health (Dowd and Zajacova 2007) and most importantly differences between national populations (Desesquelles et al. 2009).

Previous country level comparisons suffer from a major methodological limitation as, with few exceptions (Ploubidis and Grundy 2009), the issue of between-country measurement invariance has not been addressed. Measurement invariance is a set of hypotheses stating that measurement model parameters function without bias across groups—countries in this case—or occasions (Meredith 1993). Meaningful country level comparisons require consideration of the measurement invariance of the health indicators used due to the possible influence of country-specific biases on measurement. Taking into account, these methodological challenges the objective of this study is to investigate the importance of country level variation in later life health in the older population of Europe and to empirically compare five explanations for this variation.

Method

Sample

The Survey of Health Ageing and Retirement in Europe (SHARE) is a multidisciplinary cross-national survey that includes data on the health, socio-economic status and social and family networks of non institutionalised individuals aged 50 or over (<http://www.share-project.org>). An overall response rate of 61.6% was obtained, varying across countries from 38.8 to 81.0%, with the lowest response rates observed in Switzerland and Belgium. Here, we employ data from 14 European countries included in the SHARE second wave (2006–2007, release 2-3-0). The countries included are drawn from Northern (Denmark and Sweden), Western (Austria, France, Ireland, Germany, Belgium, the Netherlands and Switzerland), Mediterranean (Spain, Italy and Greece) and Eastern (Poland and Czechia) regions of Europe. Of the 27,444 interviewees at the first wave of SHARE, 2.3% had died by wave 2 whilst another 1.7% had moved out of their respective country without leaving contact details. In wave 2, a refresher sample was drawn in all first wave countries except Austria and the Flemish part of Belgium. For the refresher sample the same sampling methods were used as in SHARE Wave 1, except that cohorts born in 1955 and 1956 were oversampled to keep the sample representative of the population 50 years old and older. We note that appropriate sampling frames for individuals, dwellings or households were not available in all countries and there were thus constraints on what kind of sample survey design was chosen. Specific details on the sampling frame for each country as well as further details of the SHARE sampling methodology have been reported elsewhere (Börsch-Supan and Jürges 2005). Our

Table 1 Descriptive statistics for individual level variables

	<i>N</i>	%
Gender		
Male	15,173	45.3
Female	18,355	54.7
Missing	0	0
Living arrangements		
Living with spouse/others	24,927	74.3
Living alone	8,596	25.6
Missing	5	0.01
Functional limitations		
Not limited	18,938	56.5
Limited	14,479	43.2
Missing	99	0.3
Symptoms		
Less than 2 symptoms	19,090	56.9
2+ symptoms	14,319	42.7
Missing	102	0.3
Self-rated health		
Excellent	3,003	9.0
Very good	6,240	18.6
Good	12,347	36.8
Fair	8,258	24.6
Poor	3,559	10.6
Missing	97	0.3
Smoking		
Current or past smoker	10,404	31.0
Never smoked	22,773	67.9
Missing	351	1.0
Physical activity		
Never	4,161	12.4
At least once a week	29,006	76.6
Missing	361	1.0
Chronic illness		
Less than 2 diseases	18,561	55.4
2 + chronic diseases	14,826	44.2
Missing	100	0.3
Mobility		
No problems	17,242	51.4
1 + mobility problems	16,160	48.2
Missing	108	0.3
ADL limitations		
No ADL limitations	29,873	89.1
1 + ADL limitations	3,530	10.5
Missing	109	0.3
IADL limitations		
No IADL limitations	27,753	82.8
1 + IADL limitations	5,650	16.9
Missing	109	0.3

Table 1 continued

	<i>N</i>		%	
Obesity				
No	26,372		78.7	
Yes	6359		19.0	
Missing	797		2.4	
	<i>N</i>	Mean	Std. deviation	Median
Age	33528	65.43	10.03	
Years of education	33313	10.00	4.54	
Numeracy score	33222	3.36	1.15	
Net household Income	33222	119506.3		30399.5
	<i>N</i>	Mean	Std. Deviation	
Grip strength	30,499	34.58	12.00	
Peak flow	29,757	335.43	159.72	
Alcohol use (drinks per week)	33,165	4.52	2.22	

analytic sample comprised 33,528 respondents to SHARE wave 2 and included participants with partially missing data, which as can be seen in Table 1, mostly occurred in the observer measured health indicators.

Measures

In the first stage of the analysis, we employed a latent variable modelling approach that allowed us to combine information from self-reported and observer measured health indicators under the assumption that they are manifestations of latent (not directly observed) “true” health status, as reflected in their shared (common) variance (Ploubidis and Grundy 2011). Population health is thus viewed as a variable whose true values cannot be directly observed (Rabe-Hesketh and Skrondal 2008). Two observer measured (grip strength and peak flow, which is a measure of respiratory function) and seven self-reported health indicators (self-rated health, the presence of long standing illness, the presence of one or more functional limitations, the presence of one or more problems with activities of daily living (ADL), the presence of one or more problems with instrumental activities of daily living (IADL), the presence of one or more symptoms and the presence of one or more mobility problems), were combined in the latent health dimension representing somatic/physical health (indicators of mental health were not included in the model). All indicators were recoded so high values on the latent dimension represent good physical health.

Table 2 Distribution of country level variables

	Obesity ^a (%)	Social trust ^a	Gini ^b	GDP per capita ^c
Austria	24	5.62	0.27	123
Germany	18	5.38	0.3	116
Sweden	16	6.55	0.23	123
Denmark	15	7.34	0.23	121
The Netherlands	15	6.32	0.27	132
Switzerland	13	6.49	0.28	141
Belgium	19	5.25	0.27	116
Ireland	27	6.19	0.32	148
Spain	25	5.6	0.32	105
Italy	19	4.71	0.35	104
France	16	4.57	0.28	109
Greece	21	4.7	0.32	93
Czechia	25	5.76	0.27	80
Poland	26	5.14	0.37	54

^a Obesity and social trust derived from the SHARE dataset

^b Derived from Organization for Economic Co-operation and Development (OECD), Inequality Data http://www.oecd.org/document/0,3746,en_2649_201185_46462759_1_1_1_1,00.html

^c Derived from Eurostat, <http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/>

In the second stage of the analysis, we estimated a multilevel model in an attempt to account for country level variation in later life health. We included as individual level predictors gender, age, living arrangements, net household income (recoded to 20 equal groups and used as a continuous variable in the model), years of education, smoking status, alcohol use (how many drinks on average per week in the past 3 months), a binary indicator of obesity (BMI > 29.99), based on self-reported height and weight, an indicator of moderate physical activity in the past week (some/a lot vs. none) and a numeric ability score (range 1–5) as a proxy measure of cognitive ability. In Table 1, we present descriptive statistics of all health indicators and individual level predictors.

As country level predictors (see Table 2), we employed a categorical indicator of the prevalence of obesity (BMI > 29.99) derived from the SHARE dataset, which we aggregated by country as a proxy measure of healthy lifestyle; the Gini coefficient as a measure of income inequality, and logged GDP per capita (adjusted for purchasing power standards) as a measure of standard of living. Information on the 2007 values of GDP per capita in purchasing power standards (PPS) expressed in relation to the European Union (EU-27) average of 100 was derived from Eurostat, and the Gini values (after taxes and transfers) for the mid 2000s were taken from OECD tables. We also used a binary dummy variable indicating whether or

not the country had a social democratic welfare regime. This variable distinguished Sweden, Denmark and the Netherlands from the other countries, in line with recent regime typologies (Chung and Muntaner 2007). Finally, an aggregated by country indicator of social trust was derived from the SHARE dataset and was employed as a measure of social capital (“would you say that most people can be trusted”?). Responses to the social trust item ranged from 0 to 10, with high scores indicating that people can be trusted. We decided to employ aggregate social trust as an indicator of country level social capital, as it can be argued other indicators, such as the availability of social contacts, social network resources and interactions are endogenous to health.

Statistical modelling

The issue of unbiased health measurement is usually approached with forms of factor analytic measurement models (Rabe-Hesketh and Skrondal 2008; Skrondal and Rabe-Hesketh 2008). At the individual level, these models, when appropriately used, capture the true variance of the construct that is being measured, whilst controlling for sources of error exogenous to health. However, this approach assumes a uniform distribution of error between countries. This assumption is unlikely to be valid as there are additional sources of error arising, for example, from the translation of health items to several languages and cultural differences in the understanding of items, in response tendencies, but also differences in the instruments used in the collection of observer measured health indicators. These may vary by country so their distribution as sources of error cannot be assumed to be uniform. The establishment of between-country measurement invariance where both uniformly and not uniformly distributed sources of measurement error are controlled is therefore needed to engage to a meaningful health comparison between countries as well as to further model country level variation.

Within the generalized latent variable modelling framework, measurement equivalence or invariance is analogous to factorial invariance. Factorial invariance is a set of hypotheses stating that measurement model parameters function without bias across groups or occasions. According to Drasgow (1984), measurement equivalence exists “when the relations between observed test scores and the latent attribute measured by a test are identical across subpopulations”. Factorial Invariance can be tested with multigroup confirmatory factor analysis (CFA) (Joreskog 1971). The multiple group (global) Chi-square test statistic, and the indices of fit derived from it for this multiple group model, are used to assess invariance, or in other words assess whether a confirmatory factor analytic model with fixed model parameters amongst groups fits the data. Strict factorial

invariance is achieved when measurement parameters (thresholds, factor loadings and their associated standard errors for binary or ordinal indicators, factor loadings, intercepts and standard errors for continuous indicators) function equivalently in each group of a multigroup confirmatory factor analytic model (Joreskog and Moustaki 2001).

The measurement model for health and the tests for invariance were based on a latent variable model appropriate for the combination of binary, ordinal and continuous indicators. The associations between the latent health continuum and the binary or ordinal health indicators were modelled with a 2 parameter (factor loadings and thresholds) probit regression. In this instance, factor loadings represent the strength of the association between the indicator and latent health, whereas the thresholds represent the level of latent health that needs to be reached for a particular response category in a categorical or ordinal health indicator to be endorsed. The part of the model where continuous health indicators are linked with continuous latent factors is a traditional confirmatory factor analytic model with linear regressions linking the latent health continuum with the observed health indicators. In this instance, intercepts and factor loadings are estimated to capture the association between the health indicators and latent health.

The between-country invariant latent variable measurement model was estimated with the Weighted Least Squares, Mean and Variance adjusted (WLSMV) estimator which is appropriate for multiple group analysis and invariance tests. Factor scores that represent individual later life health were derived from this latent variable measurement model and were used in further analyses. In theory, latent scores range from $-\infty$ (minus infinity) to $+\infty$ (infinity), but in practice the range is usually from -3 to 3 , with high scores in this case representing good health. Model fit was assessed with the Comparative Fit Index (CFI), the Tucker Lewis Index (TLI) and the root mean

square error of approximation (RMSEA). We note that values >0.95 for the CFI and TLI and values <0.06 for the RMSEA indicate good model fit.

In the second stage of the analysis, we estimated a multilevel model (random intercepts only, as the slope variation was negligible) using the between-countries equivalent (invariant) latent health outcome derived in the first stage of the analysis. Multilevel models have recently found prominence in the context of international comparisons (Jen et al. 2009a) as they overcome the ecological fallacy that has troubled previous research, but they have not been widely used in analyses of health differentials between the older populations of European countries. In the context of multilevel analyses where the sample size of the level two units is inevitably small (14 countries in our case), maximum likelihood estimation which relies on large sample (asymptotic) theory may produce biased estimates and the Bayesian estimation framework offers an attractive alternative. In this instance, we employed non informative priors using the Markov Chain Monte Carlo algorithm (two chains, 50,000 Bayes iterations) based on the Gibbs sampler. Model convergence was assessed with the proportional scale reduction (PSR) criterion (values close to 1 indicate model convergence). All models were estimated in Mplus 6 (Muthen and Muthen 1998–2010). Taking into consideration that unbiased parameter estimates cannot be obtained without properly addressing the implications of incompleteness, we analysed partially incomplete datasets using Multiple Imputation under the Missing at Random (MAR) assumption (Vogler et al. 1995; Little and Rubin 2002).

Results

The strict measurement invariance model, which implies that all parameters of the measurement model used to

Fig. 1 Crude/unadjusted country level means

derive the physical health population metric are equal between countries fitted the data, CFI = 0.97, TLI = 0.96, RMSEA = 0.058, and therefore allowed us to use the latent health continuum as a between-country equivalent health metric in all further analyses. In Fig. 1, we present the crude (unadjusted) latent health means for all countries within SHARE. Sweden, Denmark, Germany, the Netherlands and Switzerland were the countries with the highest latent health score, whereas Spain, Poland and Italy were the countries with the lowest health score. A one way analysis of variance revealed that the observed between-countries differences were significant, $F(13, 33459) = 612.456$, $P < 0.001$ and that the countries formed four homogeneous subsets with respect to somatic health. The first subset (best health) includes Sweden, the second includes Denmark, Germany, the Netherlands and Switzerland, the third the Czech Republic, Greece, France, Belgium and Ireland and the fourth (lowest health), Poland, Spain and Italy. The estimated intraclass correlation from a null multilevel model (the random effects equivalent to the fixed effects ANOVA) was 0.24, indicating that country of residence accounts for 24% of the overall later life health variance in Europe.

In Table 3, we present the standardised parameters and 95% Bayesian credibility intervals derived by the multilevel model. The largest PSR value in the first 100 iterations was 1.026 and declined sharply in the following

Table 3 Standardised multilevel model parameters and 95% Bayesian credibility intervals

Individual level	
	Health
Gender (Women)	−0.372 (−0.380 to −0.364)*
Age	−0.371 (−0.379 to −0.362)*
Years of education	0.096 (0.086 to 0.106)*
Income	0.072 (0.059 to 0.086)*
Living arrangements	0.018 (0.009 to 0.027)*
Smoking	−0.017 (−0.036 to −0.011)*
Alcohol	−0.048 (−0.057 to −0.038)
Physical activity	0.172 (0.164 to 0.181)*
Obesity	−0.025 (−0.035 to −0.016)*
Numeracy score	0.103 (0.093 to 0.112)*
Country level	
	Health
Obesity	−0.288 (−0.688 to −0.083)**
Social trust	−0.037 (−0.624 to 0.566)
Regime	0.039 (−0.555 to 0.621)
GINI	−0.453 (−0.820 to −0.080)**
GDP per capita	0.170 (−0.344 to 0.0617)

* $p < 0.001$

** $p < 0.05$

iterations, indicating model convergence. In Table 3, we present parameter estimates on the association between the predictors and health as derived from the multilevel model. As expected women had worse health compared to men, age had a negative association with health, whilst years of education were positively associated with health, as was net household income and living with a spouse or partner. Obese participants had worse health compared to the non obese, smoking was negatively associated with health, as was alcohol use. Exercise and numeric ability were both positively associated with health. At the country level, the prevalence of obesity had a negative association with health, as did the Gini coefficient. However, the indicator of welfare regime typology was not associated with later life health. Similarly, we did not observe a significant association between health and GDP per capita, or between health and social trust.

Discussion

Most studies of international differences in health within Europe consistently report between-country differences (Minicuci et al. 2004; Navarro and Shi 2001; Olsen and Dahl 2007; Verropoulou 2009). Considering the demographic ageing of European populations, the objective of this study was to investigate the extent and pattern of country level variation in later life health, as well as to empirically compare five explanatory macrolevel drivers of this variation. Our results suggest that almost a quarter (24%) of the overall variation in later life physical health in Europe is due to country level differences. The Scandinavian countries along with Germany, the Netherlands and Switzerland appear to have the best health, whereas Spain, Italy and Poland received the lowest later life health score. Despite methodological differences with our study, these findings are generally in agreement with previous studies in Europe (Chung and Muntaner 2007; Coburn 2004; de Groot et al. 2004; Minicuci et al. 2004; Olsen and Dahl 2007; van den Brink et al. 2003).

Income inequality was the strongest predictor of country level variation, with less inequality being associated with better health, followed by the prevalence of obesity. The finding that country level variation in health is largely associated with differences in the level of egalitarianism of each country as measured by the Gini coefficient is in accordance with some previous findings (Wilkinson 1997; Wilkinson and Pickett 2006), but not with others (Jen et al. 2009a, b; Lynch et al. 2004). However, these studies have not focussed on the older population of Europe. The association between GDP per capita and later life health was not statistically significant, although in the expected direction. This finding is in line with previous studies

suggesting a nonlinear association between GDP and health with diminishing marginal effects once a certain standard of living is achieved (Rodgers 1979; Bloom and Canning 2007). It appears that in the high income European countries considered here, more unequal distribution of income leads to a larger proportion of older individuals with incomes less than the minimum required for healthy living (Morris et al. 2000, 2007).

We found that the prevalence of obesity was a strong predictor of country level variation in later life health, with, as expected, a lower prevalence of obesity associated with good health, a finding in accordance with previously reported results (de Groot et al. 2004). Health-related lifestyle has a well-established strong association at the individual level with morbidity and mortality (Schrijvers et al. 1999). However, external constraints influence choices, making this not only an individual choice, but also one influenced by country level contexts, as different policies on the availability of high-quality foods and availability of opportunities for exercise may have an impact on the prevalence of obesity (Cockerham 1997).

We did not observe an association between our indicator of social trust and country level variation in later life health, a finding in agreement with some previous studies (Lynch et al. 2001; Olsen and Dahl 2007). Although it has been suggested elsewhere that income inequality leads to more violence, less social cohesion and lower levels of social trust (Wilkinson and Pickett 2006), it appears that at least for the older population in Europe the mechanism through which income inequality influences health is not related to social trust (Morris et al. 2000, 2007). Differences in social capital have been reported in Europe (Beugelsdijk and Van Schaik 2005; Kaariainen and Lehtonen 2006) but the variation between European countries may be less than in studies including a wider range of countries, which may explain the different result here. Another explanation may be that dimensions of social capital other than social trust may be more important for physical health, or that the influence of social trust is manifested mostly on mental health outcomes as suggested by the social stress hypothesis (Schwartz and Meyer 2010; Seeman et al. 1997). Furthermore, our indicator of social trust was relatively crude and based on only one indicator which may also have influenced our results.

In agreement with previous studies, the direct effect of the regime typology on health was negligible (Olsen and Dahl 2007). However, the social democracies were the countries with the best health, and the unadjusted regression parameter which captures the association between type of regime and health was significant. Therefore, we conclude that social democratic countries exhibit better population health status, but this effect is largely mediated by more equal distribution of income and to a lesser extent

by lower prevalence of obesity. This was confirmed by an additional model where the Gini coefficient and the prevalence of obesity were modelled as mediators of the association between regime type and health. Both indirect effects were significant, suggesting that in social democratic regimes policies, such as strong employment and wage protection, coupled with high unemployment benefits and pensions, lead to more equal income distribution and thus to better health (Beckfield and Krieger 2009).

A particular strength of this study is the use of a population health metric derived from a combination of observer measured and self-reported health indicators, as well as the establishment of its between-country measurement equivalence. The properties of the derived population health metric imply that random as well as systematic measurement error, such as between-country differences in responding to self-reported health indicators as well as differences in the instrumentation related to the observer measured indicators were statistically controlled. However, there are some limitations that should be considered whilst interpreting our results. First, a limitation which is present in all country level analyses in Europe is the relatively small number of level 2 units (countries) in the multilevel model. Although the Bayesian approach we adopted here specifically addresses this issue we conducted sensitivity analyses with several analytic strategies that have been previously used. We estimated a random effects model with the country level predictors included on the fixed part (individual level) of the model both with robust maximum likelihood (MLR), restricted maximum likelihood (REML) and Bayesian estimators. The results of these alternative analytic strategies were similar to the ones presented here indicating that the relatively small number of level two units did not bias our results.

Our analysis was carried out using partially incomplete data, which mostly occurred in the observer measured health indicators. We estimated several models for sensitivity purposes excluding the observer measured health indicators from the analysis thus increasing the proportion of complete cases in the analysis sample, as well as models with complete data (results available from corresponding author). The results of these models were similar with the one we present here, suggesting that the missing data generating mechanism was captured adequately by our modelling strategy. Another notable limitation is the variation in the household response rates that markedly differed between the countries we considered here. Non response may bias estimates on both the individual and macrolevels as results from survey research indicate that non response is generally higher in lower socioeconomic groups and in less healthy people (Cavelaars et al. 1998; Knesebeck et al. 2006). However, this possible selection bias due to non representation of the disadvantaged groups

would result in an underestimation of socio-economic disparities in health, indicating that the effects we present here may be underestimating the impact of socio-economic inequality on both the individual and country levels. Furthermore, even though indicators for all existing explanatory theories of between countries health differences were included, there was still unexplained country level health variance. This could be partly due to measurement error in the existing macrolevel indicators or the exclusion of important indicators of the existing theories, such as dimensions of social capital other than social trust. Unmeasured confounding may have also biased our estimates on the individual level and due to the nature of multilevel models this may have implications for the estimated country level variation.

Despite these limitations, in this study, we present evidence of systematic macrolevel variation in later life health in Europe that is theoretically amenable to change. Our results indicate that within country structural changes and policies to reduce income inequality and population interventions to decrease the prevalence of obesity could improve population health such policies lie beyond the realm of medical interventions and the health system alone (Stuckler et al. 2010). They involve actions that would form part of any broadly based strategy to promote population health, such as measures to improve diet, increase physical activity, prevent cognitive decline (Doyle et al. 2009), as well as investments in social policies which are likely to reduce within country health inequalities.

Acknowledgments George B. Ploubidis is supported by a Medical Research Council (MRC) Population Health Science fellowship, G0802442.

References

- Avendano M, Glymour MM, Banks J, Mackenbach JP (2009) Health disadvantage in US adults aged 50 to 74 years: a comparison of the health of rich and poor Americans with that of Europeans. *Am J Public Health* 99(3):540–548. doi:[10.2105/ajph.2008.139469](https://doi.org/10.2105/ajph.2008.139469)
- Beckfield J (2004) Does income inequality harm health? New cross-national evidence. *J Health Soc Behav* 45(3):231–248. doi:[10.1177/002214650404500301](https://doi.org/10.1177/002214650404500301)
- Beckfield J, Krieger N (2009) Epi + demos + cracy: linking political systems and priorities to the magnitude of health inequities—evidence, gaps, and a research agenda. *Epidemiol Rev* 31(1):152–177. doi:[10.1093/epirev/mxp002](https://doi.org/10.1093/epirev/mxp002)
- Beugelsdijk S, Van Schaik T (2005) Differences in social capital between 54 Western European regions. *Reg Stud* 39(8):1053–1064. doi:[10.1080/00343400500328040](https://doi.org/10.1080/00343400500328040)
- Bloom DE, Canning D (2007) Commentary: the Preston curve 30 years on: still sparking fires. *Int J Epidemiol* 36(3):498–499. doi:[10.1093/ije/dym079](https://doi.org/10.1093/ije/dym079)
- Börsch-Supan A, Jürges H (eds) (2005) The survey of health, aging and retirement in Europe: methodology. Research Institute for the Economics of Ageing, Mannheim
- Carlson P (1998) Self-perceived health in East and West Europe: another European health divide. *Soc Sci Med* 46(10):1355–1366
- Cavelaars A, Kunst AE, Geurts JJM, Crialesi R, Grotvedt L, Helmer U, Lahelma E, Lundberg O, Matheson J, Mielck A, Mizrahi A, Rasmussen NK, Regidor E, Spuhler T, Mackenbach JP (1998) Differences in self reported morbidity by educational level: a comparison of 11 Western European countries. *J Epidemiol Community Health* 52(4):219–227
- Chung H, Muntaner C (2007) Welfare state matters: a typological multilevel analysis of wealthy countries. *Health Pol* 80(2):328–339
- Coburn D (2004) Beyond the income inequality hypothesis: class, neo-liberalism, and health inequalities. *Soc Sci Med* 58(1):41–56
- Cockerham WC (1997) The social determinants of the decline of life expectancy in Russia and Eastern Europe: a lifestyle explanation. *J Health Soc Behav* 38(2):117–130
- de Groot L, Verheijden MW, de Henauw S, Schroll M, van Staveren WA, Investigators S (2004) Lifestyle, Nutritional status, health, and mortality in Elderly people across Europe: a review of the longitudinal results of the SENECA study. *J Gerontol Ser A-Biol Sci Med Sci* 59(12):1277–1284
- Desesquelles AF, Egidi V, Salvatore MA (2009) Why do Italian people rate their health worse than French people do? An exploration of cross-country differentials of self-rated health. *Soc Sci Med* 68(6):1124–1128
- Dowd JB, Zajacova A (2007) Does the predictive power of self-rated health for subsequent mortality risk vary by socioeconomic status in the US? *Int J Epidemiol* 36:1214–1221. doi:[10.1093/ije/dym214](https://doi.org/10.1093/ije/dym214)
- Doyle Y, McKee M, Rechel B, Grundy E (2009) Meeting the challenge of population ageing. *BMJ* 339. doi:[10.1136/bmj.b3926](https://doi.org/10.1136/bmj.b3926)
- Drasgow F (1984) Scrutinizing psychological tests: measurement equivalence and equivalent relations with external variables are the central issues. *Psychol Bull* 95(1):134–135
- Eikemo TA, Bambra C, Judge K, Ringdal K (2008) Welfare state regimes and differences in self-perceived health in Europe: a multilevel analysis. *Soc Sci Med* 66(11):2281–2295
- Espelt A, Borrell C, Rodriguez-Sanz M, Muntaner C, Pasarin MI, Benach J, Schaap M, Kunst AE, Navarro V (2008) Inequalities in health by social class dimensions in European countries of different political traditions. *Int J Epidemiol* 37(5):1095–1105. doi:[10.1093/ije/dyn051](https://doi.org/10.1093/ije/dyn051)
- Grundy E (2006) Ageing and vulnerable elderly people: European perspectives. *Ageing Soc* 26:105–134
- Hebert JR, Peterson KE, Hurley TG, Stoddard AM, Cohen N, Field AE, Sorensen G (2001) The effect of social desirability trait on self-reported dietary measures among multi-ethnic female health center employees. *Ann Epidemiol* 11(6):417–427
- Jen MH, Jones K, Johnston R (2009a) Global variations in health: evaluating Wilkinson's income inequality hypothesis using the world values survey. *Soc Sci Med* 68(4):643–653. doi:[10.1016/j.socscimed.2008.11.026](https://doi.org/10.1016/j.socscimed.2008.11.026)
- Jen MH, Jones K, Johnston R (2009b) Compositional and contextual approaches to the study of health behaviour and outcomes: using multi-level modelling to evaluate Wilkinson's income inequality hypothesis. *Health Place* 15(1):198–203. doi:[10.1016/j.healthplace.2008.04.005](https://doi.org/10.1016/j.healthplace.2008.04.005)
- Joreskog KG (1971) Simultaneous factor analysis in several populations. *Psychometrika* 36(4):409–426
- Joreskog KG, Moustaki I (2001) Factor analysis of ordinal variables: a comparison of three approaches. *Multivar Behav Res* 36(3):347–387
- Kaariainen J, Lehtonen H (2006) The variety of social capital in welfare state regimes: a comparative study of 21 countries. *Eur Soc* 8(1):27–57. doi:[10.1080/14616690500491399](https://doi.org/10.1080/14616690500491399)

- Karlsson M, Nilsson T, Lyttkens CH, Leeson G (2010) Income inequality and health: Importance of a cross-country perspective. *Soc Sci Med* 70(6):875–885
- Knesebeck OV, Verde PE, Dragano N (2006) Education and health in 22 European countries. *Soc Sci Med* 63(5):1344–1351. doi: [10.1016/j.socscimed.2006.03.043](https://doi.org/10.1016/j.socscimed.2006.03.043)
- Little RJA, Rubin DB (2002) Statistical analysis with missing data, 2nd edn. Wiley, Chichester
- Lynch J, Smith GD, Hillemeier M, Shaw M, Raghunathan T, Kaplan G (2001) Income inequality, the psychosocial environment, and health: comparisons of wealthy nations. *Lancet* 358(9277):194–200
- Lynch J, Smith GD, Harper S, Hillemeier M, Ross N, Kaplan GA, Wolfson M (2004) Is income inequality a determinant of population health? Part 1. A systematic review. *Milbank Q* 82(1):5–99
- Mackenbach JP, Stirbu I, Roskam AJR, Schaap MM, Menvielle G, Leinsalu M, Kunst AE, European Union Working Group on Socioeconomic Inequalities in Health (2008) Socioeconomic inequalities in health in 22 European countries. *N Engl J Med* 358(23):2468–2481
- McKee M (1998) An agenda for public health research in Europe. *Eur J Public Health* 8(1):3–7
- Meredith W (1993) Measurement invariance, factor analysis and factorial invariance. *Psychometrika* 58(4):525–543
- Minicuci N, Noale M, Bardage C, Blumstein T, Deeg DJH, Gindin J, Jylha M, Nikula S, Otero A, Pedersen NL, Pluijm SMF, Pluijm MF, Zunzunegui MV, Maggi S, Grp CW (2003) Cross-national determinants of quality of life from six longitudinal studies on aging: The CLESA Project. *Aging Clin Exp Res* 15(3):187–202
- Minicuci N, Noale M, Pluijm SMF, Zunzunegui MV, Blumstein T, Deeg DJH, Bardage C, Jylhä M, The CLESA working group (2004) Disability-free life expectancy: a cross-national comparison of six longitudinal studies on aging. The CLESA project. *Eur J Ageing* 1(1):37–44. doi: [10.1007/s10433-004-0002-5](https://doi.org/10.1007/s10433-004-0002-5)
- Morris JN, Donkin AJM, Wonderling D, Wilkinson P, Dowler EA (2000) A minimum income for healthy living. *J Epidemiol Community Health* 54(12):885–889
- Morris JN, Wilkinson P, Dangour AD, Deeming C, Fletcher A (2007) Defining a minimum income for healthy living (MIMHL): older age, England. *Int J Epidemiol* 36(6):1300–1307. doi: [10.1093/ije/dym129](https://doi.org/10.1093/ije/dym129)
- Muthen LK, Muthen BO (1998–2010) Mplus User's Guide. 6th edn. Muthen & Muthen, Los Angeles
- Navarro V, Shi L (2001) The political context of social inequalities and health. *Soc Sci Med* 52(3):481–491
- Olsen KM, Dahl S-Å (2007) Health differences between European countries. *Soc Sci Med* 64(8):1665–1678
- Omran AR (1971) Epidemiologic transition: theory of epidemiology of population change. *Milbank Mem Fund Quart* 49(4):509–538
- Ploubidis GB, Grundy E (2009) Later-life mental health in Europe: a country-level comparison. *J Gerontol Ser B-Psychol Sci Soc Sci* 64(5):666–676. doi: [10.1093/geronb/gbp026](https://doi.org/10.1093/geronb/gbp026)
- Ploubidis G, Grundy E (2011) Health measurement in population surveys: combining information from self-reported and observer-measured health indicators. *Demography* 48(2):699–724. doi: [10.1007/s13524-011-0028-1](https://doi.org/10.1007/s13524-011-0028-1)
- Rabe-Hesketh S, Skrondal A (2008) Classical latent variable models for medical research. *Stat Methods Med Res* 17(1):5–32. doi: [10.1177/0962280207081236](https://doi.org/10.1177/0962280207081236)
- Rodgers GB (1979) Income and inequality as determinants of mortality: international cross-section analysis. *Popul Stud-J Demogr* 33(2):343–351
- Schrijvers CTM, Stronks K, van de Mheen HD, Mackenbach JP (1999) Explaining educational differences in mortality: the role of behavioral and material factors. *Am J Public Health* 89(4):535–540
- Schwartz S, Meyer IH (2010) Mental health disparities research: the impact of within and between group analyses on tests of social stress hypotheses. *Soc Sci Med* 70(8):1111–1118. doi: [10.1016/j.socscimed.2009.11.032](https://doi.org/10.1016/j.socscimed.2009.11.032)
- Seeman TE, Singer BH, Rowe JW, Horwitz RI, McEwen BS (1997) Price of adaptation: allostatic load and its health consequences: MacArthur studies of successful aging. *Arch Intern Med* 157(19):2259–2268
- Skrondal A, Rabe-Hesketh S (2008) Latent variable modelling. *Stat Methods Med Res* 17(1):3–4. doi: [10.1177/0962280207081235](https://doi.org/10.1177/0962280207081235)
- Stuckler D, Basu S, McKee M (2010) Budget crises, health, and social welfare programmes. *BMJ* 340:c3311. doi: [10.1136/bmj.c3311](https://doi.org/10.1136/bmj.c3311)
- van den Brink CL, Tijhuis M, Kalmijn S, Klazinga NS, Nissinen A, Giampaoli S, Kivinen P, Kromhout D, van den Bos GAM (2003) Self-reported disability and its association with performance-based limitation in elderly men: a comparison of three European countries. *J Am Geriatr Soc* 51(6):782–788. doi: [10.1046/j.1365-2389.2003.51258.x](https://doi.org/10.1046/j.1365-2389.2003.51258.x)
- Verropoulou G (2009) Key elements composing self-rated health in older adults: a comparative study of 11 European countries. *Eur J Ageing* 6(3):213–226. doi: [10.1007/s10433-009-0125-9](https://doi.org/10.1007/s10433-009-0125-9)
- Vogler GP, Sorensen TIA, Stunkard AJ, Srinivasan MR, Rao DC (1995) Influences of genes and shared family environment on adult body mass index assessed in an adoption study by a comprehensive path model. *Int J Obes* 19(1):40–45
- Wilkinson RG (1997) Socioeconomic determinants of health: health inequalities: relative or absolute material standards? *BMJ* 314(7080):591–595
- Wilkinson RG, Pickett KE (2006) Income inequality and population health: a review and explanation of the evidence. *Soc Sci Med* 62(7):1768–1784. doi: [10.1016/j.socscimed.2005.08-036](https://doi.org/10.1016/j.socscimed.2005.08-036)