Session Number: Parallel Session 7B
Time: Friday, August 29, AM

Paper Prepared for the 30th General Conference of The International Association for Research in Income and Wealth

Portoroz, Slovenia, August 24-30, 2008

Life satisfaction in the European Union: also a regional matter?

Andrew Gelman, M. Grazia Pittau and Roberto Zelli

For additional information please contact:

Name: M. Grazia Pittau
Affiliation: Department of Statistics
Full mailing address:
Sapienza, University of Rome P.le Aldo Moro, 5 00185 ROMA, ITALY
Email addresses: grazia@stat.columbia.edu; grazia.pittau@uniroma1.it

This paper is posted on the following website: http://www.iariw.org

Life satisfaction in the European Union: also a regional matter?

Andrew Gelman[*], M. Grazia Pittau[†], Roberto Zelli[‡]

1 Introduction

A blooming economic literature on happiness, started with Easterlin (1974), has developed ways of measuring happiness of individuals induced by life events, economic performances and other "external" factors related to the area where people live (see Frey and Stutzer, 2002, for a survey). Several studies have focused on socio-demographics, like marital status, age and education while a number of other researches have focused on micro-economic aspects including the individual level of income and the role of being unemployed (Diener et al., 1993; Clark and Oswald, 1994; Oswald, 1997; Winkelmann and Winkelmann, 1998). A somewhat different perspective is the evaluation of the effects of macro-variables that reflect the socio-economic environment where the individuals live. Di Tella *et al.* (2003) provide evidence that, after controlling for a large set of individual characteristics, the subjective well-being of Europeans is largely affected by levels and changes of country-level macroeconomic variables, like inflation, per capita GDP, unemployment rate and social welfare state indicators. Alesina *et al.* (2004) and Graham and Felton (2006) address the question of how income inequality affects individual well-being, finding different results between developed and

[*]Department of Statistics, Columbia University, NYC, gelman@stat.columbia.edu

[†]DSPSA, Sapienza Università di Roma, and Department of Statistics, Columbia University, NYC, grazia.pittau@uniroma1.it, grazia@stat.columbia.edu

[‡]DSPSA, Sapienza Università di Roma, roberto.zelli@uniroma1.it

developing countries and also between United States and Europe. Political arrangements also matter. The degree of trust and freedom of democratic institutions (Layard, 2005), as well as the degree of participating in direct democracy (Frey and Stutzer, 2000) seem to positively influence individual reported well-being.

Most of the analyzes in the European Union have focused on the effects on individual well-being of national economic indicators. On the other hand, it has been argued that measuring macro-variables at a territorial sub-national level may influence the individual well-being to a larger extent. The European Union has devoted particular attention to regional disparities among European regions, that are still wide both economically and socially. Their reduction is a target that has been made explicit in the Treaty on European Union and an increasing volume of the EU budget has been devolved towards this objective.

To the best of our knowledge, there are very few studies on the measurement and the determinants of life satisfaction in European regions[1]. Starting from this background, aim of our paper is to evaluate if regional economic factors play a significant role in explaining the subjective well-being of individuals in Europe.

The paper is articulated as follows. In the next section we present the relevant characteristics of the reported life satisfaction in our dataset, the Eurobarometer (EB) surveys, collected and harmonized by ICPSR (Inter-university Consortium for Political and Social Research) in the Mannheim Trend File, 1970-2002. Some descriptive results on life satisfaction in European regions, classified according to the first-level of nomenclature of territorial units, NUTS1, are also reported. This section also presents the econometric approach followed in modelling the data, which is based on a multilevel framework. Section 3 reports the main evidences at individual

[1]A notable exception is the recent analysis of regional well-being in Europe by using European Social Survey data of Aslam and Corrado (2007).

level and at regional and country level, using different model specifications. Throughout, we emphasize graphical summaries of the results. Finally, some concluding remarks are given in Section 4.

2 Data and model design

2.1 Measurement in the Eurobarometer surveys

Research programs aimed to evaluate correlates and determinants of subjective well-being rely, for the most part, on data collected from large surveys in which people are asked to self-report their overall level of happiness and/or life satisfaction on a one to point x scale. A rather skeptical view on its use in economic and policy literature is expressed by Bertrand and Mullainathan (2001) and Wilkinson (2007). In fact, surveys are considered fairly weak instruments for probing into people's feelings (psyches), and more fine-grained self-reported techniques, as ESM (experience sampling methods) or DRM (daily reconstruction method), have been proposed by distinguished psychologists as more accurate tools to evaluate emotional recall (see Kahneman and Krueger, 2006, for a review). Another point that has been raised is that cross-personal comparison is not possible, since different people may understand the concept of life satisfaction or happiness in a different way. In international surveys, moreover, the words happiness and life satisfaction have no precise equivalent in some languages, reflecting cultural heterogeneity (Veenhoven, 2007, has recorded more than a dozen of separate definitions). Another complication is that life satisfaction is not the same as happiness: both are broadly consistent measures of subjective well-being, but have to be considered separately. When asked how happy they are, people tend to consider the more volatile concept of current emotional state, while definition of life-satisfaction is closer to the concept of an overall and more stable living-flourishing and actualizing the best potential within oneself. A person's subjective well-being includes both these emotive

and cognitive judgements, and different people weigh them differently. This explains why several nations report low life satisfaction in the World Value Survey and at the same time high levels of happiness (as for example Nigeria). Even though questions on overall life satisfaction or happiness suffer from several limitations, their use is widely recognized. A review of some of the arguments made in the economic literature in favor of using survey "happiness data" is reported in Di Tella *et al.* (2001) and Alesina *et al.* (2004). An evaluation of the reliability of the global judgment of life satisfaction or happiness has been recently given by Krueger and Schkade (2007). Apart from the fact that large surveys allow comparisons across many different groups of people in terms of socio-economic characteristics, one argument is that global life satisfaction questions have been found to be highly correlated with a variety of relevant physical relations that can be thought of as describing true, internal happiness, as objective physiological and medical criteria (e.g. electrical readings in the brain) or individual's emotional states (e.g. smiling frequency, sleep quality) or ratings made by friends. Furthermore, when using representative population samples, idiosyncratic effects of recent events that may affect the answers are likely to average out. Kahneman and Krueger (2006) also note that respondents are not reluctant to answer global life satisfaction or happiness questions. Overall, we acknowledge the fact that available happiness data are flawed, perhaps more than most economic data in several respects, but at any rate self-reported measurements tell us a lot about the conditions under which different kinds of people are inclined to say that they are satisfied or unsatisfied with life.

Data of our analysis are drawn from the Eurobarometer (EB) surveys, collected and harmonized by ICPSR (Inter-university Consortium for Political and Social Research) in the Mannheim Eurobarometer Trend File, 1970-2002 (Schmitt and Scholz, 2005). The Eurobarometer surveys are conducted on behalf of the European Commission since the early seventies at least two times a year in all member states on a representative sample of people aged

15 and over residing in the EU. The Eurobarometer series is designed to provide regular monitoring of the social and political attitudes in the European Union publics through specific trend questions. The Mannheim Eurobarometer Trend File, a collaborative effort between the Mannheimer Zentrum fur Europaische Sozialforschung (MZES) and the Zentrum fur Umfragen, Methoden und Analysen (ZUMA), combined the most important trend questions of the Eurobarometer surveys conducted between 1970 and 2002. The file consisted of questions asked at least five times in standard Eurobarometer surveys. A total of 1,134,384 respondents from 15 European Union member nations in some years were interviewed in these surveys. The EB surveys have been extended to include European countries only after their entrance to the European Union.

Respondents were also asked for their overall satisfaction with their lives. Life satisfaction is measured on a four-point scale. The question usually asked is: "On the whole, are you very satisfied, fairly satisfied, not very satisfied or not at all satisfied with the life you lead?". Life satisfaction questions have not been asked in the 1996 surveys. The surveys also include a similar question on the level of happiness, on a three-point scale, but this question was not included in the most recent years. Demographic and other background information collected include the respondents' age, gender, and marital status, the number of people residing in the household, the number of children under 15 in the household, respondent's age at completion of education, left-right political self-placement, occupation, religion, income and region of residence.

2.2 Life satisfaction in the European regions

We first present an overall descriptive analysis of life satisfaction in Europe at national and regional level. The EB surveys have a code for the regions where individuals live. We reclassified the Eurobarometer codes according to the most recent Nomenclature of Territorial Units for Statistics (NUTS)

which is based primarily on a normative criteria, that is the institutional divisions currently in force in the Member States. We focus on the first level of the classification (NUTS1) that includes 70 regions for the 15 countries analyzed[2]. This choice seems a reasonable compromise between the need of investigating regional influences, regional data availability and sample size. We focus on the period 1992–2002, which is the last period available in the Mannheim File. Empirical results of happiness equations that make use of the Eurobarometer Surveys did cover the period from mid seventies to, at the most, the year 1992 (Blanchflower and Oswald, 2004; Di Tella *et al.*, 2003). This motivates our choice to verify the consistency of the main findings in the nineties and early 2000's.

In order to present a comprehensive picture, we treat the reported level of life satisfaction as an ordinal measure, accrediting value 1 when the answer is "not at all satisfied" until 4 when the answer is" very satisfied".

Figure 1 reports the trend of life satisfaction scores of the countries, classified in Northern countries (Denmark, Sweden, Finland, UK and Ireland), Western countries (Netherlands, Luxemburg, Belgium, France, Germany and Austria) and Southern countries (Spain, Portugal, Italy and Greece). A substantial stability of the levels of life satisfaction over time is clear for the average of the European Union, with two marked bumps in 1999 and 2001, which are common in almost all the countries. Northern countries are consistently more satisfied than the rest of Europe and are also more stable over time. Not surprisingly, Southern countries show the lowest levels of life satisfaction, but there is a sizeable increase after 1998 along with a more pronounced variability. Western countries follow the average pattern of the European countries.

Specifically, Denmark is by far the happiest country in the European Union and its level of life satisfaction is stable over time. Ireland seems to

[2]Because of data deficiency in EB, Sweden and Finland are considered as whole countries. For the same reason French Départements D'Outre-Mer is excluded.

have the most irregular pattern, with a remarkable increase in 1997 followed by two evident declines, one in 1998 and the other in 2002. Regarding Western countries, what is notable is the increasing trend of France. Similar values of life satisfaction are observed for Germany from 1997. Southern countries exhibit a decreasing pattern in the mid nineties, while from 1997 they seem to have a rise in life satisfaction levels.

Figure 1: Pattern of life satisfaction in European countries (panel a), Northern countries (panel b), Western countries (panel c) and Southern countries (panel d); 1992–2002.

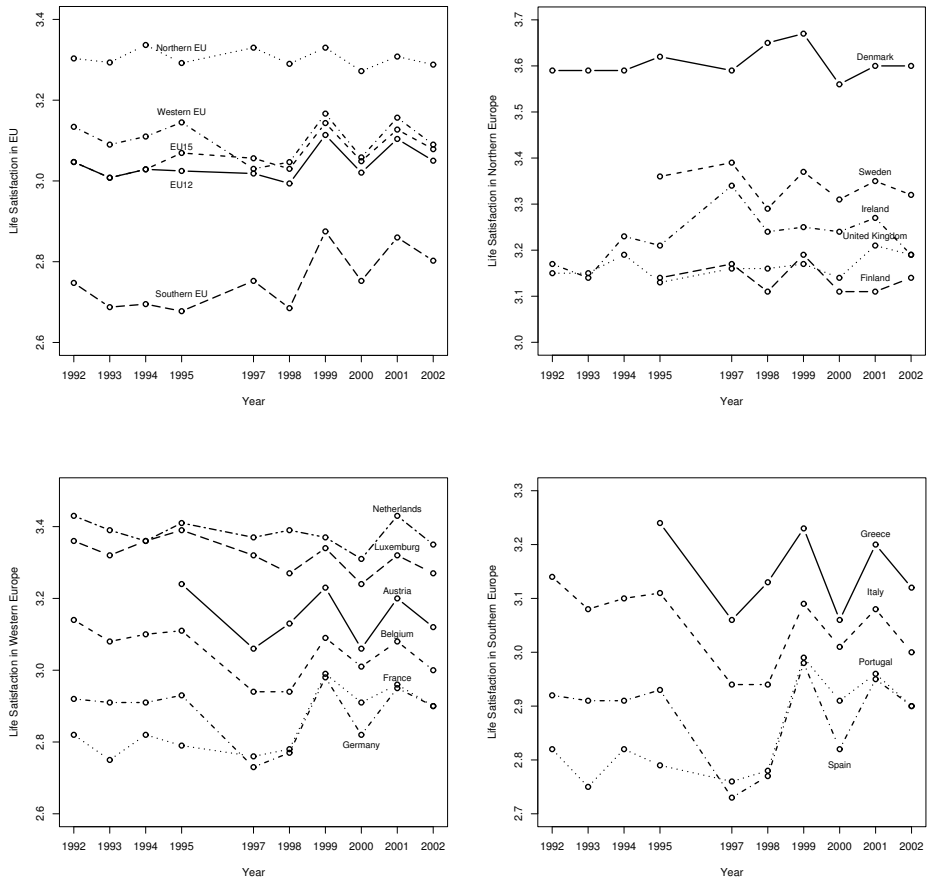


Figure 2 reports the time-average levels of life satisfaction of countries (red triangles) and regions (circles) in Europe in 1992–2002, ranking the countries from the highest satisfied level to the lowest. A marked variability across countries is evident: Scandinavian countries along with the Netherlands and Luxembourg report the highest levels of life satisfaction, while Italy, France, Portugal, and Greece report the lowest. The rank of the countries is similar to the one obtained from the World Value Survey during the period 1995–2005 (Veenhoven, 2007), which is based on combined happiness and life satisfaction scores.

A variability within countries, whereas the NUTS1 nomenclature splits the country in more than one region, is also noticeable. This is more evident in Belgium, Germany, Spain, Italy and Portugal. A table with the codes and the names of the NUTS1 regions along with their average life satisfaction levels 1992–2002 is reported in Appendix A.

Giving some details of the dispersion across regions, we look at the decile ranking that measures where regions fall in the distribution of life satisfaction. Decile rankings range from 1 to 10, with 1 meaning that regions are in the top 10 percent of the region distribution. The Flemish region Vlaams Gewest has a life satisfaction of 3.16 and it is located in the 3rd decile ranking of the distribution, while the Belgian Region Wallonne has a value of 2.89 and it is positioned in the 7th decile. In Germany the happiest region is the northern region of Schleswig-Holstein, placed in the 3rd decile of the distribution, while there are two ex East Germany regions (Thüringen and Mecklenburg-Vorpommern) in the bottom decile. The bottom decile also comprises the Greek regions and the Continental region of Portugal. Spain has one region in the 5th decile (Centro, that includes Castilla y León, Castilla-La Mancha and Extremadura), while the Communidad de Madrid is in the 9th decile. Big disparities are also present in Italy, where residents in southern regions and islands (Sud and Isole, both in the 9th decile) are less satisfied than residents in the north (Nord Est and Nord Ovest, both

in the 6th decile)[3]. More homogeneous are the Netherlands (all the four NUTS1 regions are located in the top decile, with similar values), Greece (all the four in the bottom decile) and, to a lesser extent, Austria and Portugal. French citizens are, on average, more satisfied in the Ouest, which includes Pays de la Loire, Bretagne and Poitou-Charantes, positioned in the 6th decile, and less satisfied in the Mèditerranèe region, positioned in the 9th decile ranking. UK regions are distributed between the 2nd and the 5th decile: Northern Ireland displays the highest level of life satisfaction (3.25, very similar to Ireland, 3.23), while Scotland the lowest (3.07, similar to the northern English regions, North East and North West). Another interesting finding is that regions where the capital city is located usually report levels of happiness among the lowest within their country. This is the case for London, Berlin, Paris, Madrid, Vienna and Athens.

The observed variability across European regions motivates our choice to account for both individual-and regional-level variation in a multilevel framework.

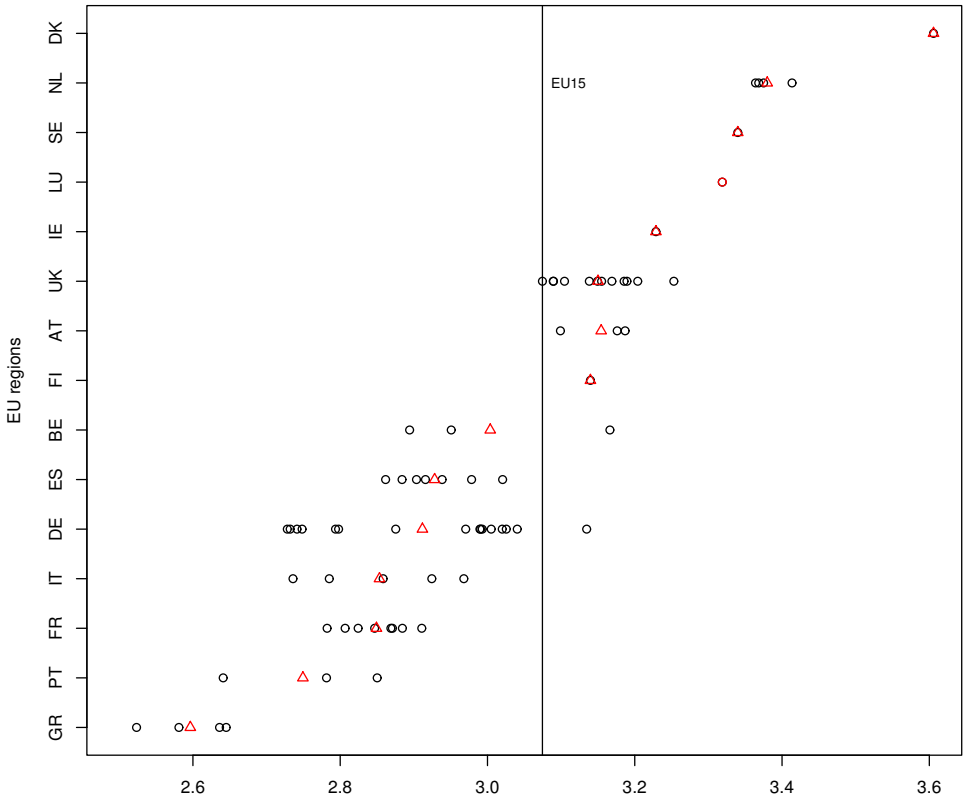
2.3 Multilevel model

As mentioned in the Introduction, aim of the paper is to explore the effects of socio-economic conditions of European citizens, both at micro and macro level, on their level of global life satisfaction. Given the nature of our data, at individual, regional and country level, the statistical model chosen in our analysis is a multilevel model, which has distinct advantages over classical regression models.

Multilevel modelling allows us to estimate patterns of variation within and between groups (in this case, regions and/or countries), taking into account the hierarchical nature of the data (individuals within regions) and also the specific characteristics of each region by allowing their intercepts,

[3]These findings are in accordance to the ones obtained by Scoppa and Ponzo (2008) who use data on Italians subjective well-being from the Bank of Italy Survey of Household Income and Wealth.

Figure 2: Pattern of life satisfaction in European regions 1992–2002. The circles represent the European NUTS1 regions, while the red triangles the European countries. The horizontal line stands for the whole European level of life satisfaction.



and eventually slopes, to vary (see, e.g., Snijders and Bosker, 1999, for a general overview of multilevel models, and Gelman and Hill, 2007, for the notation used here.).

Since life satisfaction answers are intrinsically ordinal, the natural way to treat them in an econometric model should be by ordered logit or probit equations. Because of the extent to which emotions and other individual unobserved characteristics of the well-being can influence the response variable, the goodness of fit measures of the models are in general lower than the ones economists are used to (Graham, 2008).

As discussed in Ferrer-i-Carbonell and Frijters (2004) and Frey and Stutzer (2002), ordinality or cardinality of life satisfaction scores makes little difference, so the use of ordered logit models or linear model is not expected to change the substantive findings. To give confirmation to this hypothesis, in our analysis we implement an ordered logit, a logistic and a linear model for modelling the determinants of satisfaction in life.

Despite the choice of the selected functional form – linear, logistic or ordered – our model is a multilevel varying-intercept model, that estimates life satisfaction on individual socio-demographic characteristics and regional variables.

In our basic model, life satisfaction of individual i resident in region j can be written as[4]:

$$y_i = \alpha_{j[i]} + X_i\beta + \varepsilon_i,\ i = 1, .., n, \tag{1}$$

where y is the life satisfaction level, X is the matrix of individual-level predictors, as the level of education and employment status, ε_i is the error term that includes measurement errors and unobserved personality traits, and $j[i]$ indexes the region j where person i resides.

[4]Without lacking of generality of the multilevel framework, we treat the outcome as a continuous variable. In case of generalized linear models, it is necessary to adapt model (1) to the logistic scale (Gelman *et al.*, 2008).

The second part of the model, what makes it "multilevel", is the simultaneous modelling of the region coefficients α_j:

$$\alpha_j \sim N(\mu_\alpha, \sigma_\alpha^2),\ j = 1, .., J, \tag{2}$$

where μ_α is the mean of the α_j's and σ_α is the standard deviation of the unexplained group-level errors, that can be estimated from the data.

Multilevel model can be thought as a classical regression that includes dummy variables representing group membership. The crucial multilevel modelling step is that these J coefficients are themselves given a model (most simply, as shown in (2), a common distribution for the J parameters or, more generally, a regression model for the α_j's group-level predictors).

When the number of groups is large, there is typically enough information to accurately estimate group-level variation from the data alone and, as a result, multilevel models gain much beyond classical varying-coefficient models, that suffer from reduction in degrees of freedom. Multilevel models, in this setting, estimate more accurately heterogeneous groups in small samples, avoiding the problem of large standard error related to the smallness of sample size in small area estimation procedures (Longford, 2007).

As special cases of multilevel models are the classical regression models. The limit of $\sigma_\alpha \longrightarrow 0$ yields the complete-pooling model, while $\sigma_\alpha \longrightarrow \infty$ reduces to the no-pooling model. Given multilevel data, we can estimate σ_α. Therefore, there is no reason (except for convenience) to accept estimates that arbitrarily set this parameter to one of these two extreme values (Gelman and Hill, 2007).

A further step of the model is to add group-level predictors to improve inference for the group coefficients α_j:

$$\alpha_j \sim N(\gamma_0 + U_j\gamma, \sigma_\alpha^2),\ j = 1, ..., J \tag{3}$$

where U is the matrix of region-level predictors and γ is the vector of coefficients for the region-level regression.

Group-level predictors not only are themselves of interest, but play a special role in the multilevel context, since they may reduce the unexplained group-level variation, that is the standard deviation σ_α. Reduction of unexplained group-level variation can be therefore interpreted as a measure of the importance of the predictor.

3 Empirical results

Our analysis starts fitting models that allows personal characteristics to predict life satisfaction within each region, while also allowing systematic differences between regions. This means adding into model (1) a varying regional intercept α_j according to equation (2), for better interpreting the variation between regions. The inclusion of a varying intercept helps us in understanding a reasonable amount of region-to region variation that remains unexplained after controlling for individual characteristics.

To confirm the findings in the existing literature, we estimate the basic multilevel model (see equations (1) and (2)) considering different specifications: linear, logit and ordered logit[5].

In the linear model, life satisfaction outcome is treated as a continuous variable ranging from 1 to 4; in the logistic model, we model the probability of being very satisfied ($y_i = 1$) with respect to being fairly or not very or not at all satisfied ($y_i = 0$), while in the ordered logistic regression life satisfaction is considered as a four categorical outcome. We exclude interviewers who responded "don't know" or did not respond.

[5]Estimates of the models are obtained by the `lmer` function in R (R development core team, 2006) and are based on the restricted maximum likelihood procedure (REML). The REML procedure corrects the downwards bias of the maximum likelihood estimator of variance components related to the lost of degrees of freedom in estimating the fixed effects. The name `lmer` stands for linear mixed effects in R but the function works also for generalized linear models. However some technical challenges exist in fitting multinomial models in a multilevel framework. Therefore, for the ordered logit model we use the classical no-pooling regression. The term "mixed effects" refers to random effects (coefficients that vary by group) and fixed effects (coefficient that do not vary), (Gelman and Hill 2007).

Individual characteristics used as first-level predictors are: income level, age, marital status, employment status, gender and education level[6]. For other interesting characteristics, like religion, number of children and political self-placement, we have too many missing observations to include them in the equation since for several years these variables have not been collected. The Mannheim Trend File uses twelve income categories, making this variable comparable between countries and EB surveys. Income classes are expressed in absolute values. Educational categories refer to the age when interviewers finished her full-time education and are codified as: up to 15 years old, between 16 and 19 years old, more than 20 years, and still studying.

Regional variables we use as second-level predictors are from Regio, the Eurostat's harmonized regional statistical database. It covers the main aspects of economic and social life in the European Union, classified up to the first three levels of the nomenclature of territorial units, NUTS. National accounts aggregates at NUTS level are based on data from the European System of Accounts ESA 1995, using a harmonized methodology, and are calculated by Eurostat from 1995. On the other hand, as previously mentioned, life satisfaction data are not available in 1996. Therefore, due to data availability in the EB surveys and in the European regional data set, we confine our analysis to the 1997–2002 period. In any case we did not find any significant difference, considering only the individual characteristics, when we expanded the period backward to 1992 (see Figure in Appendix B).

Figures 3, 4 and 5 report the estimated coefficients for the individual characteristics in the three different model specifications. Time indicator variables are included in the models. The zeros correspond to the "baseline" categories for each categorical variable.

[6] As pointed out by, e.g., Di Tella *et al.*(2003) and Frey and Stutzer (2006), estimated effects should be treated with cautions since some personal characteristics can be considered endogenous. Moreover if unobserved personal traits influence reported life satisfaction, results suffer from potential bias.

Figure 3: Estimated coefficients with relative ±1 standard errors of individual characteristics in the basic estimated multilevel linear model: 1997–2002.

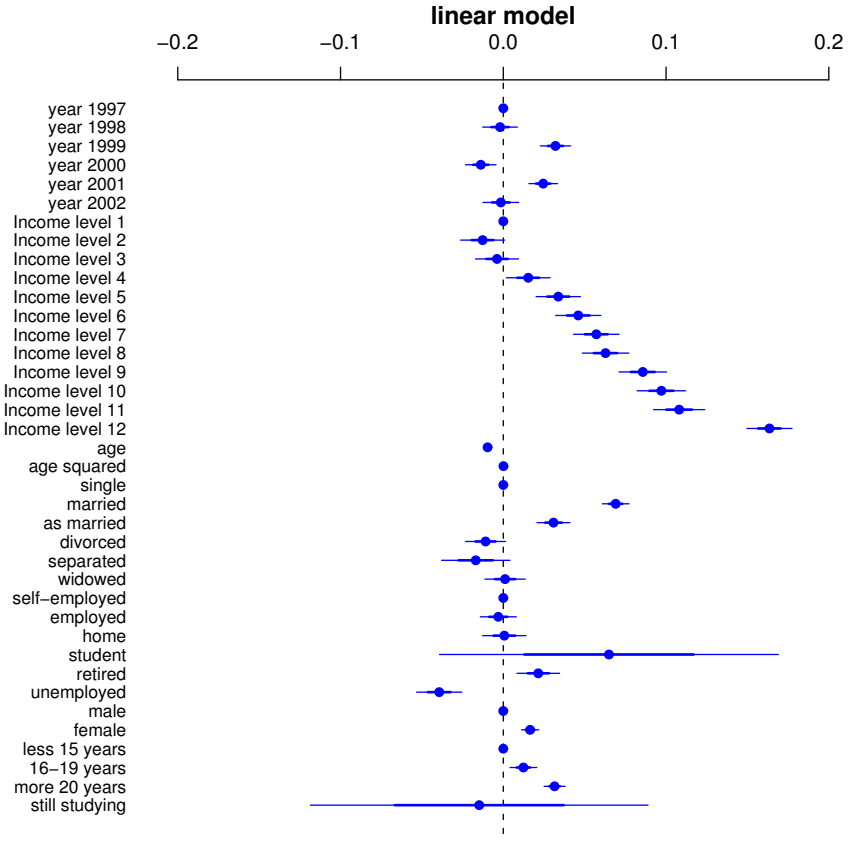


Figure 4: Estimated coefficients with relative ±1 standard errors of individual characteristics in the basic estimated multilevel logistic model: 1997–2002.

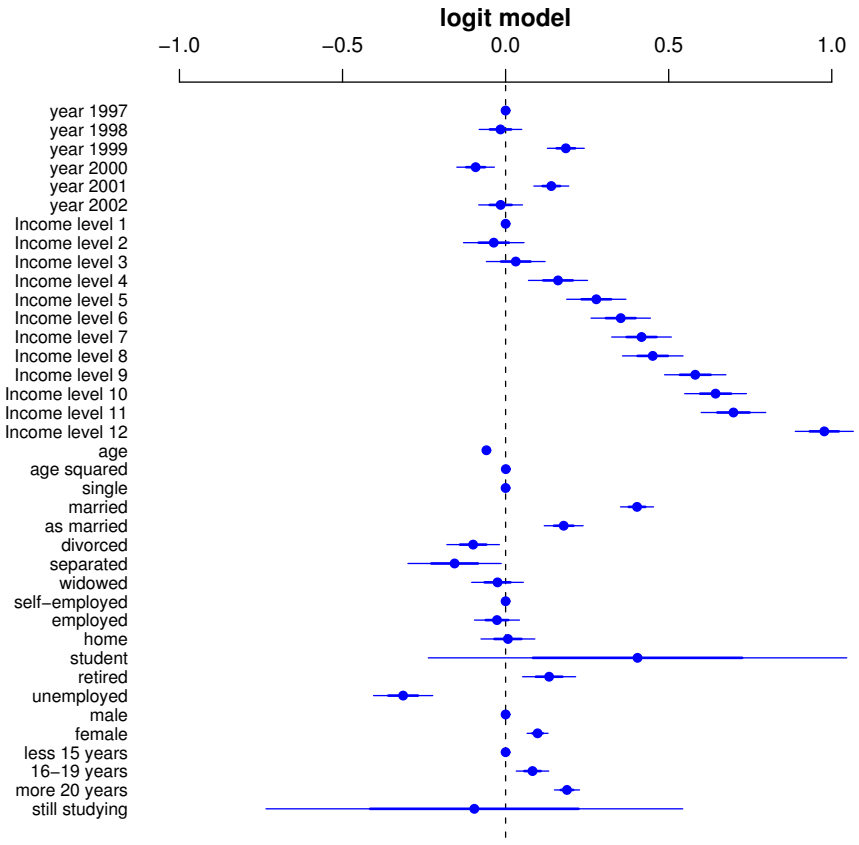
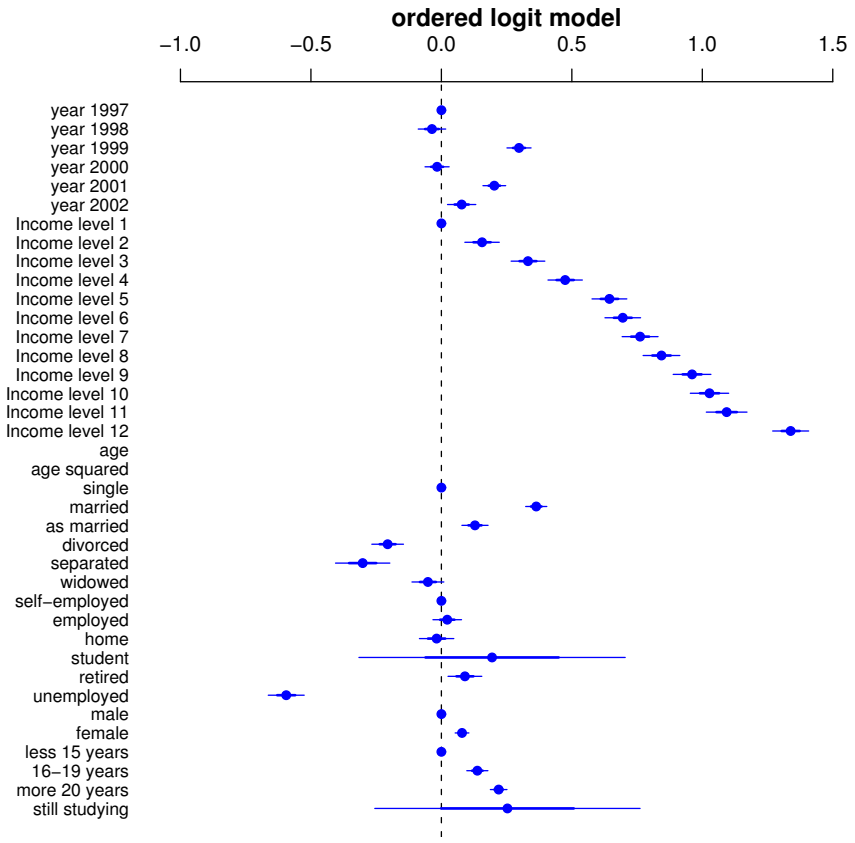


Figure 5: Estimated coefficients with relative ±1 standard errors of individual characteristics in the basic estimated ordered logistic model: 1997–2002. The estimated cutpoints are: $c_{1|2} = -2.25$, $c_{2|3} = -0.31$, $c_{3|4} = 2.88$



In accordance to the main findings in the literature the effects of income are substantial in all the three specifications. In the linear and logistic regression models there are no significant differences with respect to first income class until income level 4. But, for example, moving from the bottom class to the upper income class (level 11) increases, *ceteris paribus*, the life satisfaction score by 0.11 according to the linear model. In the logistic specification this movement increases the probability of being very satisfied by at the most 18%[7]. Interpreting the coefficients of the ordered logistic regression is not straightforward. Given the estimated cutpoints, the "baseline" individual (who is a man of average age, with income level 1, self-employed, single, who left school before 15 years old and interviewed in the year 1997) has a probability of being "very satisfied" approximately equal to 2%, of being "fairly satisfied" equal to 62%, of being "not very satisfied" equal to 37% and being "not at all satisfied" equal to 1%. If this guy increases his income level until class 11 the estimated probabilities change as follow: "very satisfied" to 3%, "fairly satisfied" to 88%, "not very satisfied" 9% and "not at all satisfied" to 0%.

There is a nonlinear effect of age despite the specification of the models: happiness starts off relatively high in early adulthood, then falls, bottoming out on average around age 45, and then rises after that year and on into old age.

Marital status appears to have a positive impact on life satisfaction for those who are in some form of relationship (married or as as married) while it has the most negative effect for those who are separated or divorced. No significant difference exists between single and widowed.

Regarding the employment status, being unemployed has a negative and highly significant effect even though we are controlling for income. Being unemployed with respect to being self-employed reduces by 0.05 life satis-

[7]We applied the "divide by 4 rule" to get un upper bound of the predictive difference in the probability of being very satisfied moving to income categories (Gelman and Hill, 2007).

faction level in the linear specification. In terms of logistic regression model, being unemployed corresponds to an approximate 8% negative difference in probability of being "very satisfied". Again, the interpretation of the ordered logistic model has to be done referring to a specific respondent. As an example, we take the "baseline" individual, whose predicted value of the underline continuous life satisfaction variable is equal to 0. His probability of being "very satisfied" if he becomes unemployed decreases from 2% to 0%, the probability of being "fairly satisfied" goes down from 62% to 40%, while the probability of being "not very satisfied" increases by 18 percentage points. His probability of being "not at all satisfied" goes up from 1% to 5%.

Life satisfaction increases also with years of education. Note that the standard errors of the categories "student" and "still studying" are very high due to collinearity.

As evident from Figures 3, 4 and 5 the estimated models give similar answers for the individual characteristics, and this is also true for the estimated varying intercepts. Our results confirm that effects of individual characteristics analyzed in micro-econometric "happiness" regressions display similar structure, across time and model specifications. Consequently, further analyzes are developed in the linear framework.

Considerable geographical differences still appear after controlling for individual observable characteristics. The group-level equation in the linear model, in fact, tells us that the regional intercepts, α_j's, have an estimated mean of 3.12 and standard deviation of 0.22.

To explore the association between the α_j's and the macroeconomic context, we plot the estimated regional intercepts against the per capita GDP and unemployment rate at regional level. These two economic variables, along with the rate of inflation, have been most thoroughly investigated and are recognized as the most influential. On the first variable, Frey and Stutzer (2002) report studies showing that life satisfaction and income are

uncorrelated over time within countries. Across countries, instead, they observe weak correlation once a certain stage of income has been reached. They also document that higher levels of unemployment reduce the average satisfaction with life, even after controlling for individual unemployment status. The GDP per capita, instead, matters among European countries, as shown in Di Tella et *al.* (2003), giving us a motivation to investigate these relationships across European regions in further details.

Figure 6 reports the estimated regional intercepts against regional (log) GDP per inhabitant at market prices converted to national purchasing power standard (PPS) to make a correction for different cost of living. Unfortunately, Eurostat does not possess comparable regional price levels which would have been enabled us to handle for regional differences in price levels within same countries. Adjusting GDP per capita to national PPS has however the effect of reducing the dispersion in the data since low income in poor regions tend to be partially counterbalanced by the lower cost of living (Pittau and Zelli, 2006). A positive relationship between estimated levels of life satisfaction, once controlled for individual characteristics, and per capita log GDP in the European regions is clearly detected until a certain level, while the relationship seems to become almost irrelevant after a threshold. Note that richer areas are generally those with the largest cities that report low levels of life satisfaction. A conjecture is that, in those areas, other factors that go along with urban agglomeration, like mobility and commuting problems, unsafe environment and perception of unsafeness, negatively affect the level of well-being.

A different picture is detected by looking at the estimated regional intercept versus regional unemployment rates (Figure 7). The clear negative correlation tells us that, after adjusting for personal characteristics, individuals in higher unemployment areas tend to be less satisfied. A conjecture behind this empirical evidence is that people living in areas with high level of unemployment worry about the possibility of becoming unemployed them-

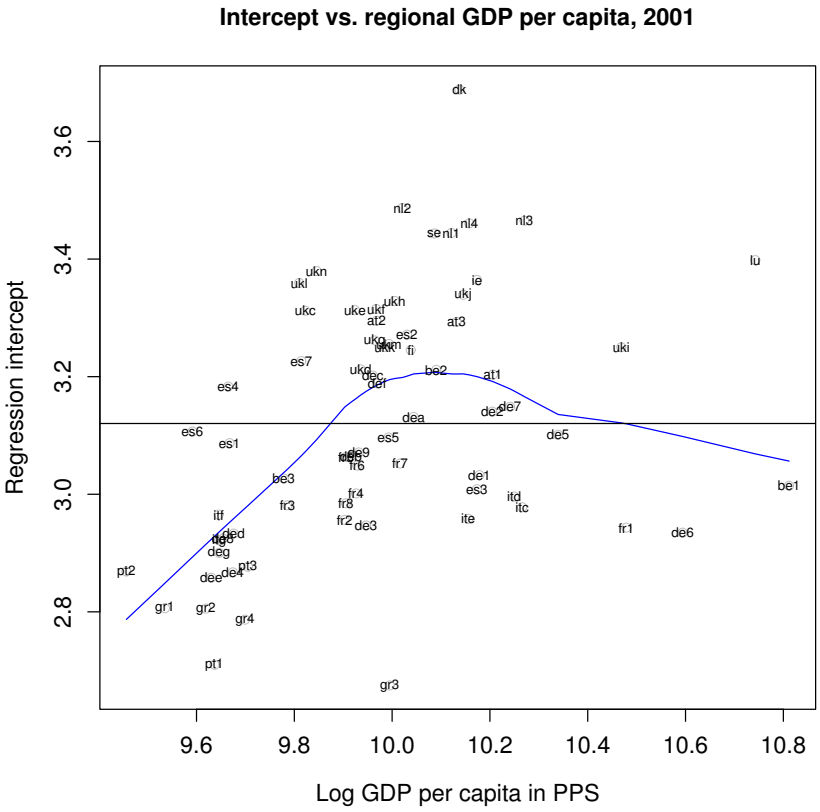


Figure 6: Estimated regional intercepts α_j in the varying intercept regression model plotted versus regional per capita GDP. A non parametric regression line is fitted to the estimates.

selves and feel unsafe because of the potential increase of social tensions.

We add as regional level predictors in the multilevel model (equation 2) the regional GDP per capita in purchasing power standards and the unemployment rate.

We start fitting the model introducing the unemployment rate, which ha the expected negative sign and it is highly significant. Its introduction in the model reduces the unexplained group-level standard deviation by 10%, confirming the relative importance of this predictor in understanding life satisfaction disparities across EU regions.

Adding per capita GDP in the model for the whole period (1997–2002) is more problematic. As pointed out by Di Tella et *al.* (2003), a potentially

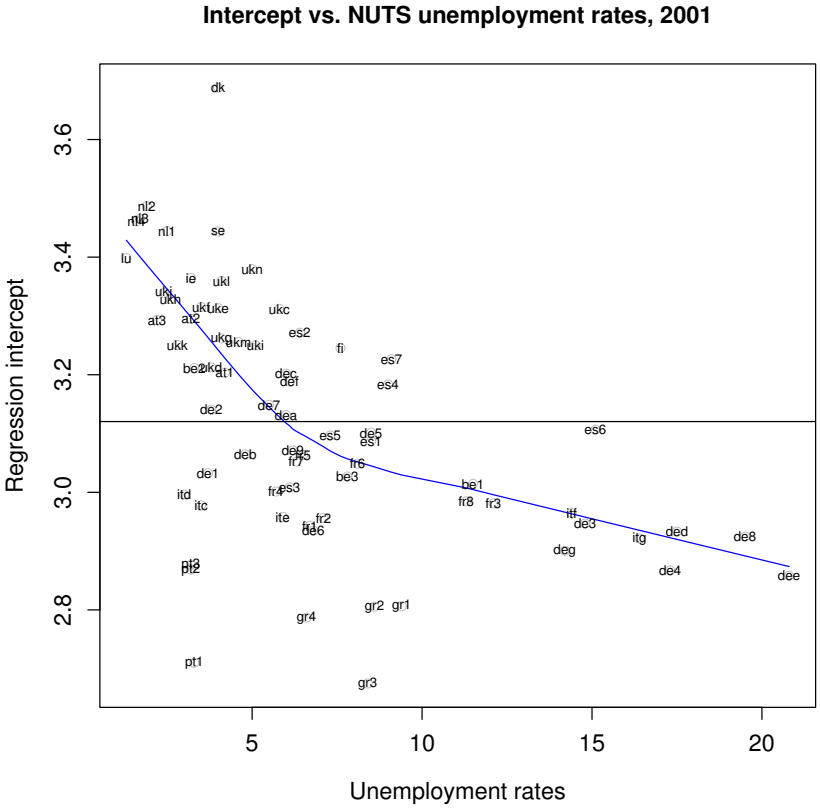


Figure 7: Estimated regional intercepts α_j in the varying intercept regression model plotted versus regional unemployment rate. A non parametric regression line is fitted to the estimates.

non stationary predictor (the GDP) is introduced to explain an outcome that is naturally stationary (the life satisfaction rated on a 4-point scale). Therefore, the estimated coefficients may be "unpersuasive" due to the inapplicability of conventional statistical procedures. The (stochastic) trend of the non stationary variable will in fact dominate all other variations.

To overcome this problem we prefer to model the time-series structure of our dataset repeating the model year-by-year. The method of repeated modelling, followed by time-series plots of estimates is rarely used as a data analytic tool but it can be very informative and easy to understand.

The estimated coefficients of equation (3) of the multilevel model are plotted in Figure 8. In the model we use as regional level predictors the GDP per inhabitant in PPS at current market prices[8] and the unemployment rate. We also add two more predictors that measure the disparity between the regional variable and its corresponding country value. The first variable is calculated as the ratio of the per capita GDP of region j and the per capita GDP of the country k where region $j[k]$ belongs to. Analogously, unemployment disparity is the ratio of the regional unemployment rate and the respective country rate.

The inclusion of these additional predictors help exploring the possibility of having also a country effect by measuring the discrepancies in GDP per capita ("GDP disparities") and unemployment ("unemployment disparities") that eventually turn out between regional and national levels.

The effect of per capita GDP is large and it remains fairly stable over time, confirming that this variable, when observed cross-sectionally, matters. Also the unemployment rate has a significant and sizeable effect, with the expected negative sign. Conversely, the coefficient of unemployment disparities is not statistically significant, indicating that what really matters is

[8]To check that the results on per capita GDP do not just reflect the interest for relative income, we use, as individual predictor, the income quartile ranking within which individuals lie. Results however do not differ from those obtained with the income absolute categories.

the regional unemployment rate rather than the corresponding national one. The coefficients of the GDP disparities, instead, are statistically significant but small in size. Their negative signs may imply that there is aversion to economic disparities within countries, that is people who live in richer areas of the country are aware of poor areas in their own country.

When we study these macroeconomic variables with a time lag, the substantive conclusions remain the same.

Another possible solution to the non stationarity problem, is to introduce as predictor the real growth rate of income. Unfortunately, Regio Eurostat source releases the real GDP growth rate only from the year 2000[9]. Limiting the analysis to the period 2000–2002, we estimate our multilevel model using as regional level predictors, the regional growth rate ("real GDP growth") and its difference with the corresponding national real GDP growth rate ("growth disparities"), the regional unemployment rate, its disparity with respect to the country level.

Figure 9 shows the estimated coefficients tracked over time. Apart from the unemployment rate that exhibits a stable pattern with its maximum value in the year 2001, all the other predictors are not statistically significant and it is not clear whether we should "believe" in their interpretation.

4 Concluding remarks

Our analysis on reported life satisfaction in Europe over the period 1992–2002 has emphasized the relevance of regional disparities, between and within countries. It has also documented that people living in big capital cities are generally less satisfied, also after controlling for socio-economic and demographic characteristics of the individuals.

Since we have 70 European regions, classified according to the first-level more recent nomenclature of the territorial units, a consistent number

[9]Note that the GDP per capita released by Regio from 1995 is at current market prices. Regional real growth rates are not available for UK regions.

Figure 8: Estimated coefficients (and ±1 standard errors) for the multilevel regression model of life satisfaction on per capita GDP, GDP disparities, unemployment rate and unemployment disparities, as fit separately from 1997 to 2002.

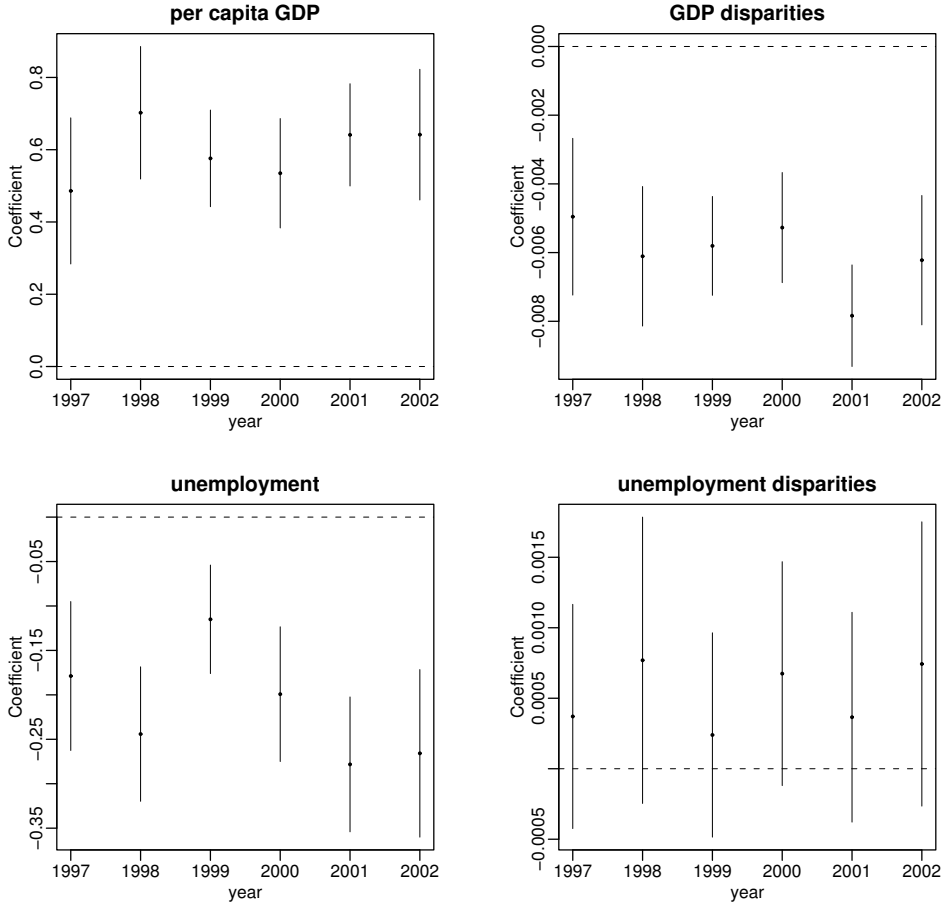
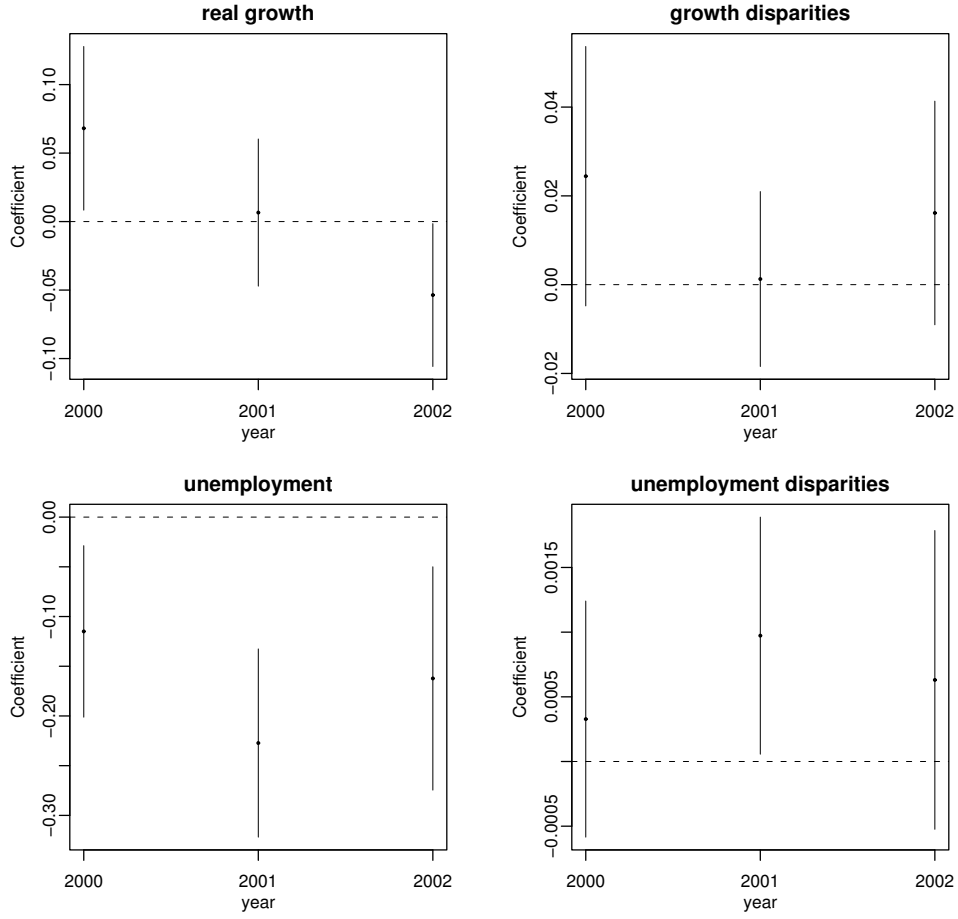


Figure 9: Estimated coefficients (and ±1 standard errors) for the multilevel regression model of life satisfaction on real GDP growth, growth disparities, unemployment rate and unemployment disparities, as fit separately from 2000 to 2002.



of indicator variables should have been introduced in the classical micro-econometric subjective well-being equations. The multilevel framework gave us a more suitable environment for modelling these hierarchical data, individuals within European regions. With respect to traditional alternatives, multilevel modelling allowed us to estimate differences between large number of groups (in terms of variation between groups) and to analyze data coming from different sources inside the same model.

The coefficients associated with individual observable characteristics are in accordance with the most important findings in the literature, with life satisfaction largely increasing with personal income and negatively affected by being unemployed, divorced and low educated. These findings are robust to alternative specifications of the models and show a stable structure.

The group-level regression on regions has documented that, across NUTS1, local per capita GDP and unemployment rates are correlated with reported life satisfaction. The effects are large and stable over time even after controlling for the personal level of income and employment status. The introduction of regional per capita GDP and unemployment rates reduces the unexplained regional-level variability by around 10% steadily over time, indicating that other factors may be equally or more important to explain differences in reported subjective well-being between regions. The effects of these macroeconomic regional variables largely sweep off the effects that the corresponding country variables have on reported life satisfaction, showing that the factors that affect the subjective well-being are essentially local.

To better understand the importance of macro-economic effects and, eventually, the different consequences on life satisfaction of region and country factors, we plan to further analyze the country source of variation, generalizing this line of research.

References

Alesina, A., Di Tella R., MacCulloch R. (2004), "Inequality and happiness: are Europeans and Americans different?" *Journal of Public Economics*, 88, 2009–2042

Aslam, A., Corrado L., (2007), "No Man is an Island: The Inter-Personal Determinants of Regional Well-being and Life Satisfaction in Europe", Cambridge Working Paper in Economics, CWPE 0717.

Bertrand, M. Mullainathan S. (2001), "Do people mean what they say? Implications for subjective survey data", *American Economic Review*, 91, 67–72.

Blanchflower, D.G., Oswald, A.J. (2004), "Well-being Over Time in Britain and the USA", *Journal of Public Economics*, 88, 1359–1386.

Clark, A., Oswald, A.J. (1994), "Unhappiness and Unemployment", *Economic Journal*, 104, 648–659.

Diener, E. Sandvik, E., Seidlitz, L., Diener, M. (1993), "The relationship between income and subjective well-being: Relative or absolute?" *Social Indicators Research*, 28, 195–223.

Di Tella, R., MacCulloch, R., Oswald, A.J. (2001), "Preferences over Inflation and Unemployment: Evidence from Happiness Surveys", *American Economic Review*, 91(1), 335–42.

Di Tella, R., MacCulloch, R., Oswald, A.J. (2003), "The Macroeconomics of Happiness", *Review of Economics and Statistics*, 85, 809–827.

Easterlin, R.A. (1974), "Does economic growth improve the human lot? Some empirical evidence", in David P., Reder, M., Nations and Households in Economic Growth, New York, Academic Press.

Ferrer-i-Carbonell A., Frijters P. (2004), "How Important is Methodology for the Estimates of the Determinants of Happiness?", *Economic Journal*, 114, 641–659.

Frey, B.S., Stutzer, A. (2000), "Happiness, Economy and Institutions", *Economic Journal*, 110, 918–938.

Frey, B.S., Stutzer, A. (2002), *Happiness and Economics: How the Economy and Institutions Affect Human Well-Being*. Princeton: Princeton University Press.

Frey, B.S., Stutzer, A. (2006), "Does Marriage Make People Happy, Or Do Happy People Get Married?" *Journal of Socio-Economics*, 35(2), 326–347.

Gelman, A., Shor B., Bafumi J., Park D. (2008), "Rich state, poor state, red state, blue state: what's the matter with Connecticut?", *Quarterly Journal of Political Science*, 2, 345–367.

Gelman, A., Hill J. (2007), *Data Analysis Using Regression and Multilevel/Hierarchical Models*, Cambridge University Press, Cambridge.

Graham, C., Felton A. (2006), "Inequality and happiness: Insights from Latin America", *Journal of Economic Inequality*, 4(1), 1569–1721.

Graham, C. (2008), "The Economics of Happiness", in S. Durlauf and L. Blume (eds) *The New Palgrave Dictionary of Economics*, MacMillan, 2 Rev.ed edition.

Layard, R. (2005), *Happiness: Lessons from a New Science*, New York, Penguin Press.

Kahneman, D., Krueger A.B. (2006), "Developments in the Measurement of Subjective Well-Being", *Journal of Economic Perspectives*, 20(1), 3–24.

Krueger A.B., Schkade D.A. (2007), "The Reliability of Subjective Well-Being Measures", NBER Workin Paper n.13027.

Longford, N.T. (2007), *Missing Data and Small-area Estimation*, Springer-Verlag, New York, NY.

Oswald, A.J.(1997), "Happiness and Economic Performance", *Economic Journal*, 107, 1815–1831.

Pittau, M.G., Zelli R. (2006), "Empirical Evidence of Income Dynamics

Across EU Regions", *Journal of Applied Econometrics*, 21, 605–628.

R Development Core Team (2006), R: A Language and Environment for Statistical Computing, Vienna, Austria: R Foundation for Statistical Computing, www.r-project.org.

Schmitt, H., and Scholz E. (2005), The Mannheim eurobarometer Trend File, 1970-2002 [Computer file]. Prepared by Zentralarchiv fur Empirische Sozialforschung. ICPSR04357-v1. Mannheim, Germany: Mannheimer Zentrum fur Europaische Sozialforschung and Zentrum fur Umfragen, Methoden und Analysen [producers]. Cologne, Germany: Zentralarchiv fur Empirische Sozialforschung/Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributors].

Scoppa, V., Ponzo M. (2008), "An Empirical Study of Happiness in Italy", *B.E. Journal of Economic Analysis and Policy*, vol.8, 1 (Contributions), article 15.

Snijders, T.A.B., Bosker R.J (1999), *Multilevel Analysis*, London, Sage.

Veenhoven, R. (2007), World Database of Happiness, Distributional Findings in Nations, Erasmus University Rotterdam, www.worlddatabaseofhappiness.eur.nl

Wilkinson, W. (2007), "In Pursuit oh Happiness Research. Is It reliable? What Does It Imply for Policy?", *Policy Analysis*, n.590.

Winkelmann, L., Winkelmann, R.(1998), "Why are the Unemployed so Unhappy?", *Economica*, 65(257), 1–15.

Appendix A:
Life satisfaction (LS) in European Regions (NUTS1), average 1992-2002

Region (NUTS1)	Label	LS
Région de Bruxelles	**be1**	2.95
Vlaams Gewest	**be2**	3.16
Région Wallonne	**be3**	2.90
Denmark	**dk**	3.61
Baden-Württemberg	**de1**	2.99
Bayern	**de2**	3.04
Berlin	**de3**	2.79
Brandenburg	**de4**	2.74
Bremen	**de5**	2.99
Hamburg	**de6**	2.88
Hessen	**de7**	3.01
Mecklenburg-Vorpommern	**de8**	2.73
Niedersachsen	**de9**	2.99
Nordrhein-Westfalen	**dea**	3.03
Rheinland-Pfalz	**deb**	2.97
Saarland	**dec**	3.02
Sachsen	**ded**	2.80
Sachsen-Anhalt	**dee**	2.75
Schleswig-Holstein	**def**	3.13
Thüringen	**deg**	2.73
Ireland	**ie**	3.23
Voreia Ellada	**gr1**	2.65
Kentriki Ellada	**gr2**	2.58
Attiki	**gr3**	2.52
Nisia Aigaiou, Kriti	**gr4**	2.64
Noroeste	**es1**	2.94
Noreste	**es2**	3.02
Comunidad de Madrid	**es3**	2.86
Centro (ES)	**es4**	2.98
Este	**es5**	2.91
Sur	**es6**	2.88
Canarias (ES)	**es7**	2.90
Île de France	**fr1**	2.85

Bassin Parisien	**fr2**	2.81
Nord - Pas-de-Calais	**fr3**	2.82
Est	**fr4**	2.87
Ouest	**fr5**	2.91
Sud-Ouest	**fr6**	2.88
Centre-Est	**fr7**	2.87
Mèditerranèe	**fr8**	2.78
Nord Ovest	**itc**	2.93
Nord Est	**itd**	2.96
Centro (IT)	**ite**	2.86
Sud (IT)	**itf**	2.78
Isole (IT)	**itg**	2.73
Noord-Nederland	**nl1**	3.36
Oost-Nederland	**nl2**	3.41
West-Nederland	**nl3**	3.37
Zuid-Nederland	**nl4**	3.37
Ostösterreich	**at1**	3.10
Südösterreich	**at2**	3.18
Westösterreich	**at3**	3.19
Continente	**pt1**	2.63
Região Autónoma dos Açores	**pt2**	2.85
Região Autónoma da Madeira	**pt3**	2.77
Finland	**fi**	3.14
Sweden	**se**	3.34
North East	**ukc**	3.15
North West (including Merseyside)	**ukd**	3.09
Yorkshire and The Humber	**uke**	3.14
East Midlands	**ukf**	3.19
West Midlands	**ukg**	3.09
Eastern	**ukh**	3.20
London	**uki**	3.10
South East	**ukj**	3.19
South West	**ukk**	3.17
Wales	**ukl**	3.16
Scotland	**ukm**	3.07
Northern Ireland	**ukn**	3.25
Luxembourg (Grand-Duché)	**lu**	3.32

Source: authors' calculations on Mannheim Eurobarometer trend file.

Appendix B: Estimated coefficients with relative ± 1 standard errors for the regression of life satisfaction on individual characteristics 1992-2002.