Investigating the macro determinants of self-rated health and well-being using the European Social Survey: methodological innovations across nations and time

Christopher Deeming and Kelvyn Jones
School of Geographical Sciences and the Centre for Multilevel Modelling
University of Bristol

Abstract

At present, there is debate over the relative importance and contribution of household income to well-being, and the link between economic growth, welfare and well-being is not fully understood. We sought to examine how changes in contextual and individual income (spanning the Great Recession) are associated with changes in self-reported well-being in the European Social Survey (ESS) 2002-2011. A multivariate multilevel analysis was performed on 237,253 individuals nested within 128 country cohorts covering 30 countries. In this article we focus specifically on the analysis and some of the methodological challenges and issues faced when making international comparisons across nations and time.

Key words:

Multilevel modelling

European Social Survey

Subjective well-being

Economic growth

Economic crisis

Welfare regimes

Inequality

Bibliography and acknowledgements

Dr Christopher Deeming (https://www.researchgate.net/profile/Christopher_Deeming) is ESRC Senior Research Fellow in the School of Geographical Sciences at the University of Bristol and Kelvyn Jones (https://www.researchgate.net/profile/Kelvyn_Jones/) is Professor of Human Quantitative Geography and Co-Director of the Centre for Multilevel Modelling (CMM) at the University of Bristol. CMM at Bristol develops new statistical methodology and the MLwiN software used here to address unsolved issues in quantitative modelling of social processes http://www.bristol.ac.uk/cmm/.

This article was presented at the 11th Annual ESPAnet (European Social Policy Analysis Network) Conference on Social Policy and Economic Development, 5-7 September 2013, Poznań University of Economics, Poland (Stream 3: 'Designing Comparative Social Policy Analysis'). We thank Jon Kvist and Olli Kangas for their invitation and the discussants. We are also grateful to IJS referees and editors for helpful comments on an earlier version of the manuscript. The writing of this article was supported by a three-year research council fellowship (UK Economic and Social Research Council award ES/K001353/1 'New Cultural Contradictions in European Societies').

Introduction

This article examines the relation between subjective well-being (SWB) and income and economic growth (GDP) in 30 European countries over five time points (spanning the Great Recession) using data from the European Social Survey (ESS, described in Appendix 1 below). Three measures of well-being—happiness, life satisfaction, and health—are examined simultaneously using a multilevel modelling framework to account for variability in nested relations including individual, country, welfare state regime and time by survey wave.

The results reveal the importance of income for health and well-being at both the individual and country levels and, how redistributional welfare policies can be adjusted to reflect the impact of relative as well as absolute wealth and income on well-being. Before describing the study methods, and results, and discussing their implications in the face of growing levels of global inequality (OECD, 2008, 2011), we consider the methodological challenges and issues we faced in making international comparisons across nations and time.

Why multilevel?

There is a natural congruence between the multilevel random coefficients approach and the nature of the research questions we wanted to address in our analysis, but it is worth providing a few more reasons why this approach has been adopted, albeit briefly, in comparison to the alternatives, particularly because this is not the standard approach. The research questions that motivated out analysis are:

- How big are between country differences in SWB; are they changing or stable over time; and is this different or the same for each of the well-being outcomes?

- How correlated are the three SWB measures at the country, country-wave, and individual level; are they measuring the same underlying dimension of well-being, and do they change consistently together over time ?

- Are changing country differentials accounted for by different welfare regimes, or equivalently how have differential welfare regimes fared in the 2000's in terms of our three outcomes?

- Are the differences between countries and welfare regime types an artefact of differential and changing country income as measured by GDP; is it simply wealth or do different institutional forms have differential impact on citizens?

- Are the differences between countries and welfare regime types simply an artefact of individual characteristics such as income and demographic variables?

The multilevel approach is ideal in this situation because it transcends the dilemma of the atomistic and ecological fallacies (Subramanian et al., 2009) by modelling both individuals and their contexts simultaneously. It would have been possible to take an aggregate approach where the means of the outcomes are simply regressed on the means of the predictors but this would have ignored within-country relationships with potentially problematic results (Jones, 1990). Another approach would be to retain the individual data and perform a 'contextual analysis' (Boyd and Iversen, 1979) by including country variables in a single-level standard regression model but this would have profound implications for the standard errors of country-level variables as they would be estimated on an apparent three quarters of a million degrees of freedom and not the true number of thirty.

Another option is to adopt a fixed-effects approach in which a set of dummy variables is included for each and every country. Indeed, it is fair to say that this approach is currently the 'default option' when using time-series-cross-sectional data (e.g. Angrist and Pischke, 2009). This is prompted by fears of potential correlation between lower-level covariates and higher-level residual differences, a case of omitted variable bias. Fixed-effects models hold constant and control for differences among countries, so that attention focusses solely on the longitudinal within-country differences. However, as argued by Bell and Jones (2015) and demonstrated in their simulation study, this bias can be mitigated as much by random effects as fixed effects if the within and between, mean-centred approach they outline is adopted. The fixed effects approach moreover, cannot handle time-invariant processes (such as being in a particular welfare regime) that can have an effect on time-varying variables, and any time-varying covariate can have time-invariant 'between' effects which can be different from time-varying effects of the same variable. Only a random effects model allows these processes to be modelled simultaneously. The central aim of this study – the analysis of time-invariant welfare regime effects – is not possible with fixed effects because the inclusion of a set of country dummies would have consumed all the degrees of freedom at that level, and would have left nothing to be modelled. By treating each country as an island we would have explained everything (a perfect fit with no residual differences) but would have gained no understanding. The random-coefficients approach not only allows the modelling of country level predictors but also permits the assessment of how much variation between and within countries remains after taking account of predictor variables.

In short, the multilevel approach handles multiple outcomes in an overall model, permits imbalance and missingness, partitions residual variance into between- and within-group components, and allows for the assessment of correlation between outcomes at each level. It also allows for un-observables at each level, corrects standard errors for mis-estimated precision, analyses micro and macro models simultaneously and thereby address our research questions in a richer and more robust manner than the alternatives. Finally, the random effect estimates (of for example country by wave differentials) are precision-weighted so that estimates that are based on a small number of observations are automatically down-weighted or shrunk in the analysis (Jones and Bullen, 1994; Jones et al., 2014).

The different types and measures of well-being

Well-being can be seen as either subjective (such as whether an individual feels in good health, happy, or satisfied with their life) or objective (involving indicators such as life expectancy, education, and per capita income, see Sudhir and Sen, 1994). In this article, we focus on subjective measures of health and well-being, thus the measures of happiness, life satisfaction and self-rated health drawn from the ESS. The ESS asks respondents:

- Taking all things together, how happy would you say you are? Please answer using this card, where 0 means extremely unhappy and 10 means extremely happy.
- All things considered, how satisfied are you with your life as a whole nowadays? Please answer using this card, where 0 means extremely dissatisfied and 10 means extremely satisfied.
- How is your health in general? Would you say it is very good (1), good (2), fair (3), bad (4) or, very bad? (5).

The 'life satisfaction' measure, for instance, forms part of the 'life evaluative' approach to SWB measurement (OECD, 2013). The ESS question asks individuals to reflect on their life and make a cognitive assessment of how their life is going overall. Our life satisfaction measure is therefore dependent on a global appraisal of quality-of-life (Diener, 2009). The 'affect' approach to SWB seeks to measure people's feelings and emotions (OECD, 2013). The general state of 'happiness', as used in this study, forms part of the more global cognitive appraisal of subjective well-being (Diener, 2009). While our self-rated health measure addresses health-related aspects of quality-of-life (Bowling, 2004). The measures – or indicators – are generally held to be distinctive, they tap into different aspects of human well-being and quality-of-life. Therefore the strength of socioeconomic associations may depend on the type of well-being indicator under consideration. Here we retain them as separate entities yet model them simulataously so as to assess how they inter-relate with each other and how they relate to other variables and how they change over space and time.

Compositional influences, income

People require incomes to be able to purchase goods and services to sustain their health and well-being; the fewer goods and services are provided publicly, the more important individual and household income becomes (Marmot, 2002). However, the relationship of SWB to poverty is not fully understood and may depend on the type of well-being indicator under consideration (Subramanian and Kawachi, 2004; Jen et al., 2009; Povey et al., 2013). For this reason Stiglitz et al. recommend that researchers and policymakers give much more prominence to the distribution of income in the assessment of well-being (Stiglitz et al., 2010: 13-15, Recommendation 4).

Although income is a metric central to many poverty measures there are often issues over definition and measurement, see for example the report by the Expert Group on Household Income Statistics (Canberra Group, 2011). Many international surveys pose major challenges for researchers hoping to correlate SWB with household income. The ESS does not capture precise information on household income for example, instead a categorical scale is used. Moreover, there have been fundamental changes to the way income is captured and coded since the survey first began (Hoffmeyer-Zlotnik and Warner, 2006). In ESS Rounds 1-3 household income was obtained in standard income categories, as 12 bands ranging from below €1.8 thousand Euro per year to over €120 thousand, defined irrespective of country. In Rounds 4-5 the ESS deployed national categories based on deciles of the actual household income range in the given country. In order to achieve continuity in the study it was necessary to 'retrofit' income in the earlier waves into deciles to obtain compatible data across multiple rounds (see Appendix 2 for further details).

Contextual influences, national per capita income

The association between income and wellbeing is complex. In a well-known contribution, Easterlin (1974) observed the lack of relationship between GDP growth changes in happiness in the United States. Presumptions about an automatic link between economic growth and the promotion of citizen well-being are once again under the spotlight (Stiglitz et al., 2010). There has been a considerable amount of work claiming to support (e.g. Ovaska and Takashima, 2006) or refute the famous Easterlin 'paradox' (e.g. Deaton, 2008; Stevenson and Wolfers, 2008). Recent research suggests both sides may be correct; in that income has both an absolute and a relative effect on well-being, but with absolute effects becoming less important as national income rises (e.g. Clark et al., 2008; Graham et al., 2010). Thus, average happiness may become less sensitive to rising GDP in richer countries.

Here we are interested in the relationship between GDP, which is a country-level predictor variable, and well-being across the diverse European societies. GDP per capita data from the World Bank provided complete coverage across time and for our study sample of European countries. Eurostat or OECD GDP figures were not an option here as they did not extend to all of the countries in our study.

Welfare regimes

Welfare and other social protection policies in European welfare regimes act to reduce health and social inequities arising from social and labour market conditions (Dahl et al., 2006; Beckfield and Krieger, 2009); welfare regimes are also one of the main mechanisms driving population health and well-being (Pacek and Radcliff, 2008; Deeming and Hayes, 2012). Of the many functions of the 'welfare state', two are particularly prominent, as Barr (2012) observes: the 'Robin Hood' function which operates to redistribute resources within society

(i.e., between members) in order to promote social well-being; and the 'piggy bank' function, which is concerned with the redistribution of resources over the lifecycle in order to promote individual well-being (i.e., 'from cradle to grave'). These functions of the welfare state are the principal mechanisms by which the advanced economies help to promote population health and well-being, collectively guarding citizens against adverse social risks, such as unemployment and poverty.

In our study we draw on the Esping-Andersen (1990, 1999) framework of welfare regimes: the Social Democratic (Nordic), Conservative (Corporatist) and Liberal worlds of welfare state capitalism that reflect the dominant political movements and forms of political economy identified during the last century (Table 1). The Nordic welfare regime is characterised by universalism and a comprehensive system of social protection in contrast to the 'residual' system of welfare provision targeted at the poorest sections of society in the Liberal regime. In Continental Europe 'societal corporatism' emerged out of industrialization, with 'Bismarkian' systems of worker insurance and company-based social protection schemes. Here welfare has traditionally been organised around the principle of subsidiarity, occupational and social insurance schemes are dominant. Scholars have extended the original framework of welfare regimes to include a Southern European (Mediterranean) country cluster (Ferrera, 1996), and an Eastern European (Post-Communist) group (Fenger, 2007), which we also include in our analysis (Table 1). The Mediterranean welfare states are described as 'rudimentary' because they are characterised by their fragmented system of welfare provision. Reliance on the family and voluntary sector is also a prominent feature. The formerly Communist countries of Eastern Europe have experienced extensive economic upheaval and have undertaken extensive social reforms. There has been a shift towards policies associated more with the Liberal welfare regime, notably the marketization of social services. The field is contested however (Arcanjo, 2011; Ferragina and Seeleib-Kaiser, 2011). Italy was Conservative in Esping-Andersen's original work, while Switzerland was cross-classified: being Liberal on his social stratification index and Conservative on his index of decommodification (i.e., the generosity of social security). Recent scholarship now situates Italy within the Southern European model, given the share of common political-economic and geographical traits (Ferrera, 2005), while Switzerland now conforms more closely to the Continental European Conservative regime type according to most experts (Obinger et al., 2010).

The aggregate level of well-being is believed to be higher within Nordic welfare states, and its distribution more equitable (Wilkinson and Pickett, 2009; Kvist et al., 2011). However, too much state intervention may impose on individual freedoms, undermine people's resilience and self-reliance − all of which might have a negative effect on population well-being (Welzel and Inglehart, 2010). People may resent having to pay higher taxes for more expansive social provision, and higher tax 'burdens' may mean there is less individual freedom to choose. This is the classic argument usually advanced by those on the right of the political spectrum, who tend to argue for smaller governments. On the other hand, people may feel dissatisfied if everyday life risks become understood as issues of personal failure rather than social risks to be addressed through collective action. The welfare regimes defined in Table 1 are time

invariant classification of countries. We are assuming quite reasonably in this context that no country has undergone a fundamental change of regime during the study period (cf. Bambra et al., 2010).

Methodology

Data

A multivariate multilevel analysis was performed on 237,253 individuals nested within 128 country cohorts covering 30 countries participating in the ESS over five waves, 2002-2011 (Table 1). The ESS is a pan-European cross sectional time series survey running every two years (described in Appendix 1 below). Unfortunately, not all countries participated in each wave of the ESS however, and this missingness needs to be handled by our modelling approach. Table 1 also shows participation rates at the country-level, and overall there is good coverage with most countries being surveyed on most waves.

<<Insert Table 1 about here>>

Our resulting study sample therefore was potentially three SWB responses for 237,253 individuals giving 711,759 records. In fact some survey respondents did not respond to the outcomes and the final number of records to be analysed was 708,241 observations. There are 30 countries and they were measured on up to 5 waves. This would give 150 country-waves but in fact only 128 were observed as not all countries were surveyed at each wave (Table 1).

In this study we are interested in modelling the underlying propensity of being Un-Happy, Dis-Satisfied and Un-Healthy. Therefore, the three ESS outcomes measures, for Happiness, Satisfaction and Health (Table 2), have been coded as binary outcomes (see Borrell et al., 2009; Deeming, 2013). For instance, ESS responses of four or below to the question about Happiness are coded as (1) 'Unhappy', while those scoring five or above are coded as (0) 'Happy'. This dichotomy was repeated for Dissatisfaction. Responses of 'bad' and 'very bad' to the question about Health are coded as (1) 'Unhealthy' ('very good', 'good' and 'fair' are coded 'Healthy'). The ESS contains a range of socio-demographic data; we include respondents' age and sex as well as their household income (Table 2) and we include GDP per capita data available from the World Bank (Appendix 3).

<<Insert Table 2 about here>>

Model structure

There are a number of key features of the research questions and the nature of available data that require careful attention and when addressed appropriately, and in combination, require highly innovative modelling. The following aspects of the data structure are of particular importance.

- There are three outcomes: happiness, satisfaction, and health which require to be modelled simultaneously. We do this through a multivariate multilevel model whereby the three outcomes at level 0 are conceptualised as nesting within the individual at level 1.

- The three outcomes are all binary and cannot adequately be estimated by Normal-theory models. Instead we use a generalised multilevel model in which the propensity to be in a discrete state is non-linearly related to the predictor variables through a logit link (Jones and Subramanian, 2013: Chapter 12). The unexplained differences between individuals are treated as a Bernoulli distribution, and a covariance structure is specified to allow correlations between the outcomes. These correlations assess the relations at the individual level in answering 'yes' to each pair of outcomes. The random coefficients multilevel methodology (Duncan et al., 1998) deals efficiently with missing observations; as long as at least one of the three outcome responses is observed on an individual we are able to include that case in our analysis.[1]

- Different individuals have been asked in different rounds of the ESS about the three responses, so that we are dealing with a repeated cross-sectional data. This can be seen quite naturally as a hierarchical structure in which responses at level 0 are nested in individuals at level 1 who are nested in waves at level 2 and countries at level 3; thus we require a four level multilevel structure to be modelled. It is worth stressing that the wave is really country-by-wave so that a differential is modelled for a country overall (level 3). Then around this overall differential, there is an allowed to vary difference at each observed wave (level 2). The random coefficient approach adopted here does not require completely balanced data so that we can analyse country differences when the country has not appeared in each and every wave (e.g. Italy, see Table 1).[2] At each of the higher levels there is a variance-covariance matrix of the unaccounted differentials; this allow us to assess the size of the country differentials for each response and how volatile these differentials are over time.

The terms in the model have so far been concerned with allowed-to-vary differentials between country, wave and individuals (i.e., the random part of the model). But we also want to include variables at each level which may account for these differentials - the fixed part of the model, which models the underlying overall (mean) relationships between the three outcomes and predictor variables. We can include variables at each of the higher levels be they measured for individuals, waves or countries. There are a number of distinctions that need to be made.

- The type of welfare regime is a time-invariant macro variable as we assume no country has undergone a fundamental change of regime during the study period (but see our cautionary note on page 31). By including this variable as a set of dummy variables and interacting this with a time variable we can analyse what Fairbrother (2014) calls societal growth curves. Although we are not dealing with a true panel survey with repeated measures of an individual, we can study societal change as we have repeated measures on individual countries. This approach comes out of studying school effectiveness as schools are assessed in terms of their changing performance as successive cohorts graduate from the school (Jones and Subramanian, 2013, Chapter 11). Technically, these growth models involve including an interaction between a time variable and the categorical predictor of type of welfare regime. We assess whether the rate of change over time depends on time-invariant predictors.

- GDP is a time varying macro variable as we can measure this for each country at each wave. Moreover, using the approach explained in Bell and Jones (2015) we can separate the cross-sectional effect of GDP on the three outcomes from the longitudinal effect which estimates the changing effect of GDP as country income increases or decreases. This is readily accomplished by calculating the mean GDP across the waves and subtracting it from the country-wave GDP and including both the mean (which is now a time–invariant country variable) and the differential in the fixed part of the model. This group–mean centred approach results in two orthogonal predictors that allow differing within and between relations. This permits possibly different relations between the outcomes and enduring and changing country wealth. The model also includes an explicit time variable to allow for the possibility of a simultaneous but unrelated time trends in both predictors and outcomes. Assessing relationships in this longitudinal way is important for validating that the relationship is not spurious and operating through some unobserved confounders. It may appear rather ambitious with only a 5 country-waves (although covering a ten year period) to attempt this sort of modelling. But according to the simulations of Fairbrother (2014) 'these techniques appear relatively robust to real-world complication' and he notes that observing even just five rather than two country-waves per country has substantial benefits in terms of parameter precision. The reliability of the effects will also be affected by the variation in the predictor variables and the substantial turbulence in GDP of European countries in the 2000's will help the precision of the estimates.

The final set of variables are the household income and demographic variables which are observed on individuals at level 1; these can therefore be included in the fixed part of the model so that we estimate an overall effect on each of the outcome variables separately. Given the above arguments, it might be considered appropriate to include individual income both as individual and as group centred mean. However, this is not necessary as effectively this is already achieved by including GDP in cross sectional and longitudinal form. We found a remarkably high correlation of 0.92 between country-wave GDP and aggregated country-wave mean income; while Country GDP across waves is correlated 0.98 with Country aggregated mean income. Consequently by including longitudinal and cross-sectional GDP we are

effectively also taking account of cross-sectional and longitudinal country income. Moreover, the individual income data we have used (see Appendix 2) is the country-wave specific deciles, so individual household income is being measured relative to a country's overall income profile.

Model specification

Schematically the model of interest can be specified as follows as a four-level model:

$$y_{1,ijk} \sim Bernouilli(1, \pi_{1,ijk})$$

$$y_{2,ijk} \sim Bernouilli(1, \pi_{2,ijk})$$

$$y_{3,ijk} \sim Bernouilli(1, \pi_{3,ijk})$$

$$logit(\pi_{1,ijk}) = \beta_{1,0} + \sum_{l=1}^{L} \beta_{1,l}\, x_{lijk} + \sum_{m=L+1}^{M} \beta_{1,m}\, x_{mjk} + \sum_{n=M+1}^{N} \beta_{1,n}\, x_{mk} + (v_{1,k} + u_{1,jk})$$

$$logit(\pi_{2,ijk}) = \beta_{2,0} + \sum_{l=1}^{L} \beta_{2,l}\, x_{lijk} + \sum_{m=L+1}^{M} \beta_{2,m}\, x_{mjk} + \sum_{n=M+1}^{N} \beta_{3,n}\, x_{nk} + (v_{2,k} + u_{2,jk})$$

$$logit(\pi_{3,ijk}) = \beta_{3,0} + \sum_{l=1}^{L} \beta_{3,l}\, x_{lijk} + \sum_{m=L+1}^{M} \beta_{3,m}\, x_{mjk} + \sum_{n=M+1}^{N} \beta_{3,n}\, x_{mk} + (v_{3,k} + u_{3,jk})$$

$$\begin{bmatrix} v_{1,k} \\ v_{2,k} \\ v_{3,k} \end{bmatrix} \sim N(0, \Omega_v), \qquad \Omega_v = \begin{bmatrix} \sigma_{v1}^2 & & \\ \sigma_{v1v2} & \sigma_{v2}^2 & \\ \sigma_{v1v3} & \sigma_{v2v3} & \sigma_{v3}^2 \end{bmatrix}$$

$$\begin{bmatrix} u_{1,jk} \\ u_{2,jk} \\ u_{3,jk} \end{bmatrix} \sim N(0, \Omega_u), \qquad \Omega_v = \begin{bmatrix} \sigma_{u1}^2 & & \\ \sigma_{u1u2} & \sigma_{u2}^2 & \\ \sigma_{u1u3} & \sigma_{u2u3} & \sigma_{u3}^2 \end{bmatrix}$$

$$CovVar \begin{bmatrix} y_{1,ijk}|\pi_{1,ijk} \\ y_{2,ijk}|\pi_{2,ijk} \\ y_{3,ijk}|\pi_{3,ijk} \end{bmatrix} = \begin{bmatrix} \pi_{1,ijk}(1-\pi_{1,ijk}) & & \\ \rho_{12} & \pi_{2,ijk}(1-\pi_{2,ijk}) & \\ \rho_{13} & \rho_{23} & \pi_{3,ijk}(1-\pi_{3,ijk}) \end{bmatrix}$$

Examining the elements of the equation, we see that here are three observed binary outcomes for Happiness, Satisfaction and Health (y_1; y_2; y_3) which are observed for individuals *i* at country-wave *j* in country *k*. These are assumed to come from a Bernoulli distribution with an underlying propensity ($\pi_{1;}\pi_2; \pi_3$) to be Unhappy, Dissatisfied and Unhealthy which is related to a set of predictors through a logit link.

There is a separate equation for each outcome in which β_0 is the estimated logit across individuals, waves and countries when all the predictor values are zero; $\sum_{l=1}^{L} \beta_l\, x_{lijk}$ are the L fixed effects of the micro individual variables (indexed by l from 1 to L); $\sum_{m=L+1}^{M} \beta_m\, x_{mjk}$ are the M fixed effects of the time-varying variables measured for countries at each wave (indexed

by m from L+1 to M) ; and $\sum_{n=M+1}^{N} \beta_n x_{nk}$ are the N fixed effects of time-invariant country variables (indexed by n from M+1 to N) . The predictor variables can range from being variable across individuals within waves within countries (and are subscripted *ijk*) or invariant within, but variable between countries (as shown by subscript *k*).

In the random part of the models, there is for each response a differential residual (v_k) for each country, which can be seen as an allowed to vary departure from β_0 ; while u_{jk} is a residual for each wave *j* within country *k*; which is a random departure from the appropriate v_k. At each level (except level 0 which is merely used to represent that responses are nested within individuals) there is a variance covariance matrix with distributional assumptions. Thus, the between-country differentials (the v_k's) are assumed to come from a joint Normal distribution so that $\sigma_{v1}^2$is the between country variance for Happiness on the logit scale; while $\sigma_{v1v2}$is the covariance between the country differentials for Happiness and Satisfaction. Similarly, σ_{u3}^2 is the variance of the between country-wave differentials for Health and σ_{u1u3} is the covariance between the country-wave differentials for Happiness and Health. These covariances when suitably scaled give the correlation at the country and country-wave for the differentials for the three responses.

At level 1, the level of the individual, each of the observed binary responses conditional on the underlying propensity is specified to have a Bernoulli variance given by $\pi_{ijk}(1 - \pi_{ijk})$, while ρ_{13}, for example, gives the correlation between the binary response for being Happy and being Healthy. All the random part differentials are conditional on taking into account what has been estimated in the fixed part, so that the variance (σ_{v1}^2) of the between-country differences for Happiness could be estimated before and after including the GDP of a country. Similarly ρ_{13}would represent the partial correlation between being Happy and being Healthy if individual income is included as a level 1 predictor in the fixed part.

Model estimation

There are a number of choices to be made in model estimation and this is particularly important because we are dealing with discrete outcomes. Essentially there are three approaches: quasi-likelihood, maximum likelihood using quadrature, and Markov Chain Monte Carlo (MCMC). The key trade-off is between the quality of the resultant estimates and the computer time taken to achieve them given that an iterative process is required as there is no available analytic solution. It is known that the (much) more computer intensive MCMC and likelihood methods give better (less biased) estimates than quasi-likelihood (Rodriguez, 2008) we have used. However the key study for this finding, that of Rodriguez and Goldman (1995) is based on a family-level analysis that is very unlike the current study in that there were large higher-level variances, and few level-1 individuals nested in many families. Even then in their study the much faster Predictive Quasi Likelihood (PQL) with second order Taylor series expansion of Goldstein and Rasbash (1996) performed relatively well in producing less-biased estimates. This PQL method which incorporates the random effects into the linearization of the logit exploits the fast algorithm of Goldstein and Rasbash's (1992) Iterative Generalised Least

Squares (IGLS) as implemented in the *MLwiN* software but even that is challenged as time dependency is proportional to the cube of the number of lower level units in the largest higher level unit and in this study the UK has some 33,000 responses.[3] Nevertheless, the analysis of the full 700,000 plus records with complex structure and multiple random coefficients at each level was undertaken with each model calibration taking a day or so on a standard desktop computer.

The next question is the likely quality of the estimates when there are relatively few higher-level units, thirty countries, on which to base the analysis and likelihood procedures are based on asymptotic theory. Stegmueller (2013) presents simulation results that address the current problem with few higher-level units and a large number of lower level units. Our reading of his findings is that that even for cross-level interactions involving micro and macro variables that likelihood-based procedures are relatively unproblematic in terms of point estimates and confidence intervals once there are more than 15 countries as there are here (n=30).

Post-estimation

The model estimates are given in terms of logits and detailed and large estimate tables are available from the authors. However, to facilitate interpretation and dissemination we have mainly converted the logits of the fixed part estimates into relative odds and displayed these in graphical form. In particular we have set the lowest values of the odds to one, and compared the relative odds against this base. We have displayed, whenever possible without cluttering the graph, the 95% confidence intervals and used the same scale for the vertical axes for the graphs for each of the outcomes. Where this is not possible we have said so in the accompanying text. We have used the recently introduced customised predictions facility of MLwiN 2.27 to achieve this (Rasbash et al., 2009).

The variance-covariances at the higher levels are also estimated on the logit scale. The covariances we have turned to dimensionless correlation coefficients by dividing them by the product of the square root of the associated variances. To get a feel for the relative size of the variances we have used the Median Odds Ratio (MOR) of Larsen and Merlo (2005). The MOR transform the higher between-country variance on the logit scale to a much more interpretable odds scale than can be compared to the odds ratio of terms in the fixed part of the model. MOR can be conceptualised as the increased odds (on average, hence the median) that would result from moving from a lower to a higher 'risk' area if two countries were chosen at random from the distribution with the estimated level 2 variance.

Modelling strategy

The primary focus is on the effects of welfare regimes as European countries move through a major depression. We have fitted a sequence of four models of growing complexity to assess how the size and nature of the effects change as predictor variables are included. Model 1 is

the empty or null one and has the specification of the schematic multilevel model above except that no predictor variables are included in the fixed part. It allows us to assess the correlation between the outcomes at each level and the extent and changing nature of between country differences. Model 2 additionally includes welfare regime and linear time interactions so that gives a societal growth curve for each outcome for each type of regime over the period. Model 3 additionally includes longitudinal and cross-sectional GDP so that we can assess the growth curves conditional on GDP as well as the effects of country wealth *per se*. The final Model 4 additionally includes age by sex interactions and individual decile income. The latter, because it is an ordinal variable, is fitted as a non-linear second-order orthogonal polynomial (Rasbash et al., 2009).

Results

Country differences over time

We observe large differences in SWB between countries and comparing the best with the worst, we find that citizens in Bulgaria have 23.5 higher odds of being Unhappy compared to citizens in Denmark. Ukrainians are 35.5 times more Dissatisfied than citizens in Denmark. While the differences for Health are smaller, they are still substantial with citizens from Ukraine reporting 10.8 higher odds of being Unhealthier compared to citizens in Ireland, the country with generally the best reported self-health. Another way of looking at these results is as a MOR statistic and the variances on the logit scale convert to MORs of 2.16, 2.28 and 1.75 for Happiness, Satisfaction and Health respectively which indicates quite substantial heterogeneity between counties with more than double the risk of being unhappy and unsatisfied on average between countries.

Figure 1 shows the results obtained from Model 1 as modelled country trends in each of the three outcomes on a logit scale where the value 0, shown by the dotted line, represents 50% of respondents saying ‘yes’ to an outcome with Ukraine and Bulgaria around this figure for Dissatisfaction. Examining the graphs more carefully, there are elements of stability over time especially for Health but there are changes too. These are shown more clearly in Figure 2 where only the country-wave differentials from the overall country differences are shown, again on a logit scale. The value of 0 on this graph represents the cross wave average for each country. The relative lack of volatility in the Health outcome is now more apparent, as is the changing levels of Happiness and Satisfaction. Ireland and Greece report the most change. Ireland experienced a severe recession over the period (Whelan et al., 2015) and the Greek government-debt crisis that began in 2009 is still ongoing (Matsaganis, 2012). With the unemployment rate reaching 27%, Greece has some of the highest levels of unemployment in the Eurozone.

<<Insert Figures 1 and 2 about here>>

We find the correlations between the three outcomes at each level are positive, there is a general tendency for all three types of well-being to go together. This is particularly the case for countries, where all three correlations are above 0.75. Countries whose population are Unhappy are also Dissatisfied and Unhealthy. At the individual level, Happiness and Satisfaction have a tendency to go together, but the correlations with both these variables and the Health outcome are weaker at around 0.2. Not all Unhealthy people are Dissatisfied and Unhappy. The correlations at the country-wave level show an interesting pattern. Differential short-term changes in a country's Happiness and Satisfaction go together, but the strength of the association is weaker with Health, particularly for Satisfaction and Health.

Summarising the results from Model 1 and answering the first two research questions, there are substantial between-country differences; the country where you live matters. While Health differences are unsurprisingly relatively unchanging in the short term, both Happiness and Satisfaction are changeable over countries even in the space of a few years. There is a generally tendency of these three aspects of well-being to go together and this is particularly marked at the country level. There are certainly sufficient differences between these European countries to try and account for these differences by including predictor variables. At the individual level the three outcomes represent somewhat different dimensions as the correlation between Health and the other two is not high (Happiness=0.21 and Satisfaction=0.20). The lower level correlations are net of the effect at higher levels so these individual correlations take into account differences between countries.

Trends experienced by different welfare regimes

Attention now focusses on the differences in SWB for the different welfare regimes. Figure 3 shows the time-invariant differences for each outcome. The plot shows the relative odds for each regime against the Social Democrat type which has been set at 1. The Conservative and Mediterranean country groups are somewhat similar, and the best outcomes are experienced by the Liberal and Social Democratic countries. The Post-Communist countries have the worst experience on all three outcomes. Citizens in Eastern Europe have 5 times the odds of reporting Unhappiness compared to people in the Nordic countries, they also have 6 times the odds of reporting Dissatisfaction and have 3.5 times the odds of being Unhealthier, compared to people in the Nordic countries. The differences may be greater for Satisfaction and smaller for Health but the underlying patterns are very similar. The 95% confidence bands may appear rather wide but it must be remembered that these are based (for each outcome) on estimating 5 differential logits based on only 30 countries. At the extremes and given the effective sample size there is evidence of significant differences here.

<<Insert Figure 3 about here>>

These graphs show the time-invariant result while Figure 4 shows the different experience of different regimes over time. This figure is derived from Model 2 with no other predictor variables in the model. The new feature here is that different regimes have undergone different trends. The most marked feature is that Post-Socialist states have improved in the 2000's but are still comparatively worse than the other Mediterranean states for each outcome. The Social Democrat regime is consistently the best and health and social well-being has not changed much during the Great Recession. The Liberal states have worsened over the decade and this is more noticeable for Happiness and Satisfaction. Similarly, Copeland et al. (2015) find Sweden's welfare state offers better population health protection during recent recessions, compared to the English welfare state. The Mediterranean states generally show high levels of problems. Overall, and due to the relative improvements in the Post-Communist states, differences have reduced over time.

<<Insert Figure 4 about here>>

While Figure 4 shows the trends for different regimes without any other predictors in the model, Figures 5 and 6 shows the results conditional on GDP, and additionally individual characteristics. The changes resulting from including GDP are substantial with the major differences between welfare regimes now being found for Health, while the differences for Happiness and Satisfaction for different regimes, given how rich a country is, are much smaller. The Nordic regime remains the best for Happiness and Satisfaction and also does comparatively very well for Health, but the experience of Liberal states have worsened considerably for Health and Happiness after taking account of national wealth. The Post-Communist counties are behaving more like our group of hybrid regimes except that their self-reported health is markedly worse albeit improving. When age, sex and individual income are additionally included there are some changes, but the basic picture stays the same. The differences between regimes are relatively small for Unhappiness and Dissatisfaction with the Nordic welfare model being consistently the best; Liberal countries have become slightly worse during the decade, but the overall differences are not very substantial. The differences for Health however remain substantial with those countries with a Liberal experiencing the lowest rates. The Eastern European countries are substantially worse although they have shown improvements during the decade. Inequality between welfare regimes has declined but remains substantial.

<<Insert Figures 5 and 6 about here>>

In summary, the story is a complicated one; and what you take into account matters for comparing welfare states. Without conditioning on any variables, the Health outcome has the smallest differences between regimes but when you take account of how rich a country is, the largest differences are for Health. The Social Democratic regime consistently does well across

all three outcomes. The biggest differences are for Health where, when individual characteristics are taken into account, the experience of the Post-Communist group remains considerably worse than any other regime.

Effects of longitudinal and cross-sectional GDP

Given the observed changes in the effects of welfare regime as GDP is included, we can expect relationships between the outcomes and this measure of a country's wealth. Moreover we can expect a differential relationship as the Health differences between regimes were relatively unaffected by this conditioning. Figure 7 and Figure 8 shows the cross-sectional effect and longitudinal effect respectively and this is net of welfare regime and of age, sex and individual income. The graphs show the effect between the 5 and 95% percentile of both variables as relative odds with 95% confidence intervals.

<<Insert Figures 7 and 8 about here>>

Long-term GDP has a powerful effect for both Happiness and Satisfaction and people in the poorest countries are markedly more Unhappy and Dissatisfied than the richest. And this is of course an ecological effect even after taking account of individual income. But the Health of a country is effectively unrelated to long-term GDP – it matters little how rich a country is; and now we understand why the welfare regime differences did not reduce when this variable was included. For longitudinal GDP the effects are much smaller and there is a great deal of uncertainty around the relationship. This must be seen in the context of having only at most 5 waves on which to estimate the relationship on, over a relatively short period of a decade, albeit a volatile one. All the effects are such that as GDP reduces from its long term mean (the negative values on the horizontal axis) there is more Unhappiness, Dissatisfaction and Unhealth. The difference between a country that is some 11 below its mean and 14 above it, is small; even for the strongest effect for Health is only 20% higher in terms of relative odds. Short-term change may have some effect but it is enduring country wealth that is more important for Happiness and Satisfaction.

Effects of individual variables

The final sets of fixed-part results pertain to the individual factors of age, sex and income. Age was included in the model as a second-order polynomial and in interaction with sex, and all the results are shown net of welfare regime and GDP; that is they are based on Model 4. Figure 9 shows the key results for the demographic variables, note that a different scale has been used for the vertical axis for Health in comparison to the other two outcomes. It is the young that are Happy, Satisfied and Healthy with (unsurprisingly) very big differences in the relative odds

for being Healthy. The gender interactions are such that differences between the sexes are more marked at older ages when it is consistently women who are the more Unhappy, Dissatisfied and Unhealthier. We cannot tell from these data whether this is an age or cohort effect as the individuals are only measured at one point in time (Bell and Jones, 2014).

<<Insert Figure 9 about here>>

The effects for income as shown in Figure 10 are remarkably consistent; indeed the correlation for each pair of relative odds is over 0.99. It is the poor who suffer the worst outcome for each of the responses. The 'shape' of the relationship is such (and a third order polynomial showed no change and was not significant at the 95% level) that the effect tapers off at the higher deciles of income. This gives a great deal of scope for income redistribution as a reduction in the topmost incomes would bring little change to the outcomes, but an increase at the bottom end would lead to substantial improvement. The effects for the extremes of individual income are smaller than that for cross-sectional GDP for both Happiness and Satisfaction, but the individual effects on Health are more substantial for individual income, but with these data we cannot rule out reverse causation.

<<Insert Figure 10 about here>>

Changes in unexplained variation

The final graphs in Figure 11 show what happens to the unexplained variation when predictor variables at the different levels are included in the models. The results make intuitive sense in the light of what we have seen.[4] The unexplained differences between waves are relatively small, particularly for Health and these do not change a great deal as regime, GDP, and individual income, age and sex are included. The between-country differences are substantially larger especially for Happiness and Satisfaction. The variances for these two responses decrease as welfare regime, and GDP are included but change little when individual variables are included. The between country differences in Health are smaller and reduce only when welfare regime is taken into account and not when GDP and the individual variables are included. For all three responses there is a 75% reduction in the country variance as we move from Model 1 to 4, so quite a lot of the differences between countries has been accounted for by a combination of welfare regime and GDP.

<<Insert Figure 11 about here>>

Discussion and Conclusion

In this article we focus specifically on the analysis and the methodological challenges and issued faced when making international comparisons between economic growth, welfare and well-being across nations and time. We sought to examine how changes in contextual and individual income (spanning the Great Recession) are associated with changes in self-reported well-being in the European Social Survey (ESS) 2002-2011. A multivariate multilevel analysis was performed on 237,253 individuals nested within 128 country cohorts covering 30 countries. Our results suggest that distributed income at the individual and country-level matters for good health and positive well-being. Long-term GDP has a powerful effect for both Happiness and Satisfaction and people in the poorest countries are markedly more Unhappy and Dissatisfied than the richest. Moreover this is an ecological effect after taking account of individual income. Individual income matters for all three outcomes, especially for the relatively poor. Long-term GDP matters less for Health, established welfare state counts – although in practice these macro factors may be related, as Wilensky (1975) famously observed.

Europe is clearly divided and the share of well-being is far from being evenly distributed or equal since we observe stark inequalities in the health and well-being of European citizens, between nations and also between the established welfare states of Europe and the weaker welfare states of Southern Europe and those emerging in Post-Communist Europe. The findings and observations presented here are, to a large extent, inherited from the past, a result of the great divergence – and continuing legacy – in the economic and political development of the European societies (Castles, 1995). The need for social protection in parts of Southern Europe is now much greater than ever before (Matsaganis, 2011, 2012). In the face of growing levels of inequality (OECD, 2008, 2011), redistributive polices are typically justified by social justice arguments (Kangas, 2000). Our results should be interpreted to indicate some directions in which policymakers could proceed when trying to incorporate deliberations of social justice with empirical evidence considering subjective well-being as on an outcome (cf. Chapple, 2011). We observe that modest income redistributions, from the top-end of the income distribution to the bottom-end would lead to substantial improvements in well-being in many countries without disadvantaging the well-being of the more wealthy sections of society.

Our findings present major challenges for European policymakers interested in a more equitable distribution of health and well-being; but as Marmot et al. observe, Euope does not need to be so divided (Marmot et al., 2012). Social policies that ensure economic growth is met with social investment and equity are required; implicated are social investment strategies that seek to regenerate welfare states, promote social inclusion and work, and help address the problems posed by economic restructuring, globalisation, demographics of population ageing and climate change (Morel et al., 2012). The European Commission's 2020 strategy calls for inclusive growth and the strengthening of social protection systems (European Commission, 2014). The scale of the challenge facing policymakers is considerable in the face of such diversity. At present, we find significant variations in the institutional setup, conditionality and generosity of European social protection schemes (Marx and Nelson, 2012). The issues here are inherently complex, not least because the EU Commission cannot stipulate specific social

protection policies. As a consequence, it has been considering whether it is desirable or even possible to set minimum social protection standards across EU Member States to help bridge social divides. A cooperative Union may help guide the development of national welfare states, while leaving the ways and means of social policy to Member States. However, there is much uncertainty as to whether a more binding policy framework on minimum standards can, in fact, raise the quality and efficiency of domestic welfare systems to ensure a fairer and more equitable distribution of social well-being (Chapple, 2011). Some non-EU countries like Ukraine (reporting the lowest level of life satisfaction in Europe) face particular challenges. With the country in crisis, the EU agreed an €11 billion financial package in 2014 to help stabilise the economy and democratic institutions of civil society, as it makes further preparations for EU membership at a time when the EU's resolve to enlarge appears to be waning. Finally these differences and changes need to be considered in a context where there are strong political forces in many countries arguing for lowering the social budget, with less redistribution, as part of an austerity package. But as sophisticated modelling of growth and debt shows it is lack of growth that leads to debt and not vice-versa; the clamour for austerity rests on a fragile evidential base (Bell et al., 2015).

Notes

[1]The facility to model unbalanced data is based on the rather undemanding 'Missing at Random' assumption (Little and Rubin, 2002) so that the missingness itself is un-informative and the non-response process depends only on observed but not on unobserved variables. In particular it is assumed that the probability of missing does not depend on responses we would have observed had they not been missing.

[2]Again this is based on the MAR assumption, and we will not get biased results providing that there is not something about the outcomes that has resulted in a particular survey wave in a country not being undertaken.

[3]Estimation involves the inversion of matrices of this order in an iterative fashion until convergence is achieved. Fairbrother (2014) used maximum likelihood with a Laplace approximation in his study of trust and post-materialism in the World Values Survey (WVS). However, he did not use a Bernoulli model at the individual level but employed the binomial method of Subramanian et al. (2001) thereby aggregating over individuals. This was appropriate for his study, but not for ours, because no level 1 predictor variables were specified and therefore the modelling is based on the much smaller dataset of the proportions aggregated to waves and countries. He was therefore able to use the more computationally demanding likelihood methods whereas we have used the computationally efficient quasi- likelihood ones in their least biased PQL-2nd order form (Goldstein and Rasbash, 1996).

[4]It is perfectly possible for the higher-level variance to increase as lower level predictors are included (Jones, 1992). This is particularly the case when dealing with Bernoulli models as the level 1 variance cannot reduce as predictor variables are included (Jones and Subramanian, 2013). This does not happen here except for a very small increase with Model 4.

References

Angrist, J. D. and J.-S. Pischke (2009) *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ and Oxford: Princeton University Press.

Arcanjo, M. (2011) 'Welfare State Regimes and Reforms: A Classification of Ten European Countries between 1990 and 2006', *Social Policy and Society* 10: 139-150.

Bambra, C., G. Netuveli and T. A. Eikemo (2010) 'Welfare State Regime Life Courses: The Development of Western European Welfare State Regimes and Age-Related Patterns of Educational Inequalities in Self-Reported Health', *International Journal of Health Services* 40: 399-420.

Barr, N. (2012) *Economics of the Welfare State*. 5th edition. Oxford: Oxford University Press.

Beckfield, J. and N. Krieger (2009) 'Epi + demos + cracy: Linking Political Systems and Priorities to the Magnitude of Health Inequities—Evidence, Gaps, and a Research Agenda', *Epidemiologic Reviews* 31: 152-177.

Bell, A., R. Johnston and K. Jones (2015) 'Stylised fact or situated messiness? The diverse effects of increasing debt on national economic growth', *Journal of Economic Geography* 15: 449-472.

Bell, A. and K. Jones (2014) 'Another 'futile quest'? A simulation study of Yang and Land's Hierarchical Age-Period-Cohort model', *Demographic Research* 30: 333-360.

Bell, A. and K. Jones (2015) 'Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data', *Political Science Research and Methods* 3: 133-153.

Borrell, C., A. Espelt, M. Rodriguez-Sanz, B. Burstrom, C. Muntaner, M. I. Pasarin, J. Benach, C. Marinacci, A. J. Roskam, M. Schaap, E. Regidor, G. Costa, P. Santana, P. Deboosere, A. Kunst and V. Navarro (2009) 'Analyzing differences in the magnitude of socioeconomic inequalities in self-perceived health by countries of different political tradition in Europe', *International Journal of Health Services* 39: 321-341.

Bowling, A. (2004) *Measuring Health: A Review of Quality of Life Measurement Scales*. 3rd Edition. Buckingham: Open University Press.

Boyd, L. and G. Iversen (1979) *Contextual Analysis: Concepts and Statistical Techniques*. Belmont, CA: Wadsworth.

Canberra Group (2011) *Handbook on Household Income Statistics*. Second Edition. Geneva: United Nations.

Castles, F. G. (1995) 'Welfare state development in Southern Europe', *West European Politics* 18: 291-313.

Chapple, S., (ed.) (2011) *Subjective well-being and social policy*. Luxembourg, European Commission.

Clark, A. E., P. Frijters and M. A. Shields (2008) 'Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles', *Journal of Economic Literature* 46: 95-144.

Copeland, A., C. Bambra, L. Nylén, A. Kasim, M. Riva, S. Curtis and B. Burström (2015) 'All in It Together? The Effects of Recession on Population Health and Health Inequalities in England and Sweden, 1991–2010', *International Journal of Health Services* 45: 3-24.

Dahl, E., J. Fritzell, E. Lahelma, P. Martikainen, A. Kunst and J. P. Mackenbach (2006) 'Welfare state regimes and health inequalities', pp. 193-222 in J. Siegrist and M. Marmot (eds.) *Social Inequalities in Health: New evidence and policy implications*. Oxford: Oxford University Press.

Deaton, A. (2008) 'Income, Health, and Well-Being around the World: Evidence from the Gallup World Poll', *Journal of Economic Perspectives* 22: 53-72.

Deeming, C. (2013) *Addressing the Social Determinants of Subjective Wellbeing: Evidence from the ONS Opinions Survey*. London: Brettenham House.

Deeming, C. and D. Hayes (2012) 'Worlds of Welfare Capitalism and Wellbeing: A Multilevel Analysis', *Journal of Social Policy* 41: 811-829.

Diener, E., (ed.) (2009) *Assessing Well-Being*. Dordrecht, Springer.

Duncan, C., K. Jones and G. Moon (1998) 'Context, composition and heterogeneity: Using multilevel models in health research', *Social Science & Medicine* 46: 97-117.

Easterlin, R. A. (1974) 'Does Economic Growth Improve the Human Lot? Some Empirical Evidence', pp. 89-125 in P. A. David and M. W. Reder (eds.) *Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz*. New York, NY: Academic Press.

Esping-Andersen, G. (1990) *The Three Worlds of Welfare Capitalism*: Cambridge: Polity Press and Princeton, NJ: Princeton University Press.

Esping-Andersen, G. (1999) *Social Foundations of Post-industrial Economies*. Oxford: Oxford University Press.

European Commission (2014) *Taking stock of the Europe 2020 strategy for smart, sustainable and inclusive growth*. Brussels: European Commission.

Fairbrother, M. (2014) 'Two Multilevel Modeling Techniques for Analyzing Comparative Longitudinal Survey Datasets', *Political Science Research and Methods* 2: 119-140.

Fenger, H. J. M. (2007) 'Welfare regimes in Central and Eastern Europe: Incorporating post-communist countries in a welfare regime typology', *Contemporary Issues and Ideas in Social Sciences* 3: http://journal.ciiss.net/index.php/ciiss/article/viewFile/45/37.

Ferragina, E. and M. Seeleib-Kaiser (2011) 'Thematic Review: Welfare regime debate: past, present, futures?', *Policy & Politics* 39: 583-603.

Ferrera, M. (1996) 'The 'Southern Model' of Welfare in Social Europe', *Journal of European Social Policy* 6: 17-37.

Ferrera, M., (ed.) (2005) *Welfare State Reform in Southern Europe: Fighting Poverty and Social Exclusion in Greece, Italy, Spain and Portugal*. London and New York, NY, Routledge.

Goldstein, H. and J. Rasbash (1992) 'Efficient computational procedures for the estimation of parameters in multilevel models based on iterative generalised least squares', *Computational Statistics & Data Analysis* 13: 63-71.

Goldstein, H. and J. Rasbash (1996) 'Improved Approximations for Multilevel Models with Binary Responses', *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 159: 505-513.

Graham, C., S. Chattopadhyay and M. Picon (2010) 'The Easterlin and Other Paradoxes: Why Both Sides of the Debate May Be Correct', pp. 247-290 in E. Diener, D. Kahneman and J. Helliwell (eds.) *International Differences in Well-Being*. Oxford: Oxford University Press.

Hay, C. and D. Wincott (2012) *The Political Economy of European Welfare Capitalism*. Basingstoke: Palgrave Macmillan.

Hoffmeyer-Zlotnik, J. H. P. and U. Warner (2006) 'Methodological Discussion of the Income Measure in the European Social Survey Round 1', *Metodološki zvezki* 3: 289-334.

Jen, M. H., K. Jones and R. Johnston (2009) 'Global variations in health: Evaluating Wilkinson's income inequality hypothesis using the World Values Survey', *Social Science & Medicine* 68: 643-653.

Jones, K. (1990) *Problems of aggregate data: a critique and review of selected studies on educational attainment*. Bristol: Centre for Multilevel Modelling, University of Bristol.

Jones, K. (1992) 'Amendment to 'Specifying and Estimating Multi-Level Models for Geographical Research' by Kelvyn Jones (1991) Trans. Inst. Br. Geogr. 16: 148-160', *Transactions of the Institute of British Geographers* 17: 246.

Jones, K. and N. Bullen (1994) 'Contextual Models of Urban House Prices: A Comparison of Fixed- and Random-Coefficient Models Developed by Expansion', *Economic Geography* 70: 252-272.

Jones, K., D. Owen, R. Johnston, J. Forrest and D. Manley (2014) 'Modelling the occupational assimilation of immigrants by ancestry, age group and generational differences in Australia: a random effects approach to a large table of counts', *Quality & Quantity* 10.1007/s11135-014-0130-8: 1-21.

Jones, K. and V. S. Subramanian (2013) *Developing multilevel models for analysing contextuality, heterogeneity and change*. Bristol: Centre for Multilevel Modelling, University of Bristol.

Kangas, O. (2000) 'Distributive Justice and Social Policy: Some Reflections on Rawls and Income Distribution', *Social Policy and Administration* 34: 510-528.

Kvist, J., J. Fritzell, B. Hvinden and O. Kangas, (eds.) (2011) *Changing Social Equality: The Nordic Welfare Model in the 21st Century*. Bristol, Policy Press.

Larsen, K. and J. Merlo (2005) 'Appropriate Assessment of Neighborhood Effects on Individual Health: Integrating Random and Fixed Effects in Multilevel Logistic Regression', *American Journal of Epidemiology* 161: 81-88.

Little, R. J. A. and D. B. Rubin (2002) *Statistical Analysis with Missing Data*. 2nd Edition. New York, NY: John Wiley.

Marmot, M. (2002) 'The Influence Of Income On Health: Views Of An Epidemiologist', *Health Affairs* 21: 31-46.

Marmot, M., J. Allen, R. Bell, E. Bloomer and P. Goldblatt (2012) 'WHO European review of social determinants of health and the health divide', *The Lancet* 380: 1011-1029.

Marx, I. and K. Nelson, (eds.) (2012) *Minimum Income Protection in Flux*. Basingstoke, Palgrave Macmillan.

Matsaganis, M. (2011) 'The welfare state and the crisis: the case of Greece', *Journal of European Social Policy* 21: 501-512.

Matsaganis, M. (2012) 'Social policy in hard times: The case of Greece', *Critical Social Policy* 32: 406-421.

Morel, N., B. Palier and J. Palme, (eds.) (2012) *Towards a social investment welfare state? Ideas, policies and challenges*. Bristol, Policy Press.

Obinger, H., P. Starke, J. Moser, C. Bogedan, E. Gindulis and S. Leibfried (2010) *Transformations of the Welfare State: Small States, Big Lessons*. Oxford: Oxford University Press.

OECD (2008) *Growing Unequal? Income Distribution and Poverty in OECD Countries*. Paris: OECD Publishing.

OECD (2011) *Divided we stand: why inequality keeps rising*. Paris: OECD Publishing.

OECD (2013) *OECD Guidelines on Measuring Subjective Well-being*: OECD Publishing. http://dx.doi.org/10.1787/9789264191655-en

Ovaska, T. and R. Takashima (2006) 'Economic policy and the level of self-perceived well-being: An international comparison', *The Journal of Socio-Economics* 35: 308-325.

Pacek, A. and B. Radcliff (2008) 'Assessing the Welfare State: The Politics of Happiness', *Perspectives on Politics* 6: 267-277.

Povey, J., P. Boreham and W. Tomaszewski (2013) 'The development of a new multi-faceted model of social wellbeing: Does income level make a difference?', *Journal of Sociology* Published online before print, doi: 10.1177/1440783313507491.

Rasbash, J., C. Charlton, K. Jones and R. Pillinger (2009) *Manual Supplement to MLwiN v2.1*. Bristol: Centre for Multilevel Modelling, University of Bristol.

Rodriguez, G. (2008) 'Multilevel Generalized Linear Models', pp. 335-376 in J. de Leeuw and E. Meijer (eds.) *Handbook of Multilevel Analysis*. Dordrecht: Springer: data and papers available from http://data.princeton.edu/multilevel/.

Rodriguez, G. and N. Goldman (1995) 'An Assessment of Estimation Procedures for Multilevel Models with Binary Responses', *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 158: 73-89.

Stegmueller, D. (2013) 'How Many Countries for Multilevel Modeling? A Comparison of Frequentist and Bayesian Approaches', *American Journal of Political Science* 57: 748-761.

Stevenson, B. and J. Wolfers (2008) 'Economic growth and subjective well-being: reassessing the Easterlin paradox', *Brookings Papers on Economic Activity* Spring: 1-102 http://www.nber.org/papers/w14282.pdf?new_window=1.

Stiglitz, J., A. Sen and J.-P. Fitoussi (2010) *MIS-Measuring Our Lives: Why GDP Doesn't Add Up*. New York, NY: The New Press.

Subramanian, S. V., C. Duncan and K. Jones (2001) 'Multilevel perspectives on modeling census data', *Environment and Planning A* 33: 399-417.

Subramanian, S. V., K. Jones, A. Kaddour and N. Krieger (2009) 'Revisiting Robinson: The perils of individualistic and ecologic fallacy', *International Journal of Epidemiology* 38: 342-360.

Subramanian, S. V. and I. Kawachi (2004) 'Income Inequality and Health: What Have We Learned So Far?', *Epidemiologic Reviews* 26: 78-91.

Sudhir, A. and A. Sen (1994) *Human development Index: Methodology and Measurement*. New York: United Nations Development Programme.

Welzel, C. and R. Inglehart (2010) 'Agency, Values, and Well-Being: A Human Development Model', *Social Indicators Research* 97: 43-63.

Whelan, C., H. Russell and B. Maître (2015) 'Economic Stress and the Great Recession in Ireland: Polarization, Individualization or 'Middle Class Squeeze'?', *Social Indicators Research*: 1-24.

Wilensky, H. L. (1975) *The Welfare State and Equality: Structural and Ideological Roots of Public Expenditures*. Berkley: University of California Press.

Wilkinson, R. and K. Pickett (2009) *The Spirit Level: Why More Equal Societies Almost Always Do Better*. London: Allen Lane.

Appendices

Appendix 1: European Social Survey (ESS)

The European Social Survey (ESS) is a biennial multi-country survey covering over 30 nations. The first round was fielded in 2002/2003, the fifth in 2010/2011. The project is funded jointly by the European Commission, the European Science Foundation and academic funding bodies in each participating country, and is designed and carried out to exceptionally high standards. The project is directed by a Core Scientific Team at the Centre for Comparative Social Surveys, City University, London. The questionnaire includes two main sections, each consisting of approximately 120 items; a 'core' module which remains relatively constant from round to round, plus two or more 'rotating' modules, repeated at intervals. The core module aims to monitor change and continuity in a wide range of social variables, including health and well-being; social and public trust; political interest and participation; socio-political orientations; governance and efficacy; moral; political and social values; social exclusion, national, ethnic and religious allegiances; human values; demographics and socio-

economics. The ESS collects a wide range of methodological data, including tests of reliability, call records, data on interview settings and event data. Upon registration the ESS data are available free of charge and without restrictions, for not-for-profit and academic research purposes.

Appendix 2: Treatment of ESS income in Rounds 1-3

The first three rounds of the ESS were based on respondents replying to the same categories of household income across all countries, the income codes from the survey are shown below.

CARD 56

YOUR HOUSEHOLD INCOME

	Approximate WEEKLY	Approximate MONTHLY	Approximate ANNUAL	
J	Less than €40	Less than €150	Less than €1800	J
R	€40 to under €70	€150 to under €300	€1800 to under €3600	R
C	€70 to under €120	€300 to under €500	€3600 to under €6000	C
M	€120 to under €230	€500 to under €1000	€6000 to under €12000	M
F	€230 to under €350	€1000 to under €1500	€12000 to under €18000	F
S	€350 to under €460	€1500 to under €2000	€18000 to under €24000	S
K	€460 to under €580	€2000 to under €2500	€24000 to under €30000	K
P	€580 to under €690	€2500 to under €3000	€30000 to under €36000	P
D	€690 to under €1150	€3000 to under €5000	€36000 to under €60000	D
H	€1150 to under €1730	€5000 to under €7500	€60000 to under €90000	H
U	€1730 to under €2310	€7500 to under €10000	€90000 to under €120000	U
N	€2310 or more	€10000 or more	€120000 or more	N

In the next two rounds, country specific income groups were used instead, and respondents had to place themselves in the income decile groups specific to that country, the revised income codes are shown below.

1 [national currency] = x.xx €.

The intervals of the categories in CARD 53 are:

CARD 53

YOUR HOUSEHOLD INCOME

Showcard code	Approximate WEEKLY	Approximate MONTHLY	Approximate ANNUAL	Data file code
J	Less than €xx	Less than €xxx	Less than €xxx	01
R	€xx to under €xx	€xxx to under €xxx	€xxx to under €xxx	02
C	€xx to under €xx	€xxx to under €xxx	€xxx to under €xxx	03
M	€xx to under €xx	€xxx to under €xxx	€xxx to under €xxx	04
F	€xx to under €xx	€xxx to under €xxx	€xxx to under €xxx	05
S	€xx to under €xx	€xxx to under €xxx	€xxx to under €xxx	06
K	€xx to under €xx	€xxx to under €xxx	€xxx to under €xxx	07
P	€xx to under €xx	€xxx to under €xxx	€xxx to under €xxx	08
D	€xx to under €xx	€xxx to under €xxx	€xxx to under €xxx	09
H	More than €xx	More than €xxx	More than €xxx	10

For the analysis we have created equivalent decile groups for the earlier rounds by a three stage procedure. For each across country Euro income group, we have generated individual income data according to uniform distribution between the lowest and highest value of the group. At the second stage, these values are sorted according to the income stress variable (HINCFEL). At stage three and working with a country-wave at a time, country-specific decile groups were created. Table 3, for instance, shows the distribution of income in the Netherlands in 2006 recast as decile income groups for the Netherlands in 2006. There are 167/166 people in each of the constructed 10 decile groups; decile 1 has to be extended from €<1.8k to €12k Euros to get enough people in the bottom decile as the Netherlands is a high-income country. The people in the 6-12 band have to be split between decile 1 and 2 and we have used the income stress variable as auxiliary information in determining this split.

<<Insert Table 3 about here>>

Table 4 below gives the cut-offs for the population deciles for the Netherlands in 2005 based on over 7 million households compared to the group based on our procedure. The approximate nature of the procedure is apparent. Generally the accordance is good except for the highest deciles which given the skewness of income data is to be expected in a sample survey. Nevertheless because we modelling income as an orthogonal polynomial, this should not be too huge a problem as an overall smooth is being imposed across the groups and ordinality has been preserved by our method.

<<Insert Table 4 about here>>

The imputed cut-offs for all country-waves are given in Table 5.

<<Insert Table 5 about here>>

Appendix 3: GDP per capita 2002-2010

Our GDP per capita data are available from the World Bank website (http://data.worldbank.org/indicator/NY.GDP.PCAP.CD (consulted June 2015)). GDP per capita is gross domestic product divided by midyear population. GDP is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources. Data in are in current U.S. dollars.

Table 1. ESS country-wave sample and the political economy of European welfare capitalism

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Totals
Northern European/'Nordic'						
Denmark (DK)	1,506	1,487	1,505	1,610	1,576	7,684
Norway (NO)	2,036	1,760	1,750	1,549	1,548	8,643
Finland (FI)	2,000	2,022	1,896	2,195	1,878	9,991
Sweden (SE)	1,999	1,948	1,927	1,830	1,497	9,201
Continental European/'Corporatist'						
Austria (AT)	2,257	2,256	2,405	-	-	6,918
Belgium (BE)	1,899	1,778	1,798	1,760	1,704	8,939
France (FR)	1,503	1,806	1,986	2,073	1,728	9,096
Germany (DE)	2,919	2,870	2,916	2,751	3,031	14,487
Luxembourg (LU)	1,552	1,635	-	-	-	3,187
Netherlands (NL)	2,364	1,881	1,889	1,778	1,829	9,741
Switzerland (CH)	2,040	2,141	1,804	1,819	1,506	9,310
Liberal						
Ireland (IE)	2,046	2,286	1,800	1,764	2,576	10,472
United Kingdom (GB)	2,052	1,897	2,394	2,352	2,422	11,117
Eastern European/'Post-Communist'						
Bulgaria (BG)	-	-	1,400	2,230	2,434	6,064
Czech Republic (CZ)	1,360	3,026	-	2,018	2,386	8,790
Croatia (HR)	-	-	-	1,484	1,649	3,133
Estonia (EE)	-	1,989	1,517	1,661	1,793	6,960
Hungary (HU)	1,685	1,498	1,518	1,544	1,561	7,806
Poland (PL)	2,110	1,716	1,721	1,619	1,751	8,917
Russia (RU)	-	-	2,437	2,512	2,595	7,544
Slovakia (SK)	-	1,512	1,766	1,810	1,856	6,944
Slovenia (SI)	1,519	1,442	1,476	1,286	1,403	7,126
Ukraine (UA)	-	2,031	2,002	1,845	1,931	7,809
Southern European/'Mediterranean'						
Cyprus (CY)	-	-	995	1,215	1,083	3,293
Greece (GR)	2,566	2,406	-	2,072	2,715	9,759
Italy (IT)	1,207	1,529	-	-	-	2,736
Israel (IL)	2,499	-	-	2,490	2,294	7,283
Portugal (PT)	1,511	2,052	2,222	2,367	2,150	10,302
Spain (ES)	1,729	1,663	1,876	2,576	1,885	9,729
Turkey (TR)	-	1,856	-	2,416	-	4,272
Totals	42,359	48,487	43,000	52,626	50,781	237,253

Source: Adapted from Hay and Wincott (2012) and ESS Waves I-V.

Table 2. Independent variables, ESS questions, and valid responses

Variable	Variable label	Monitoring question	Response counts (percent)		
health	Subjective general health.	How is your health in general? Would you say it is... ...very good=1, good=2, fair=3, bad=4, or very bad=5.	Un- 21,788 (9.2)	Healthy 215,121 (90.7)	Missing 348 (0.1)
stflife	How satisfied with life as a whole.	All things considered, how satisfied are you with your life as a whole nowadays? Where 0 means extremely dissatisfied and 10 means extremely satisfied.	Dis- 36,620 (15.4)	Satisfied 199,205 (84.0)	Missing 1,428 (0.6)
happy	How happy are you.	Taking all things together, how happy would you say you are?	Un- 23,546 (9.9)	Happy 211,965 (89.3)	Missing 1,742 (0.7)

Source: ESS Waves I-V.

Table 3. Model and data specifications

Variables	Description	Specification in the study	Source
stflife	How satisfied with life as a whole	Binary coded dependent variable: Dis-Satisfied.	ESS R1-5 (see Table 2).
happy	How happy are you with life as a whole	Binary coded dependent variable: Un-Happy.	ESS R1-5 (see table 2).
health	Subjective general health	Binary coded dependent variable: Un-Healthy.	ESS R1-5 (see Table 2).
cntry	Country	30 European countries across the five ESS waves, table 1.	ESS R1-5.
essround	Round	SWB may vary over time. We consider all five rounds of the ESS: 2002-3, 2004-5, 2006-7, 2008-9 and 2010-11.	ESS R1-5.
agea	Age of respondent, calculated	Individual-level independent variable age may help to explain well-being in European populations.	ESS R1-5.
gndr	Gender	Gender may influence health and well-being. In the ESS this is a binary male/female question.	ESS R1-5.
Hinctnt/ Hinctnta	Household's total net income, all sources	Distributed household income may help to explain health and well-being. ESS records income deciles.	ESS R1-5.[1]
GDP per capita	Gross domestic product (GDP) per head.	Country-wealth may influence population well-being. Here we use the country-level variable Gross Domestic Product (GDP) per head in US dollars, for each ESS Round.	The World Bank 2000-10.[2]
Unemployment	Unemployment rate	Per cent of total labor force. Unemployment refers to the share of the labor force that is without work but available for and seeking employment.	The World Bank 2000-10
Welfare	Welfare regime	Hay and Wincott's (2012) framework of European welfare capitalism.	Table 1.

Source: ESS Waves I-V.

Table 4. Correlations between the three outcomes for Countries, Waves and Individuals

		Correlations		
		Happiness	Satisfaction	Heath
Countries	Happiness	1.00	-	-
	Satisfaction	0.96	1.00	-
	Health	0.76	0.80	1.00
Country-Wave	Happiness	1.00	-	-
	Satisfaction	0.86	1.00	-
	Health	0.36	0.20	1.00
Individual	Happiness	1.00	-	-
	Satisfaction	0.50	1.00	-
	Health	0.21	0.20	1.00

Table 5. The relative odds of being well across European welfare regimes

		Relative Odds	
	Un-Happy	Dis-Satisfaction	Un-Healthy
Post-Communist	5.40	6.34	3.48
Mediterranean	2.91	3.69	2.23
Continental Europe	2.35	2.75	1.67
Liberal	1.53	1.60	0.86
Nordic	1.00	1.00	1.00
		Table correlations	
Happiness	1.00		
Satisfaction	0.98	1.00	
Health	0.97	0.99	1.00

Figure 1. Modelled country-outcomes shown as a matrix plot of relative odds

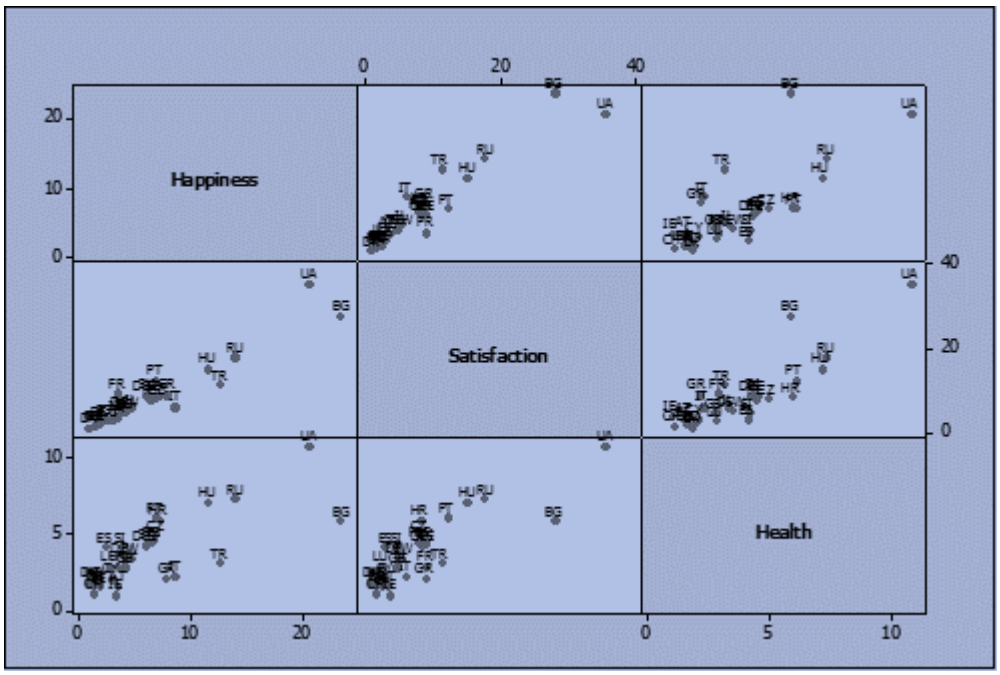

Figure 2. Modelled country-wave differentials with the three outcomes on a logit scale

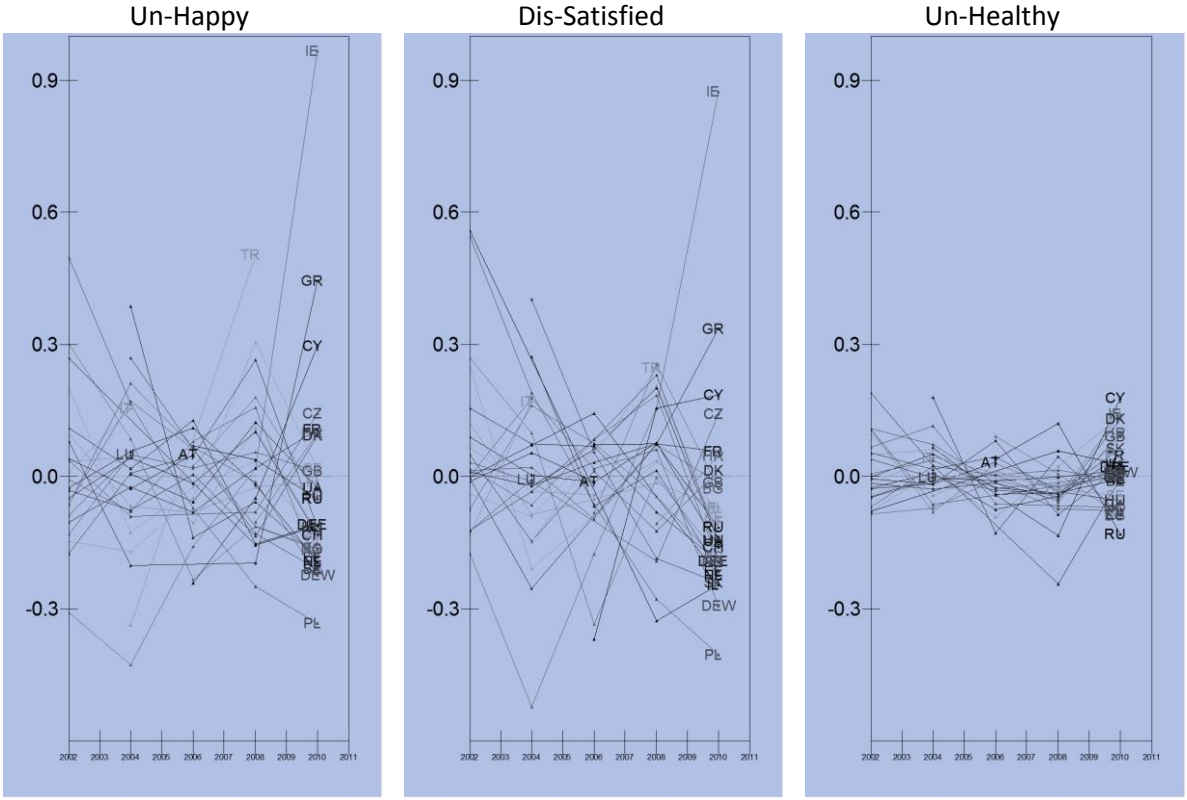

Figure 3. Differences in well-being across European welfare regimes: relative odds and 95% CI's

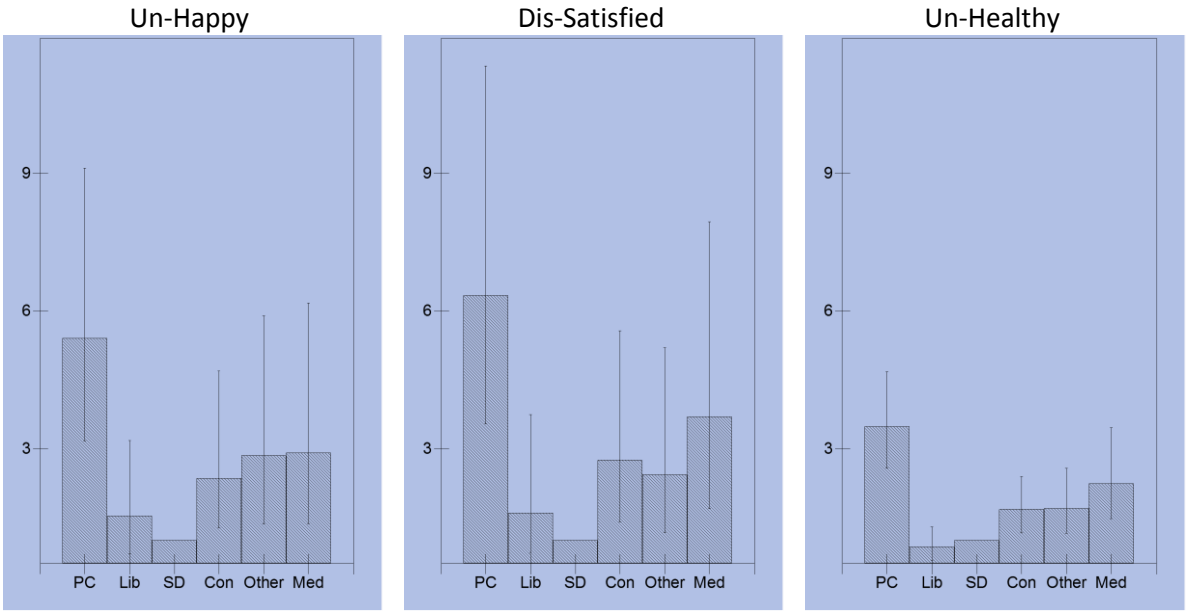

Figure 4. Trends in well-being across European welfare regimes: relative odds based on Model 2 with no other predictors

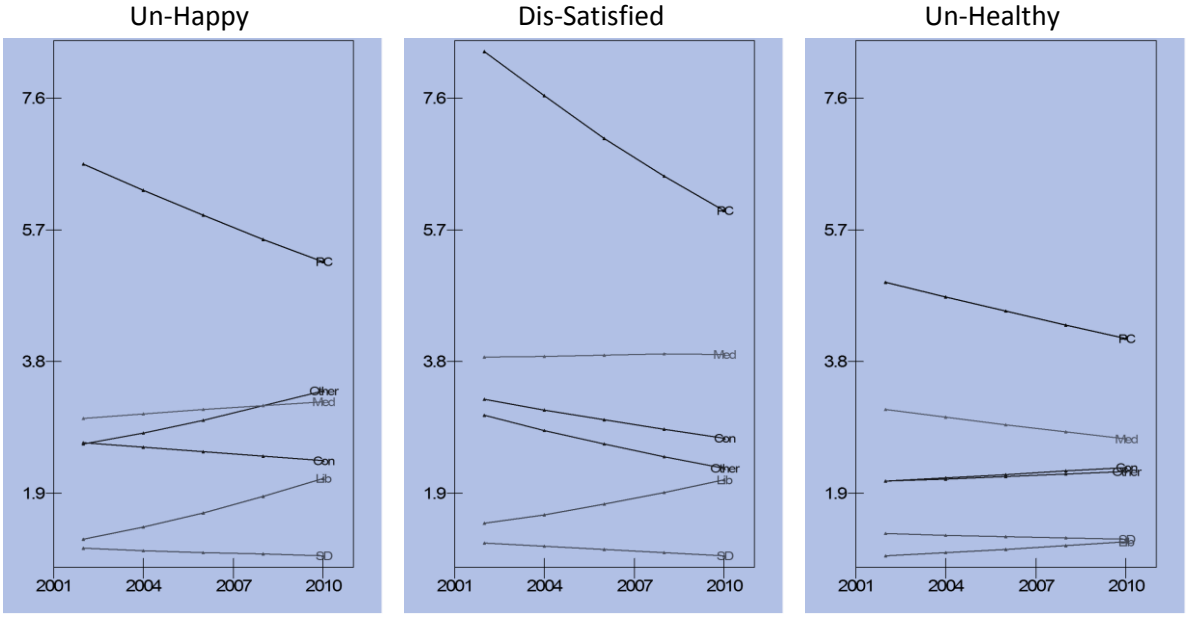

Figure 5. Trends in well-being across European welfare regimes: relative odds based on Model 3 including GDP

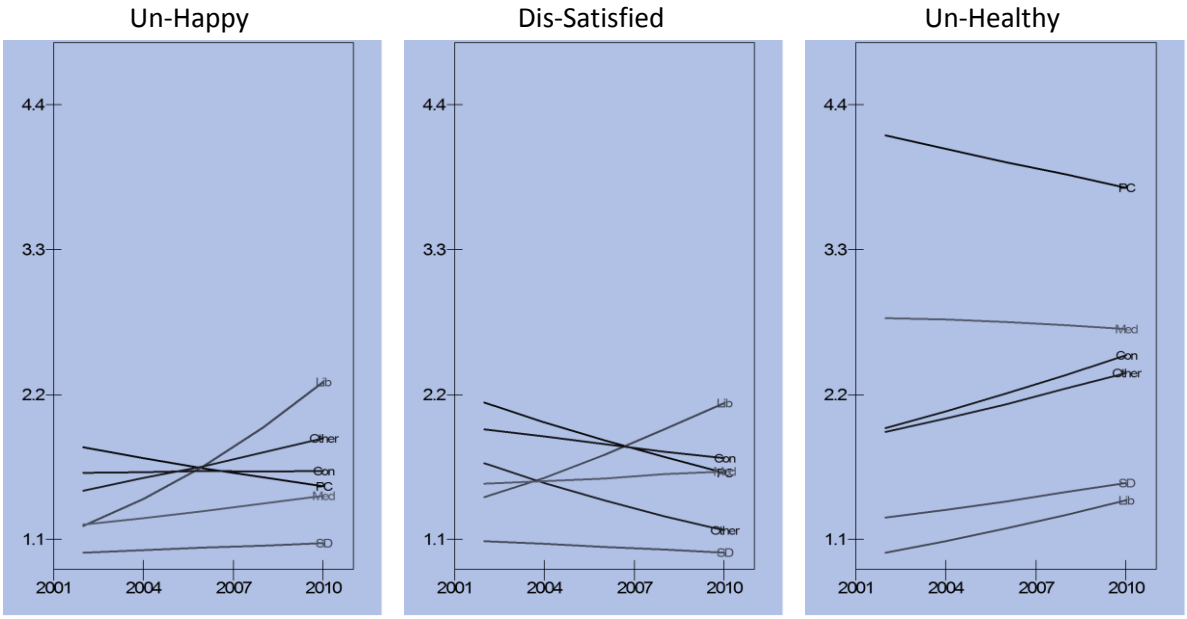

Figure 6. Trends in well-being across European welfare regimes: relative odds based on Model 4 including Age, Sex and Individual Income with GDP

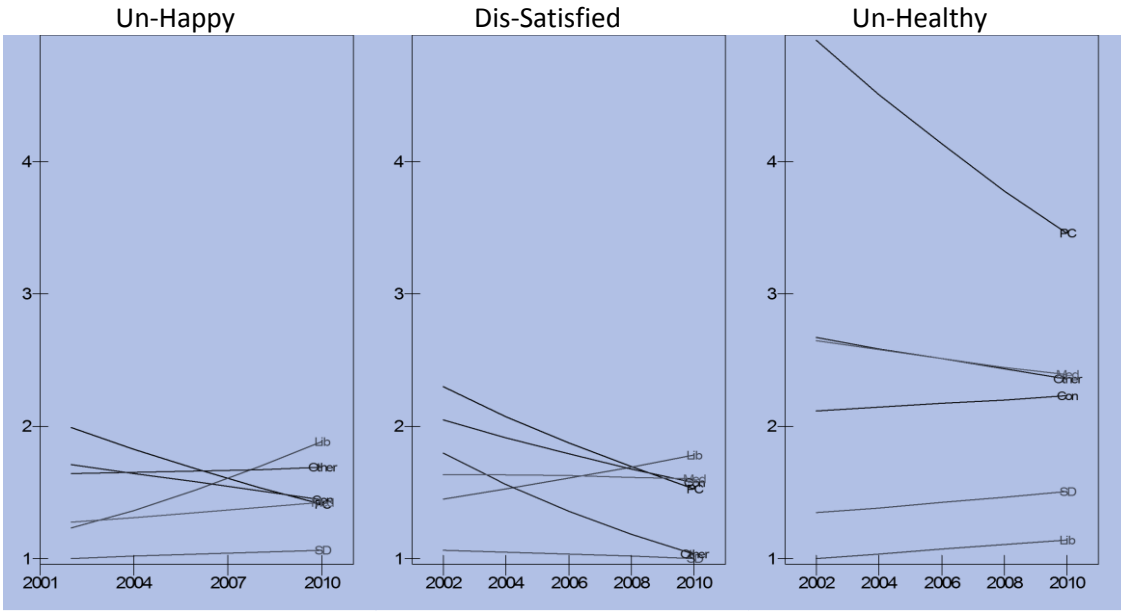

Figure 7. The effect of cross-sectional GDP on well-being: relative odds

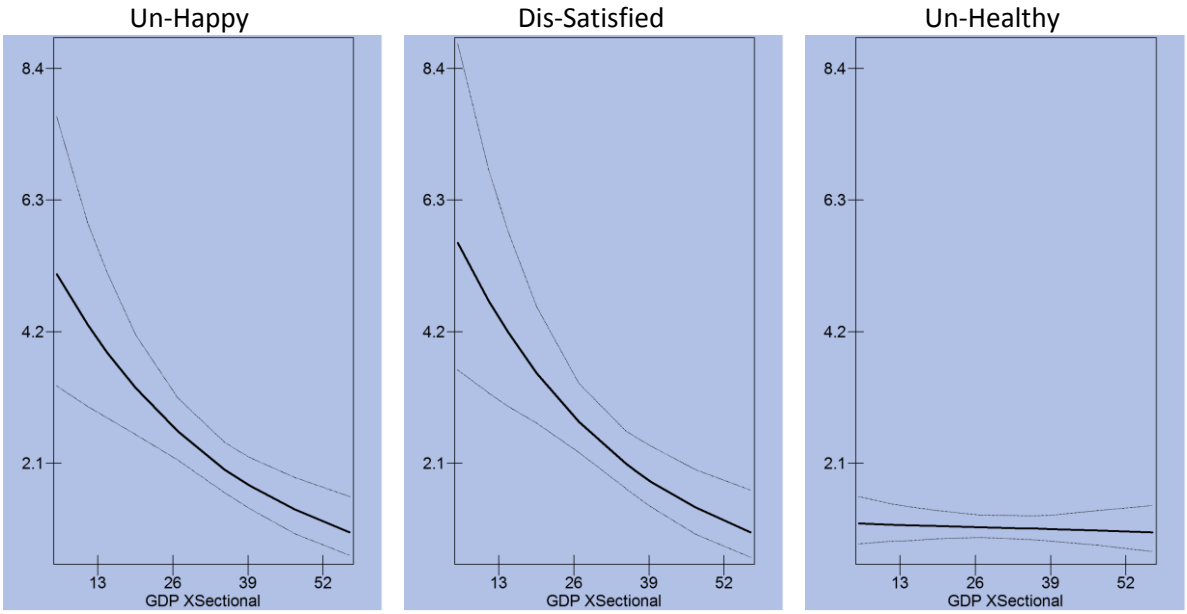

Figure 8. The effect of longitudinal GDP on well-being: relative odds

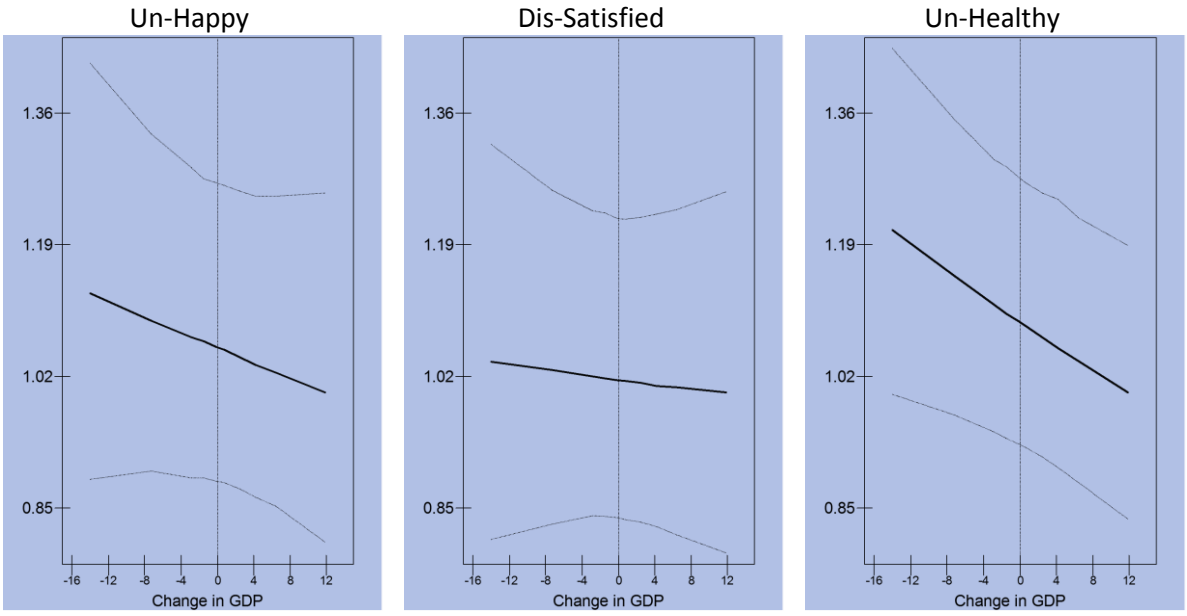

Figure 9. The effects of household income on well-being: relative odds