Growth-Innovation-Competitiveness
Fostering Cohesion in Central and Eastern Europe

Effects of Income Inequality on Population Health and Social Outcomes at the Regional Level in the EU - Differences and Similarities between CEE and Non-CEE Regions.

Sebastian Leitner

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 290657

Sebastian Leitner
The Vienna Institute for International Economic Studies
WP5 T2, P5.2.
Final version: October 20, 2014

Effects of Income Inequality on Population Health and Social Outcomes at the Regional Level in the EU - Differences and Similarities between CEE and Non-CEE Regions.

Abstract: This paper analyses the relationship between various measures of income inequality and variables describing population health and social outcomes at the regional level in the EU. Differences between the New EU Member States (NMS) and Non-NMS EU countries are highlighted. Applying fixed and random effects cross-region regressions we find negative relationships between income inequality and life expectancy, infant mortality, standardised death rates on various causes, rates of violent and property crime, rates of non-activity and early leave from education of youngsters. The results indicate that redistributive policies might be a proper measure to reduce social harm and increase population health.

1. Introduction

Inequality of disposable incomes increased starting in the 1980s in almost all European and furthermore OECD countries (OECD, 2011). In the Central and East European New EU Member States[1] (NMS) the transitional crisis at the beginning of the 1990s raised inequality levels strongly in countries like Bulgaria, Romania and the Baltic States, while these remained at relatively low levels in e.g. the Czech Republic and Slovenia (Holzner and Leitner, 2008).

There is a controversial discussion in the literature whether income inequality is, in addition to income levels, an important determinant of population health and further social outcomes. About twenty years ago particularly epidemiologists started to analyse correlations between those variables first on the cross-national level and later on also on the regional level, particularly for the US (for an overview see e.g. Babones, 2008). In 2009 Wilkinson and Pickett popularised the hypothesis that higher levels of income inequality lead to social harm in various aspects particularly in high income countries in their book 'The Spirit Level: Why More Equal Societies Almost Always Do Better'. However, evidence which points towards positive conditional correlations between inequality and social harm and possible causal relationships in cross-country and regional analysis particularly for the US and some European countries remained not unchallenged. Research for the whole EU is scarce and has to our best knowledge been performed at the regional level so far only by Elia et al (2013), who however analyse only bivariate unconditional correlations at the NUTS 1 level.

In this paper we analyse the relationship between income inequality and variables describing social outcomes across the EU NUTS 2 regions where available (and NUTS 1 level in other EU countries) and highlight the differences between the Central and Eastern European New EU Member States (NMS)

[1] Bulgaria, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania, Slovak Republic and Slovenia.

and Non-NMS EU countries. The quantitative (regression) analysis is based on EU regional data retrieved from Eurostat and the OECD. In addition we use EU SILC microdata to calculate inequality indicators not provided by the former two data sources.

The paper goes as follows. In section 2 we give a brief overview of the discussions going on in the literature on possible interrelations between income inequality and social outcomes. Section 3 follows with a description of the data used for the econometric analysis. In section 4 we describe the methodology applied and present the results of our regression analysis in section 5. Section 6 concludes and sums up with policy recommendations.

2. Literature overview

The possible effects of income inequality on various population health and social outcome indicators are analysed in different strains of the literature. Most of the studies have been performed at the cross-country level or if at the regional level most often only for individual countries. Most prominently Wilkinson and Pickett (see e.g. Wilkinson and Picket, 2009) claim in their overview that in societies where income is more equally distributed not only population health outcomes (measured e.g. by life expectancy) tend to be better, but also many other social problems, including violence, imprisonment, lack of trust, teenage births, drug abuse and poor educational performance of schoolchildren are less common. This hypothesis is however not undisputed in the literature.. Comprehensive studies on the situation in the EU at the regional level have not been done so far. Only one research paper has been published by Elia et al (2013) who analyse bivariate correlations between indicators of income inequality and social outcomes for the EU NUTS 1 regions. Thus they have not applied control variables and did not control for non-observed country heterogeneity. Many of the indicators of social outcomes used in the analysis are subjective (based on self-reported information). Moreover, the use of NUTS 1 level data (instead of NUTS 2 level data) prevents the analysis from including inter-regional differences in the Czech Republic, Denmark, Slovenia, Slovakia and in practical terms also in Finland.

Torre and Myrskylä (2014) point out that in the literature concerning **population health** effects it was discussed already for years (see e.g. Preston, 1975 and Rodgers, 1979) that mortality rates do not particularly differ between economically developed countries and regions. However, income inequality might explain these differences in social outcomes. This ‘income inequality-health hypothesis’ is based on the idea that within a population an income inequality decreasing transfer from the rich to the poor would increase population average health levels as the marginal health gain is higher at lower income levels compared to the marginal health loss at higher income levels (see also Deaton, 2003). Lynch, Smith, Kaplan and House (2000) have classified the possible mechanisms through which income inequality and population health could be related as follows. (i) The ‘Individual (or absolute) income interpretation’ asserts that individual absolute income accounts for all the health effects of income distribution and population health effects of income inequality are therefoe merely the sum of individual income effects (e.g. Gravelle 1998; Heerink, Mulatu and Bulte 2001; Gravelle et al. 2002). (ii) The ‘Psychosocial (or relative income) interpretation’ argues that individual health is affected by the perceptions that people have on their relative position in the social hierarchy, and lower positions may be related to lower investment in human capital, lack of social cohesion, feelings of insecurity (Wilkinson 1992; Porta, Borrell and Copete 2002;) and thus to chronic stress that translates into an unhealthier life style (Leigh et al, 2009). (iii) The ‘Neo-material interpretation’ suggests that income inequality affects health mainly through the lack of resources held by individuals and the systematic underinvestment of the society in social and environmental conditions (Lynch et al., 2000). This implies that greater heterogeneity hinders societies to agree on investments in public goods, e.g. in the health sector (e.g. Alesina et al, 1999). Much of the research on income inequality and mortality has been done at community and individual levels, as discussed in several review papers (Wagstaff and van Doorslaer, 2000; Macinko et al, 2003; Lynch et al, 2004; Wilkinson and Pickett, 2006). However, as pointed out by Torre and Myrskylä (2014), studying the income inequality – population health link at the population level might be more useful since income inequality is by definition a property of the population and not of the individual (see also Kaplan et al, 1996). Moreover, by studying the link at the population level social and environmental factors behind the association can be taken into account (Lynch et al. 2004). A majority of the studies at the population level done so far seem to find a negative relationship between income inequality and life expectancy or mortality. The results were most consistent for infant mortality, but less so for adult mortality. However some recent studies found negative or non-robust relationships (see e.g. De Vogli et al, 2005; Dorling et al, 2007; Gravelle et al, 2002). Torre and Myrskylä (2014) claim that many of these cross-sectional studies however suffer from one or more shortcomings: the use of simple bivariate methods without appropriate controls; no consideration of the possibility of unobserved country heterogeneity; the use of measures of income distribution that are often not internationally comparable. Moreover, they criticise that cross-country studies often pool countries of different income levels, although the relationship between inequality and population health indicators might be very dissimilar in groups of economic developed and less developed countries. Elia et al (2013) find no significant bivariate correlations between income inequality and the applied subjective indicators of 'self-assessed health status' and 'presence of chronic diseases' for the EU NUTS 1 regions.

The economic theories on **criminality** date back to Becker (1968) who has described the criminal act as a result of a cost-benefit analysis. According to this theory higher income inequality would increase the probability that crime is committed in a society, since the relative benefit would rise. In epidemiology the favoured explanatory theories have also been based on psycho-social processes (Lynch et al, 2001) such as socio–economic position, social status, disrespect, social support, anxiety, trust and community cohesion. These affect social interactions and behaviours and lower (or raise) the inhibitions of an individual to commit crime (Case and Katz, 1991). Though in general crime rates are objective measures, one should be aware the data collected is however that of reported crime and not that of incidents executed. People might deter from notifying the incident if they expect the crime not to be solved, or if the cost of reporting is high compared to the loss caused. In addition, the quality of the data tends to depend on the legal differences across countries in the way crimes are defined and on country-specific police and justice systems. Homicide and robbery rates tend to be more reliable figures compared to other types of crime since the violence associated with such criminal acts tends to increase the proclivity for the victim to officially declare the crime to the police (Fajnzylber et al, 2002a). The most prominent cross-country studies on homicide and robbery rates were performed by Fajnzylber et al (2002a, 2002b) who found in their studies using data on 40 countries over 1970-1994 that income inequality measured by different indicators robustly increases crime rates. Similar findings are stated by Messner et al (2002), while Neumayer (2005) claims that if one controls for country-specific fixed effects, then income inequality no longer is a statistically significant determinant of violent crime. A number of studies analyse the interrelation on the US county and state level. Most find positive correlations (e.g. Choe, 2008), while Brush (2007) finds no association when estimating in first differences and Kelly (2000) finds only significant relations for violent crime but not for property crime. Machin and Meghir (2004) find robust correlations for the level of wage income of the 25th percentile and crime. For Sweden Nillson (2004) finds a positive influence of poverty rates on property crime but not for other inequality measures and not on assault. Results of Entdorf and Spengler (2000) who apply static and dynamic panel estimations on data covering the German states (Länder) show no significant results for robbery. Elia et al (2013) find positive bivariate correlations between income inequality and domestic burglary of a magnitude of about 0.3 for the EU NUTS 1 regions.

Analysing the intergenerational transmission of educational outcomes and differences in **educational attainment**, various authors have pointed to the fact that in the presence of imperfect credit markets, the actual wealth distribution should affect the distribution of investments in human capital in a society (see e.g. Perotti, 1993; Banerjee and Newman, 1993; Galor and Zeira, 1993). Parts of the population with lower income or wealth possess fewer resources to access education and do not find financial markets to borrow these resources to send their children to higher education. In this case redistribution from high to low income families would raise overall attainment rates in a country or region since such transfers would allocate funds from individuals with lower marginal rates of return to liquidity constrained agents with high rates of return (Cecchi, 2003). Cecchi tests a panel of 108 countries for the period 1960-95 and finds a negative dependence of enrolment rates on the Gini index of income. Similar results have already been presented by Flug et al (1998). However, there might be more reasons for not investing in the education of children. The effective returns to educational investments might strongly depend upon factors of family background of individuals other than family income as e.g. described by Aakvik et al (2005). Widening income inequality that may raise the cultural differences between lower and upper 'classes' could thus also increase differences between returns to education of population subgroups in the society. Lower investments in education in lower income families would result therefrom and the efficiency gains of distributive measures would take longer to materialise. In their study on EU NUTS 1 regions Elia et al (2013) find positive bivariate correlations between income inequality and the rate of early school leavers of a magnitude of 0.35 in the case of the Gini index and 0.42 in the case of the quintile share ratio (S80/S20).

3. Data on social outcomes and income inequality at the EU regional level

In order to compare the situation in the EU between regions we analyse various aspects of social outcomes but have confined the analysis to indicators of population health, crime and educational attainment of youngsters providing objective measures for a cross-section of European regions. Regional data at the NUTS 2 (and NUTS 1) level have been collected from the Eurostat database and the OECD well-being dataset. For the inequality measures Gini coefficients and income quintile share ratios (explained in detail below) at the regional level are available only for the year 2010 and the poverty rates are also only available for a short time span. NUTS 2 level data (according to the NUTS 2010 classification) have been available for the following EU countries: Austria, Czech Republic, Denmark, Finland, France, Italy, Slovak Republic, Slovenia, Spain, Sweden and United Kingdom. In the case of the following EU Member States the country comprises of just one NUTS 2 region: Cyprus, Estonia, Latvia, Lithuania, Luxembourg and Malta. Only NUTS 1 level data is available for the following countries: Belgium, Germany, Greece, Hungary, Poland, the Netherlands, Bulgaria, Romania and Ireland (in the latter case the NUTS 1 region comprises the whole country). Croatia and Portugal have not been included in the dataset due to non-availability of data. Some regions have been excluded completely from the analysis since they are not situated in Europe in geographical terms or due to lack of data: these are in the case of France the overseas regions of Guadeloupe, Martinique, Guyane and

Réunion and the island of Corsica (FR91-FR94, FR83), in the case of Spain the autonomous cities of Ceuta and Melilla and the Canarias (ES63, ES64, ES70) and the small region of Åland (FI20) in the case of Finland.

Regional data have been collected for the following social indicators that were used individually as dependent variables in the subsequent regression analysis:

Population health indicators: Life expectancy at birth, infant mortality rate (Number of deaths of children <1 year of age per thousand live births in the same year) and standardised death rates (age structure adjusted): assault, drug dependence, diseases of the circulatory system and mental diseases. No data on standardised death rates were available for the regions of Denmark.

Crime indicators: Homicide rates and robbery rates for violent crime and rates of domestic burglary and theft of motor vehicles for property crime. None of the crime rates were available for the regions of Greece and the United Kingdom. In the case of the Netherlands no data on homicide rates were available.

Non-participation of youngsters in the labour market and education: Share of youngsters aged 15-24 not in employment, education or training in the population of the same age; Rate of early leavers from education (percentage of the population aged 18 to 24 having attained at most lower secondary education and not being involved in further education or training).

In order to characterise the level of income inequality in the EU regions we apply three different indicators, first the Gini coefficient, which is most sensitive to inequalities in the middle part of the income spectrum, second the (at-risk-of) poverty rate focusing on the dispersion between low and medium income earners and the income quintile share ratio highlighting the dispersion between low and high income earners.

Gini coefficient: Measure of dispersion of equivalised disposable income in the reference population. A Gini coefficient of 0 would denote total equality of income; a Gini coefficient of 100 total inequality, i.e. one person would accrue all income received by the population.

Poverty rate: Share of persons in the population with an equivalised disposable income below 60% of the national median income.

Income quintile share ratio (S80/S20): Ratio of total disposable equivalised income received by the 20% of the region’s population with the highest income (top quintile) to that received by the 20% of the regions's population with the lowest income (lowest quintile).

Certainly, it would be useful to apply in addition to the Gini coefficient and the poverty rate an inequality measure that is sensitive to inequalities at the upper part of the income distribution like a Generalized Entropy indicator. However, the choice was limited by the availability of data.

In addition to inequality measures regional income levels are used as an explanatory control variable, precisely the regional level of GDP per capita based on purchasing power parity. Additional reasonable explanatories at the regional level were not available except for doctors per thousand inhabitants. However, this variable turned out to be insignificant in the regressions.

In order to get more robust indicators we calculated three year averages wherever possible, mainly of the years 2009-2011. However, in the case of the Gini coefficient and the income quintile share ratio only data for the year 2010 was available. In the case of non-availability of NUTS 2 level inequality indices for a country from the Eurostat database or OECD regional well-being dataset we calculated those based on EU SILC microdata where possible. A detailed list of data sources and time periods used for the calculations of variables is provided for in the Appendix Tables A.1 and A.2.

4. Methodology

In order to analyse the unconditional correlations of the inequality measures and social outcome variables one can start with simple Ordinary least squares regressions (OLS):

$$y_i = \alpha + \beta x_i + \varepsilon_i; \qquad \varepsilon_i \sim N(0, \sigma_y^2).$$

The best approach to analyse the true effect of income inequality on social outcome indicators would then be to include as many variables as possible that capture additional effects on the dependent variables into the regression model. However, in our case apart from GDP p.c. further variables (e.g. on the infrastructure in areas like health, schooling, etc.) are not available at the same, i.e. regional level. The only additional variable to be found was doctors per thousand inhabitants. However, the inclusion of this explanatory did not lead to significant coefficients for this variable.

Applying the simple OLS regression model will most probable lead to biased estimates of β since the observations for regions i are grouped into units (i.e. countries) j.

$$y_i = \alpha + \beta x_{j[i]} + \varepsilon_i; \quad \varepsilon_i \sim N(0, \sigma_y^2)$$

Most commonly, the unit effects α_j are associated with x, so variation in α_j must be modelled in order to avoid faulty inferences about β. Two standard approaches for modelling variation in α_j are the fixed effects and the random effects model. In the case of the latter the average unit effect is estimated by μ_α and σ_α^2 describes by how much the individual unit effects vary around that value:

$$y_i = \alpha_{j[i]} + \beta x_i + \varepsilon_i; \qquad \alpha_j \sim N(\mu_\alpha, \sigma_\alpha^2); \qquad \varepsilon_i \sim N(0, \sigma_y^2)$$

In the case of the fixed effects specification it is assumed that the intercepts are distributed with infinite variance:

$$y_i = \alpha_{j[i]} + \beta x_i + \varepsilon_i; \qquad \alpha_j \sim N(\mu_\alpha, \infty); \qquad \varepsilon_i \sim N(0, \sigma_y^2).$$

Thus with the application of one of these two models we can control for variables that would explain country differences but are not included in the model and can thus reduce the omitted-variable bias of β. The choice between fixed or random effects models is one between bias and variance. The fixed effects model will produce unbiased estimates of β, but those estimates can be subject to high sample-to-sample variability. The random effects model will, except in rare circumstances, introduce bias in estimates of β, but can greatly constrain the variance of those estimates, leading to estimates that are closer, on average, to the true value in any particular sample. In order to guide our decision of model choice we applied for all individual regressions the Hausman test for model specification. All regression models have been tested for multicollinearity by analysing matrices of pairwise correlations. No incidence of multicollinearity has been detected.

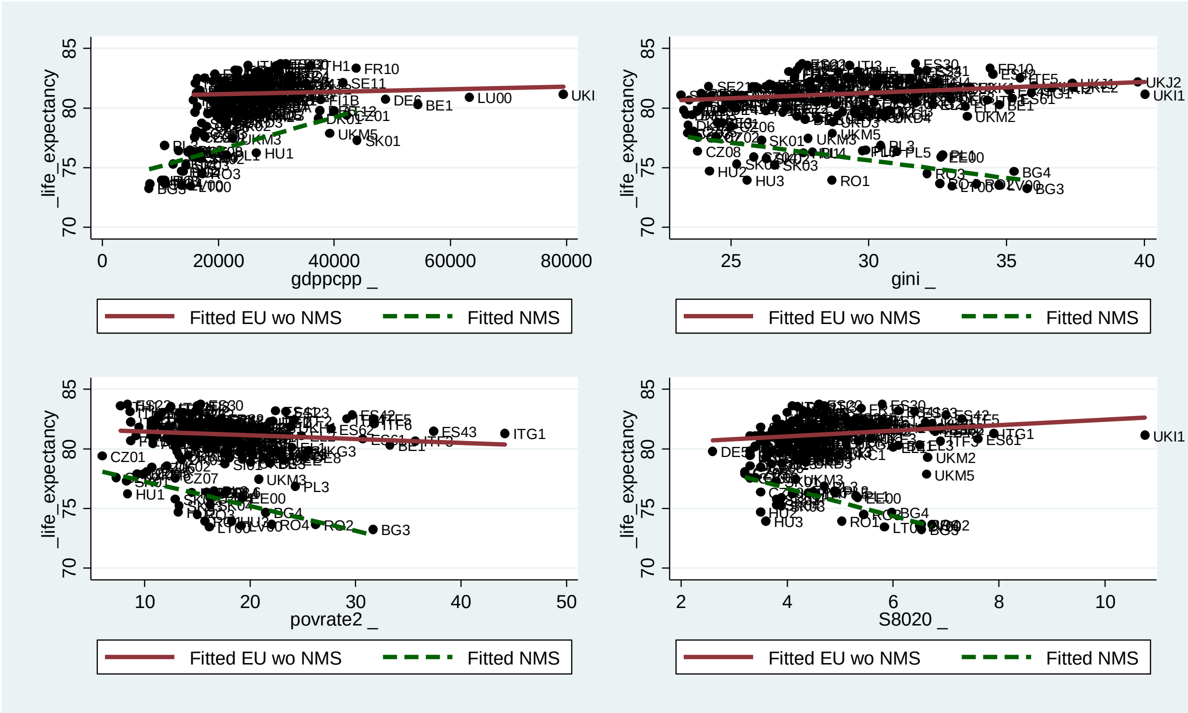
5. Results

In this section we report the results of the empirical analysis for all indicators of population health and social outcomes collected. For a number of indicators the GDP level and inequality indicators however have low explanatory power for the variations of the variables to be explained. All fixed effects and random effects regression results are presented in the Appendix Tables A.3 to A.14, while in the text only the results of those regression models are reported that were chosen following the Hausman tests. The results of the Hausman tests are reported in the Appendix Table A.15.

5.1. Life expectancy

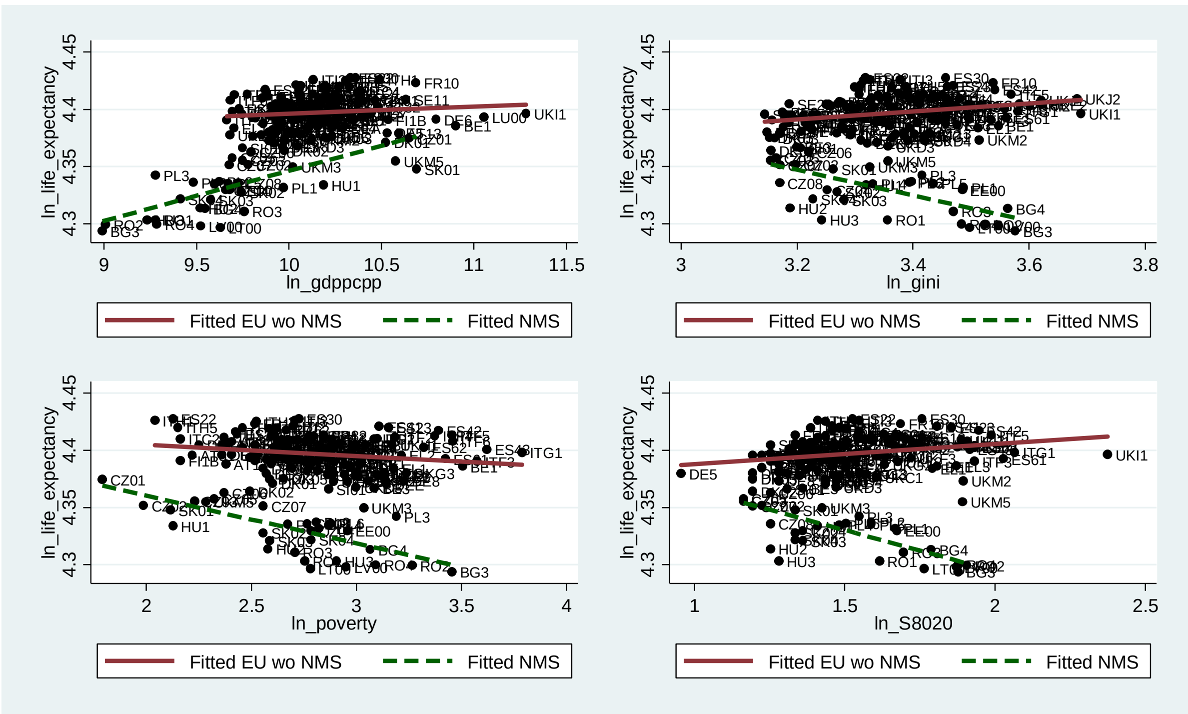
Particularly Deaton (2003) has shown the non-linear relationship between GDP per capita and life expectancy. Given this non-linear relationship, the absolute income hypothesis concerning the relationship between income inequality and health would suggest that also within countries or regions transfers from high-income earners to low-income earners would increase the health situation or life expectancy of low-income earners more than the possible loss of the high-income earners. Thus the average conditional life expectancy should be higher in regions with lower income inequality. In Figure 1 we start drawing scatter plots using the NUTS 2 raw data (and NUTS 1 data in the case of BE, DE, EL, HU, PL, NL, BG, RO and IE) for life expectancy and our explanatories GDP per capita, Gini coefficient, poverty rate and the income quintile share ratio (S80/S20). First we can see that average life expectancy at birth ranges between 73.2 years for Northern and Eastern Bulgaria (BG3) and 83.7 years for Madrid (ES30). GDP per capita spreads from a minimum of only about EUR 8 thsd. (at PPP) for Northern and Eastern Bulgaria (BG3) to the tenfold of EUR 80 thsd. (at PPP) for Inner London (UKI1). The Gini coefficient ranges from 23.2 for the Swedish region of Upper Norrland (SE33) to 40 for Inner London, the poverty rate from 6% for Prague (CZ01) to 44.2% for Sicily (ITG1) and the S80/S20 ratio from 2.6 for Bremen (DE5) to 10.8 again for Inner London. Looking at the graphs one can see that most of the regions situated in the Central and East European New EU Member States (CEE NMS) feature not only a lower GDP p.c. level but also a lower level of live expectancy at birth. Thus we split the sample into two subgroups, the NMS regions and the Non-NMS regions and draw two separate linear prediction plots. Since in the literature the relationship between income levels and life expectancy is described to be concave we transformed the data into natural logs. Moreover, when taking the logs of the variables the regression results are easier to interpret and in our regression analysis it appeared that the fit also improved. The scatter plots with logs of the variables are shown in Figure 2. From graphical inspection we suppose that most probably the level of GDP per capita, the Gini coefficient and the S80/S20 ratio might not be useful explanatory variables in the case of the regions in the Non-NMS EU group, while for the NMS group correlations between life-expectancy and all inequality explanatories might be indicative. One of the reasons why the relationships are different in the two country groups may be the lower levels of health expenditures in the NMS countries not only per capita at PPP but also as a share of GDP. Drawing on national data provided by Eurostat one can observe that in the NMS countries on average 7% of the GDP is spent in total on health care, while in the Non-NMS countries the share amounts to 9.4%. In addition in the Non-NMS countries about 25% of total health care expenditure is spent by the private sector, while in the NMS countries this amounts to almost 30%.

Figure 1. Scatter plots: Life expectancy versus GDP per capita at PPP and inequality indicators



Source: Eurostat, OECD regional well-being dataset, EU SILC, own calculations.

Figure 2. Scatter plots: Life expectancy versus GDP p.c. at PPP and inequality indicators (in logs)



Source: Eurostat, OECD regional well-being dataset, EU SILC, own calculations.

Table 1. Regression results for life expectancy (in logs)

VARIABLES	EU regions excluding CEE NMS	CEE NMS regions

ln_gdppcpp	0.0144***	0.00646	0.0162***	0.0291***	0.00788	0.0272***
	(0.00401)	(0.00429)	(0.00401)	(0.00506)	(0.00728)	(0.00466)
ln_gini	0.0102			-0.0609**		
	(0.0103)			(0.0262)		
ln_poverty		-0.0102***			-0.0245**	
		(0.00351)			(0.00913)	
ln_S8020			-0.00357			-0.0444***
			(0.00574)			(0.0154)
Constant	4.217***	4.357***	4.238***	4.251***	4.321***	4.133***
	(0.0453)	(0.0496)	(0.0391)	(0.0893)	(0.0914)	(0.0490)
Observations	149	149	148	32	32	32
Number of countries	16	16	15	10	10	10
R-squared within	0.117	0.163	0.113	0.524	0.657	0.566
R-squared between	0.0059	0.0489	0.0493	0.658	0.329	0.656
R-squared overall	0.0332	0.0361	0.0039	0.607	0.427	0.626
model	fixed	random	fixed	random	fixed	random

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Source: Eurostat, OECD regional well-being dataset, EU SILC, own calculations.

The results of the regressions for Non-NMS regions and NMS regions are presented in Table 1. The applied regression models are indicated in the last row. In our choice of model we followed the results of the Hausman test that gives an indication which of the estimators is likely to be more efficient (A summary of results of the Hausman tests are to be found in the Appendix Table A.15).

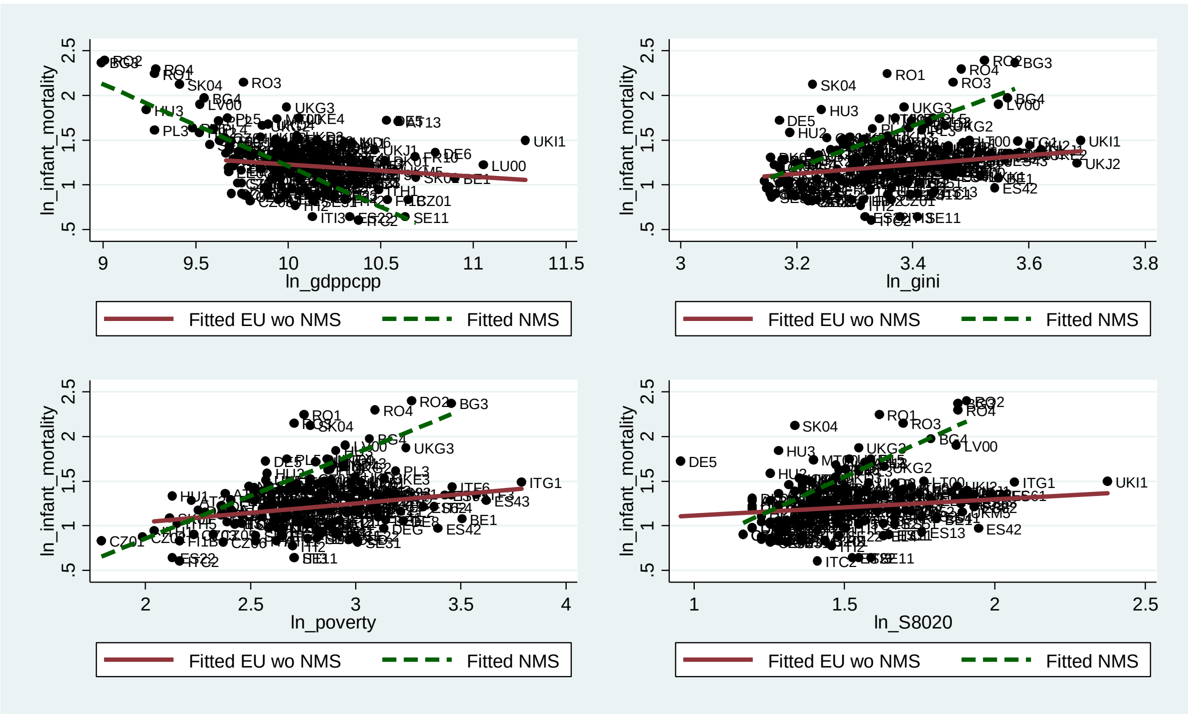
As expected the results for conditional correlations for Non-NMS regions show that we cannot find significant coefficients for the Gini coefficient and the S20/S80 ratio, while for poverty we find a significant negative slope, i.e. life expectancy falls with rising poverty rates. The effect of GDP per capita (at PPP) is positive, however not significant if in the regressions where inequality is described by the poverty rate. However, the explanatory power of the regression models as stated by the overall R-squared is quite weak. This does not mean that the relation between poverty and life expectancy is non-existent. But obviously the phenomena that influence the inter-regional differences in years of life expectancy are complex and we can explain with our independent variables only a rather small part of the dependent variable variation.

In the case of the NMS regions the explanatory power of the regressions is much higher, which we would have expected looking at the scatter plots in Figure 2. An increase of the GDP per capita (at PPP) by 1% is correlated with an increase of life expectancy by 0.02%, while an increase of the Gini coefficient by 1% is correlated with a fall of life expectancy of 0.06%. Thus, tough the size of the coefficients is small, these are nevertheless not negligible. A 10% rise of GDP is associated with an increase of the average life expectancy of about a quarter of a year, while a 10% rise in the Gini with an increase of about half a year for regions with an average life expectancy in the NMS country group. The GDP p.c. is positively correlated with life expectancy, with a slope being steeper compared to the regressions for the Non-NMS regions, but again not significant in the case of the regression including the poverty rate as explanatory. Both the poverty rate and the S80/S20 ratio are negatively correlated as expected with life expectancy.

5.2. Infant mortality

Another health indicator which is often used in the literature testing the ‘income inequality-health hypothesis’ is the infant mortality rate. In our case we applied the mortality rates of children below the age of 1 year (death of children < 1 year per thousand born children in the same year). In the EU regions these rates range from 1.8 for the Italian Aosta valley (ITC2) to 11 in the North- and South-East NUTS 1 region of Romania (RO2). Graphical inspection of the logarithmised data shown in the scatter plots in Figure 3 leads to the hypothesis that infant mortality rates are negatively correlated with GDP per capita and positively correlated with all three inequality indices. Moreover, we assume that the relationships are similar in both groups of EU regions, the NMS and the Non-NMS. However, in the group of NMS regions the slopes of the regression lines are much steeper, although the relationship in this country group might be driven exclusively by Bulgarian and Romanian regions.

Figure 2. Scatter plots: Infant mortality versus GDP p.c. at PPP and inequality indicators (in logs)



Source: Eurostat, OECD regional well-being dataset, EU SILC, own calculations.

Table 2. Regression results for infant mortality (< 1 year) (in logs)

VARIABLES	EU regions	EU regions exluding CEE NMS	CEE NMS regions

ln_gdppcpp	-0.291***	-0.00171	-0.280***	-0.129**	0.0821	-0.124*	-0.505***	-0.208	-0.490***
	(0.0542)	(0.0669)	(0.0541)	(0.0640)	(0.0705)	(0.0645)	(0.115)	(0.171)	(0.102)
ln_gini	0.631***			0.454***			0.877		
	(0.159)			(0.163)			(0.738)		
ln_poverty		0.277***			0.262***			0.365	
		(0.0581)			(0.0574)			(0.215)	
ln_S8020			0.353***			0.227**			0.781*
			(0.0907)			(0.0928)			(0.435)
Constant	2.126***	0.511	3.602***	0.985	-0.374	2.119***	3.471	2.545	5.100***
	(0.683)	(0.778)	(0.537)	(0.751)	(0.816)	(0.643)	(2.207)	(2.150)	(1.036)
Observations	183	183	182	151	151	150	32	32	32
Number of countries	26	26	25	16	16	15	10	10	10
R2 within	0.113	0.182	0.115	0.0633	0.150	0.0572	0.499	0.531	0.538
R2 between	0.486	0.159	0.473	0.0463	0.0008	0.0123	0.657	0.572	0.649
R2 overall	0.320	0.114	0.288	0.0915	0.0714	0.0521	0.684	0.563	0.727
model	random	fixed	random	random	random	random	fixed	fixed	fixed

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Source: Eurostat, OECD regional well-being dataset, EU SILC, own calculations.

The regression results presented in Table 2 show that particularly in the NMS regions infant mortality tends to rise strongly with lower GDP p.c. levels. In the rest of the EU regions with higher inequality levels tend to have also higher infant mortality rates, while there are no significant results for the NMS regions in this respect. The explanatory power of the regression models is satisfactory in the case of the NMS regions, while quite low in the case of the Non-NMS regions.

5.3. Standardised death rates by causes

In Table 3 we report the regression results for standardised (age adjusted) death rates for various causes of death. One of the problems dealing with cause-specific death rates however may be that comorbidity introduces a bias in the reported figures. In addition different national health systems may follow diverse report strategies. In the case of **assault** both for the NMS regions and the Non-NMS regions death rates tend to rise with higher inequality levels. However, only for the regressions including the Gini coefficient and the S80/S20 ratio the explanatory power is satisfactory in the case of the NMS regions. For the rate of death due to **drug dependence** (and toxicomania) we get significant results (for a regression model with low explanatory power) only in the case of the poverty rate and the S80/S20 ratio for the Non-NMS regions. However, the coefficients have a non-expected sign, i.e. deadly drug use tends to be higher in regions with lower income inequality. For death due to **diseases of the circulatory system** (including particularly heart attack) rates tend to rise with increasing poverty and the S80/S20 ratio in the case of Non-NMS regions and poverty in the case of NMS regions. In all regression models (which however do not have high explanatory power) death rates tend to fall with higher GDP p.c. levels. For **mental diseases** significant coefficients are only to be found in the case of the Non-NMS regions. However, the signs for the inequality indices are non-expected, i.e. deathrates rise with falling inequality levels.

Table 3. Regression results for standardised (age adjusted) death rates (in logs)

VARIABLES	EU regions			EU regions excluding CEE NMS			CEE NMS regions		
Dependent variable: Deathrate assault (in logs)									
ln_gdppcpp	-0.190	0.409***	-0.172	-0.065	0.439***	-0.077	-0.249	0.59***	-0.158
ln_gini	1.267***			0.999*			2.78***		
ln_poverty		0.631***			0.613***			1.03***	
ln_S8020			0.869***			0.725***			1.91***
Constant	-2.305*	-6.399***	0.477	-3.013	-6.514***	-0.604	-6.407**	-7.96***	-0.859
R2 within	0.0587	0.155	0.0968	0.0643	0.148	0.102	0.0173	0.248	0.0608
R2 between	0.241	0.0511	0.241	0.0719	0.117	0.0724	0.555	0.136	0.567
R2 overall	0.00509	0.00365	0.0173	0.0217	0.0509	0.00157	0.418	0.185	0.461
model	random	fixed	random	random	random	random	random	random	random
Dependent variable: Deathrate drug dependence, toxicomania (in logs)									
ln_gdppcpp	0.936*	0.422	0.942*	0.754	0.498	0.794	0.402	-0.215	0.382
ln_gini	-0.834			-0.635			-0.628		
ln_poverty		-0.266*			-0.281**			-0.705	
ln_S8020			-0.72***			-0.65***			-0.337
Constant	-8.087	-4.507	-9.858*	-6.642	-5.379	-8.202	-4.223	1.554	-5.641*
R2 within	0.0541	0.0575	0.0694	0.0580	0.0618	0.0737	0.0277	0.0775	0.0303
R2 between	0.423	0.418	0.412	0.0932	0.121	0.115	0.189	0.269	0.184
R2 overall	0.117	0.105	0.135	0.0272	0.0428	0.0493	0.191	0.224	0.183
model	random	fixed	random	random	random	random	random	random	random
Dependent variable: Diseases of the circulatory system (in logs)									
ln_gdppcpp	-0.23***	-0.16***	-0.23***	-0.24***	-0.17***	-0.25***	-0.22***	-0.12***	-0.21***
ln_gini	0.040			0.035			0.248		
ln_poverty		0.090***			0.090***			0.10***	
ln_S8020			0.075*			0.071**			0.195
Constant	8.188***	7.429***	8.272***	8.266***	7.490***	8.360***	7.97***	7.56***	8.44***
R2 within	0.409	0.463	0.423	0.387	0.443	0.404	0.538	0.594	0.548
R2 between	0.580	0.507	0.554	0.000763	0.00257	1.77e-05	0.540	0.441	0.504
R2 overall	0.262	0.173	0.237	0.00103	0.0169	3.27e-06	0.402	0.302	0.417
model	fixed	fixed	fixed	random	random	random	random	fixed	random
Dependent variable: Mental diseases (in logs)									
ln_gdppcpp	0.053	-0.138	0.058	0.170	-0.067	0.185*	-0.458**	0.152	-0.443**
ln_gini	-0.468**			-0.544***			0.577		
ln_poverty		-0.178			-0.259***			0.828*	
ln_S8020			-0.392***			-0.454***			0.439
Constant	3.400***	4.775***	2.336**	3.258**	4.579***	1.957*	3.591	-2.639	4.717**
R2 within	0.0158	0.0225	0.0378	0.0627	0.109	0.127	0.0542	0.119	0.0578
R2 between	0.0740	0.239	0.0890	0.130	0.0363	0.155	0.0524	0.0375	0.0434
R2 overall	0.0109	0.145	0.0148	0.0238	0.00735	0.0354	0.0223	0.0221	0.0154
model	random	fixed	random	random	random	random	random	random	random

*** p<0.01, ** p<0.05, * p<0.1

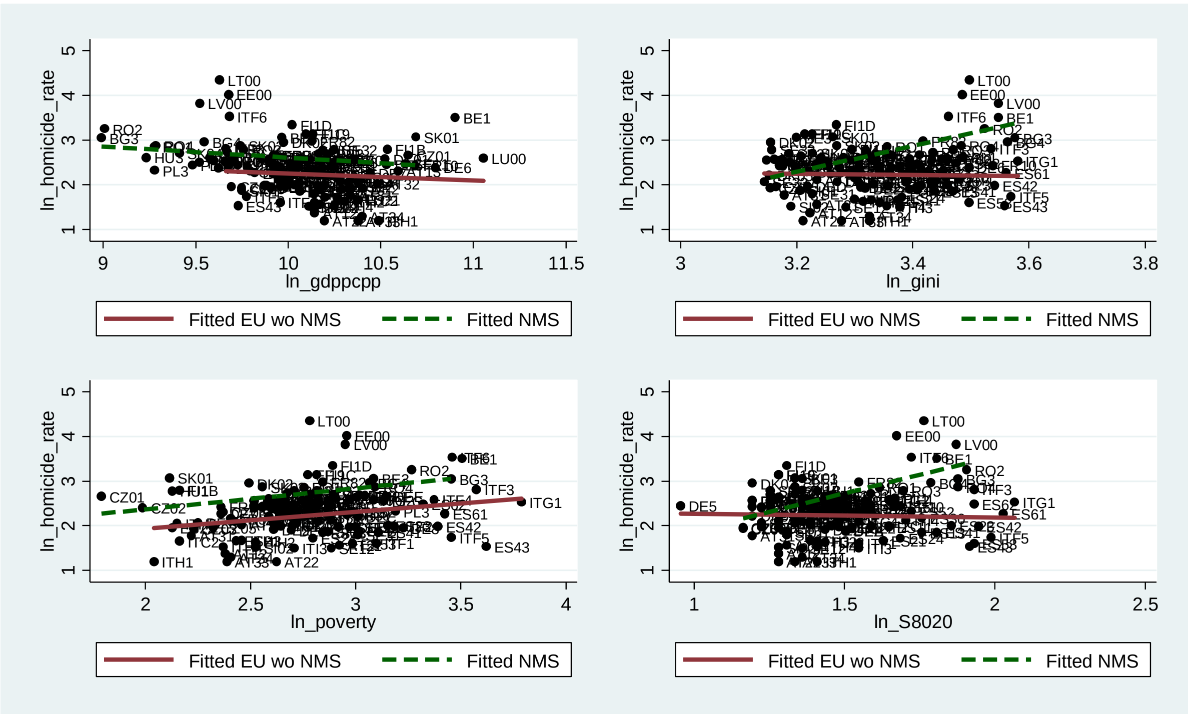
Source: Eurostat, OECD regional well-being dataset, EU SILC, own calculations.

5.4. Crime indicators: Violent and property crime

In general crime rates are objective measures, however the data collected is that of reported crime and not that of incidents executed. People might deter from notifying the incident if they expect the crime not to be solved, or if the cost of reporting is high compared to the loss caused. Homicide and robbery rates tend to be more reliable figures compared to other types of crime since the violence associated with such criminal acts tends to increase the proclivity for the victim to officially declare the crime to the police (see e.g. Fajnzylber et al, 2002). However, apart from data for violent crime we also use property crime rates for domestic burglary and theft of motor vehicles for the analysis below. No data on crime are available for Greece and the United Kingdom at the regional level. In the case of the Netherlands data on homicide rates are missing.

Homicide rates (cases per 100 thousand inhabitants per year) range from 33 for the Austrian region of Tyrol (AT33) to 773 for Lithuania. All three Baltic States report particularly high homicide rates. The first non-Baltic region is the Italian region of Calabria (ITF6) with 342 cases. A first hint on the relationship between homicide rates and our explanatories for income levels and inequality is presented by the scatter plots in Figure 4. We would expect that homicide rates fall with an increase of GDP p.c. and rise with the increase of poverty rates both in the Non-NMS and NMS regions of the EU. In the case of the Gini index and the S80/S20 ratio we would guess to find no correlation for the Non-NMS regions and a positive relation for the NMS regions.

Figure 4. Scatter plots: Homicide rates versus GDP p.c. at PPP and inequality indicators (in logs)



Source: Eurostat, OECD regional well-being dataset, EU SILC, own calculations.

Table 4. Regression results for homicide rates (in logs)

VARIABLES	EU regions excluding CEE NMS			CEE NMS regions		
ln_gdppcpp	-0.301*	0.170	-0.241	0.0143	0.475***	0.0962
	(0.160)	(0.191)	(0.158)	(0.116)	(0.184)	(0.100)
ln_gini	0.805*			2.546***		
	(0.422)			(0.652)		
ln_poverty		0.499***			0.452**	
		(0.134)			(0.230)	
ln_S8020			0.592**			1.632***
			(0.247)			(0.376)
Constant	2.776	-0.749	3.981**	-5.773***	-2.839	-0.489
	(2.028)	(2.203)	(1.688)	(2.128)	(2.342)	(1.056)
Observations	107	107	106	32	32	32
Number of countries	13	13	12	10	10	10
R-squared within	0.0957	0.149	0.118	0.322	0.241	0.418
R-squared between	0.0575	0.0960	0.0966	0.567	0.0223	0.540
R-squared overall	0.0007	0.0786	0.0000	0.452	0.0412	0.454
model	random	random	random	random	random	random

Standard errors in parentheses

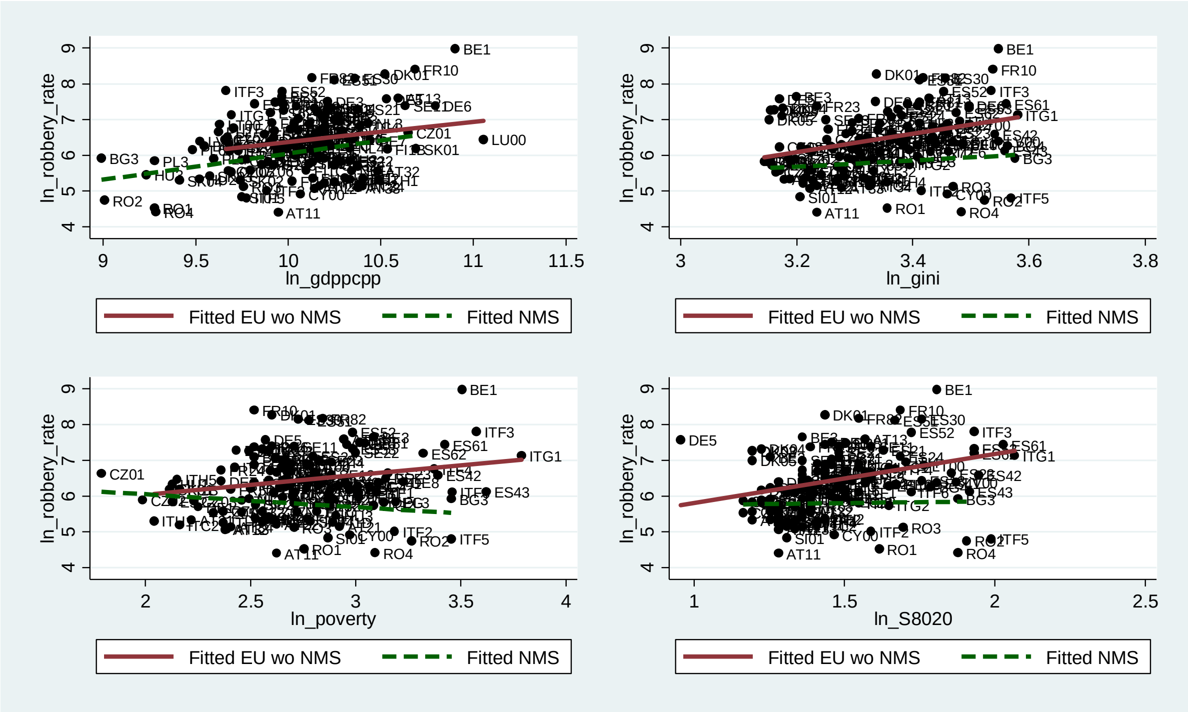
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Source: Eurostat, OECD regional well-being dataset, own calculations.

The regression results presented in Table 4 show that for the Non-NMS regions the explained variation between country groups and within those is rather low. Nevertheless in the case of the S80/S20 ratio and the poverty rate homicide rates tend to rise with an increase of income inequality. The explanatory power of the regressions for the NMS regions is much higher for the Gini coefficient and the S80/S20 ratio. In the case of the Gini indicator a rise of the Gini coefficient by 1% is correlated with an increase of the homicide rate by 2.5%, which corresponds to about 4 deaths per year per 100 thousand inhabitants for the average NMS region. A similar relationship is found in the case of the S80/S20 ratio. Also in the case of poverty rates a positive significant correlation could be detected, however the explanatory power of the regression model is quite low.

Robbery rates (cases per 100 thousand inhabitants per year) range from 812 for the Austrian region of Burgenland (AT11) to 79 thousand for the region of Brussels (BE1). The scatter plots presented in Figure 5 give an indication of the potential relationships between robbery rates in the EU regions and our explanatories. The slopes of the regression lines show that robbery rates in general tend to rise with increasing income level and ascend with higher inequality levels in the Non-NMS regions. A reasonable explanation for robbery rates rising with GDP p.c. levels is that wealthier societies tend to possess not only absolutely but in addition relatively more valuables and durables. The unconditional correlations for the NMS regions however would lead to the assumption that there is no relationship to be found between inequality indicators and robbery rates.

Figure 5. Scatter plots: Robbery rates versus GDP p.c. at PPP and inequality indicators (in logs)



Source: Eurostat, OECD regional well-being dataset, EU SILC, own calculations.

Table 5. Regression results for robbery rates (in logs)

VARIABLES	EU regions			EU regions excluding CEE NMS			CEE NMS regions		
ln_gdppcpp	0.811***	1.632***	0.997***	0.781***	2.085***	1.065***	0.693***	1.080***	0.752***
	(0.180)	(0.221)	(0.179)	(0.269)	(0.325)	(0.262)	(0.167)	(0.272)	(0.154)
ln_gini	3.068***			3.396***			2.077***		
	(0.578)			(0.709)			(0.793)		
ln_poverty		1.081***			1.282***			0.438	
		(0.191)			(0.225)			(0.335)	
ln_S8020			1.988***			2.219***			1.275**
			(0.340)			(0.409)			(0.504)
Constant	-12.1***	-13.1***	-6.71***	-12.9***	-18.4***	-7.67***	-7.82***	-5.730*	-3.342**
	(2.546)	(2.568)	(1.912)	(3.382)	(3.740)	(2.791)	(2.857)	(3.425)	(1.624)
Observations	143	143	142	111	111	110	32	32	32
Nr of countries	24	24	23	14	14	13	10	10	10
R2 within	0.300	0.329	0.318	0.289	0.335	0.308	0.570	0.554	0.598
R2 between	0.0910	0.143	0.166	0.0112	0.0551	0.0528	0.431	0.0135	0.259
R2 overall	0.210	0.237	0.219	0.132	0.141	0.157	0.267	0.231	0.174
model	random	random	random	random	random	random	random	random	random

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Source: Eurostat, OECD regional well-being dataset, EU SILC, own calculations.

Regression results presented in Table 5 show that the conditional relationships between robbery rates and the explanatories are quite similar in the Non-NMS and the NMS regions. Incidents of robbery tend to be more prevalent in regions with higher average income levels. The coefficients for income inequality are higher in the Non-NMS regions compared to the NMS regions. A rise of the Gini-index of 1% is associated with an increase of 3.4% of the crime rate in the Non-NMS regions and 2.1% in the NMS regions. These percentage increases correspond to a rise of about 320 incidences in the case of Non-NMS regions and 83 cases for NMS regions per 100 thousand inhabitants. Also the S80/S20 ratio is significantly positive correlated with robbery rates, while in the case of poverty the positive relationship is significant only in the case of Non-NMS regions.

For property crime we collected rates of domestic burglary and theft of motor vehicles. The results of the regressions are presented in Table 6. In the case of Non-NMS regions **domestic burglary** rates tend to rise in general with inequality, in the case of NMS regions only with the Gini index and the S80/S20 ratio. All regressions (which have in general only low explanatory power) show that rates of domestic burglary correlate positively with levels of GDP p.c. In the case of **theft of motor vehicles** rates tend to rise with higher levels of inequality only for Non-NMS regions. For NMS regions we find a significant coefficient only for the poverty rate. The sign of the coefficient is however negative, which was not expected, i.e. theft rates tend to be higher in regions with lower poverty rates.

Table 6. Regression results for property crime (in logs)

VARIABLES	EU regions			EU regions excluding CEE NMS			CEE NMS regions		
Dependent variable: Domestic burglary rates (in logs)									
ln_gdppcpp	0.749***	1.137***	0.824***	0.695***	1.228***	0.812***	0.635***	0.600**	0.698***
ln_gini	1.606***			1.586***			1.828**		
ln_poverty		0.472***			0.507***			-0.213	
ln_S8020			0.869***			0.834**			1.122**
Constant	-5.34***	-5.156**	-2.036	-4.559*	-6.141**	-1.708	-5.224**	1.894	-1.394
R2 within	0.228	0.206	0.218	0.183	0.176	0.170	0.637	0.640	0.654
R2 between	0.337	0.277	0.249	0.0876	0.104	0.0562	0.270	0.177	0.157
R2 overall	0.215	0.222	0.197	0.0722	0.0777	0.0612	0.237	0.0201	0.189
model	random	random	random	random	random	random	random	random	random
Dependent variable: Motor vehicle theft rates (in logs)									
ln_gdppcpp	0.264	1.377***	0.452	-0.443**	0.551**	-0.236*	1.709***	1.225***	1.694***
ln_gini	2.826**			2.650***			-0.266		
ln_poverty		1.015***			0.978***			-0.717**	
ln_S8020			1.899***			1.621***			-0.036
Constant	-4.976	-9.74***	-0.234	3.140	-0.928	7.484***	-9.59***	-3.832	-10.2***
R2 within	0.174	0.213	0.216	0.206	0.236	0.224	0.628	0.636	0.629
R2 between	0.0413	0.255	0.0625	0.00364	0.00373	0.00350	0.433	0.529	0.449
R2 overall	0.00360	0.216	0.00556	0.0254	0.0627	0.0148	0.557	0.630	0.544
model	fixed	random	fixed	random	random	random	random	random	random

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Source: Eurostat, OECD regional well-being dataset, EU SILC, own calculations.

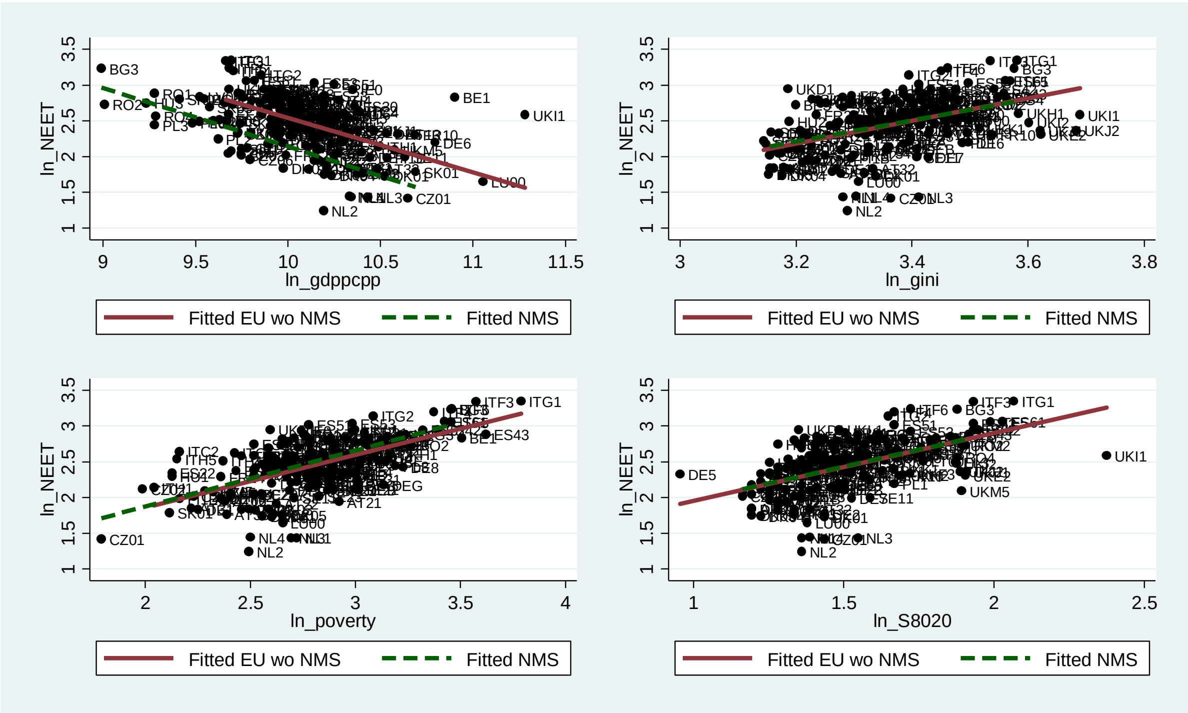
5.5. Youngsters: Non-participation in education and employment

One of the reasons for regions with higher inequality levels to feature lower participation rates in education might be the effect of intergenerational transmission of educational choice. Children whose parents have relatively low incomes (due to unemployment or due to wage inequality) could be discouraged from own investments in education or might lack of resources for those.

Analysing the relationship between overall educational attainment levels for the total population and income inequality obviously one would have to deal with the presence of endogeneity between the two variables. However in our analysis we would not expect the share of youngsters being non-active according to labour status or not in education to influence the level of income inequality of the total population in a region.

Shares of **youngsters not in employment, education or training - NEET (age 19-24)** range from 3.5% for the Eastern Netherlands (NL2) to 28.5% for the Italian region of Sicilia (ITG1). From the scatter plots presented in Figure 6 we expect that the relationship between NEET rates and our explanatories is quite similar in the Non-NMS and NMS regions. An unconditional negative correlation can be observed between income level and NEET rates. Higher NEET rates tend to appear with increased inequality levels.

Figure 6. Scatter plots: NEET rates (age 15-24) versus GDP p.c. at PPP and inequality indicators (in logs)



Source: Eurostat, OECD regional well-being dataset, EU SILC, own calculations.

Table 7. Regression results for rates of youngsters NEET – aged 15-24 (in logs)

VARIABLES	EU regions excluding CEE NMS			CEE NMS regions		
ln_gdppcpp	-0.494***	-0.0728	-0.507***	-0.721***	-0.527***	-0.693***
	(0.0743)	(0.0672)	(0.0726)	(0.110)	(0.179)	(0.112)
ln_gini	0.598***			0.849***		
	(0.186)			(0.307)		
ln_poverty		0.516***			0.316	
		(0.0529)			(0.195)	
ln_S8020			0.422***			0.486**
			(0.103)			(0.191)
Constant	5.434***	1.738**	6.942***	6.552***	6.675***	8.397***
	(0.836)	(0.771)	(0.709)	(1.686)	(2.177)	(1.210)
Observations	150	150	149	32	32	32
Number of countries	16	16	15	10	10	10
R-squared within	0.259	0.536	0.291	0.575	0.613	0.576
R-squared between	0.448	0.502	0.594	0.842	0.772	0.823
R-squared overall	0.437	0.366	0.511	0.703	0.663	0.698
model	fixed	fixed	fixed	random	random	random

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

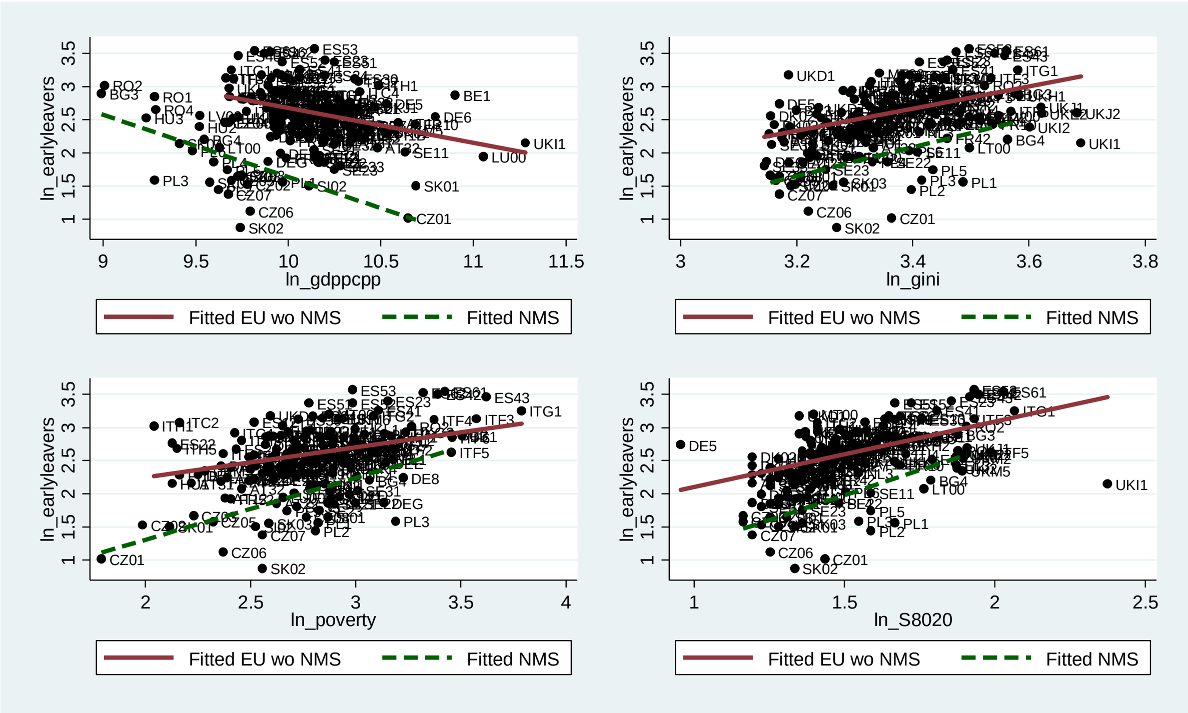
Source: Eurostat, OECD regional well-being dataset, own calculations.

In general the explanatory power of the regression models presented in Table 7 is quite good and particularly strong in the case of the NMS regions. NEET rates tend to rise with falling GDP p.c. levels. Inequality levels are conditionally correlated positively with the dependent variable; the coefficient of the Gini index is significantly higher in the NMS regions compared to the Non-NMS regions. A rise of the Gini coefficient of 10% is correlated with an increase of 6% in the Non-NMS region which corresponds on average to a rise of the NEET rate of 0.7 percentage points. In the case of the NMS regions a 10% increase of the Gini is expected to lift the NEET rate by 0.8% which corresponds to a rise by 1 percentage points on average. The relationship is similar for the S820/S20 ratio and for the poverty rate in the case of the Non-NMS regions, while the positive coefficient for the latter indicator is non-significant for the NMS regions.

The **rates of early leavers from education (age 18-24)** range from 2.4% for the region of Western Slovakia (SK02) to 35.5% for the Spanish Balearic islands (ES53). The scatter plots presented in Figure 7 show that according to unconditional correlations the relationship between our explanatories and the rate of early leavers from education might also be quite similar for the Non-NMS and NMS regions. The rates tend to fall with income and rise with inequality levels. However, the rates are lower in the

NMS regions at the same level of income but also the same level of inequality expressed by all three indicators, the Gini, the poverty rate and the S80/S20 ratio.

Figure 7. Scatter plots: Rate of early leavers from education (age 18-24) versus GDP p.c. at PPP and inequality indicators (in logs)



Source: Eurostat, OECD regional well-being dataset, EU SILC, own calculations.

Table 8. Regression results for rates of early leavers from education – aged 18-24 (in logs)

VARIABLES	EU regions excluding CEE NMS			CEE NMS regions		
ln_gdppcpp	-0.257***	-0.0529	-0.240***	-0.655***	-0.141	-0.594***
	(0.0827)	(0.0937)	(0.0834)	(0.188)	(0.317)	(0.186)
ln_gini	0.512**			1.411**		
	(0.207)			(0.643)		
ln_poverty		0.214***			0.462	
		(0.0737)			(0.397)	
ln_S8020			0.209*			1.015***
			(0.119)			(0.376)
Constant	3.491***	2.547**	4.726***	3.615	2.071	6.225***
	(0.930)	(1.075)	(0.815)	(2.958)	(3.980)	(2.005)
Observations	150	150	149	32	32	32
Number of countries	16	16	15	10	10	10
R-squared within	0.0863	0.102	0.0658	0.174	0.260	0.196
R-squared between	0.540	0.431	0.572	0.650	0.542	0.665
R-squared overall	0.335	0.149	0.348	0.479	0.389	0.526
model	fixed	fixed	fixed	random	fixed	random

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Source: Eurostat, OECD regional well-being dataset, own calculations.

The regression results presented in Table 8 show that in both groups of regions income levels are negatively correlated with rates of early leavers from education. The slope is steeper in the case of the NMS regions, while the coefficients are not significant in the regressions containing the poverty rate as explanatory. All inequality indicators are positively correlated with the social indicator as expected. However in the case of the S80/S20 ratio the significance level is quite low and no significance can be found in the case of the poverty rate for the NMS regions.

6. Summary and Conclusions

In this paper we analysed the correlations between income inequality and various indicators of population health and social outcomes on the level of EU regions. As control variable we applied the level of GDP p.c. at EUR purchasing power parities. For the majority of countries we could make use of NUTS 2 level data, for a couple of countries (Belgium, Germany, Greece, Hungary, Poland, the Netherlands, Bulgaria, Romania and Ireland) we had to resort to NUTS 1 level data. Income inequality was measured by three different indicators, the Gini coefficient, which is most sensitive to inequalities in the middle part of the income spectrum, the (at-risk-of) poverty rate focusing on the dispersion between low and medium income earners and the income quintile share ratio highlighting the dispersion between low and high income earners. Due to the limited time span for which inequality indicators at the regional EU level were available we were not able to perform a panel data analysis, but had to resort to a cross-region regression analysis. In order to control for unobserved country characteristics we applied fixed effects or random effects regression models. Our choice between those two was guided by the application of the Hausman test for each individual regression. We performed regressions for the whole group of EU regions, for the group of regions of the Central and East European New EU Member States (NMS) and for the group of regions excluding the NMS (Non-NMS regions).

Our analysis shows that indices of income inequality conditional on GDP p.c. levels show significant correlations with a number of social indicators. For an overview of coefficients of inequality indicators in the regressions see Table 9 below. For life expectancy at birth, infant mortality rates, two standardised death rates (assault, diseases of the circulatory system – heart attack), homicide rates, robbery rates, rates of domestic burglary, rates of youngsters (age 15-24) not in employment, education or training (NEET) and rates of early leavers (age 18-24) from education we found significant results for the hypothesis that higher inequality levels tend to lead to a worsening of social outcome variables. The level of R-squared and thus the explained variation of many of the regressions is however quite low for robbery and domestic burglary rates and for the Non-NMS regions also for life expectancy, infant mortality and homicide rates. This does not mean that in those cases the relation between income inequality and the latter social indicators is non-existent. But obviously the phenomena that influence the inter-regional variations are rather complex and we can explain only a small part of the variations of the dependent variables with our explanatories. The results could be in those cases more vulnerable to change if additional explanatory variables were included. The results often differ in magnitude for the NMS regions and Non-NMS regions; however, the direction of the relationship between inequality and social outcomes is almost always the same. No significant results could be found for infant mortality rates and age specific death rates for the specific causes drug dependence and mental diseases in the case of the NMS regions. The coefficients for the latter two rates show significant negative signs in the case of the Non-NMS regions, which is counter-intuitive, i.e. death rates tend to be higher in regions with more equally distributed income. One of the problems using death rates for the analysis may be that classification strategies are divers in various countries and cases of comorbidity (e.g. heart attack due to drug abuse) are dealt with differently. For the NMS regions we obtained one counter-intuitive result for theft rates of motor vehicles, which correlate negatively with poverty rates.

Table 9. Conditional correlations between social outcomes and inequality indicators (in logs)

Dependent variables	EU regions			EU regions exl. CEE NMS			CEE NMS regions		
	Gini	Poverty	S80/S20	Gini	Poverty	S80/S20	Gini	Poverty	S80/S20
Population health									
Life expectancy		-			-		-	-	-
Infant mortality	+	+	+	+	+	+			
Standardised death rates									
Assault	+	+	+		+	+		+	+
Drug dependence			-		-	-			
Circulatory system		+			+	+		+	
Mental diseases	-		-	-	-	-			
Crime									
Homicide	+	+	+	+	+	+	+	+	+
Robbery	+	+	+	+	+	+	+		+
Domestic burglary	+	+	+	+	+	+	+		+
Theft of motor vehicles	+	+	+	+	+	+		-	
Non-participation in labour market or education									
NEET rates	+	+	+	+	+	+	+		+
Early leavers from education	+	+		+	+		+		+

+/- sign of coefficient

significant coefficient, expected sign, high explanatory power (R2) of regression model

significant coefficient, expected sign, low explanatory power (R2) of regression model

significant coefficient, non-expected sign, high explanatory power (R2) of regression model

significant coefficient, non-expected sign, low explanatory power (R2) of regression model

Conditional significant correlations with satisfactorily high explanatory power for at least two of the three inequality indices have been found for the NMS for life expectancy and homicide rates, for NEET rates and early leavers from education. For the Non-NMS this is the case for NEET rates and the rate of early leavers from education. However as indicated above also for most other social indicators applied we find significant coefficients for the inequality measures.

The analysis shows that redistributive policies aimed at reducing income inequality might lead not only to improved population health but also to general positive spillover effects in the form of lower crime rates and increased activity and participation rates of youngsters in education. The split of the sample into NMS and Non-NMS regions reveals that although the effect of GDP p.c. is mostly stronger for the

NMS regions also the slopes of the conditional correlations of the inequality indicators tend to be steeper. This suggests that for the NMS countries not only absolute growth of GDP levels is expected to lead to better outcomes in population health and other social phenomena. More redistributive policies would most probably lead to improvements particularly in those countries. Concerning population health this is no surprise since total health expenditures as a share of GDP are on average lower in the NMS countries compared to the Non-NMS group. Thus it may very well be even more important in the former countries how the distribution of scarce means looks like. In the case of crime rates we find positive correlations both for violent and property crime (except for theft of motor vehicles in the NMS). We obviously cannot identify if the reasons for that are higher expected relative gains from crime or if the income dispersion leads to lower inhibitions to commit crime. However, higher crime rates are per se a fact of widening rifts in the social fabric. A low commitment to redistribution and social and health expenditures may thus lead to higher costs for internal security in a society. In both NMS and Non-NMS regions non-activity rates of youngster and early leave from education are strongly correlated with income inequality. We do not expect these regressions to be completely devoid of endogeneity however, the highlighted relationships show that the danger of transmission of difficult material living conditions to the young generation is higher in more unequal societies.

6. Bibliography

Alesina A., Baqir R. and W. Easterly, 1999, Public Goods and Ethnic Divisions, Quarterly Journal of Economics, Vol. 114(4), p. 1243-1284.

Babones S.J., 2008, Income inequality and population health: correlation and causality, Social Sience and Medicine, Vol. 66 (7), p. 1614-1626.

Banerjee A.V. and A.F. Newman, 1993, Occupational Choice and the Process of Development, Journal of Political Economy, Vol. 101(2), p. 274-298.

Becker G. S., 1968, Crime and punishment: An economic approach, Journal of Political Economy, Vol. 76, p. 169-217.

Brush J., 2007, Does income inequality lead to more crime? A comparison of cross-sectional and time-series analyses of United States counties, Economic Letters, Vol. 96, p. 264-268.

Case A.C. and L.F. Katz, 1991, The Company you keep: The effects of family and neighbourhood on disadvantaged youths, NBER Working Paper No. 3705.

Checchi D., 2003, Inequality in Incomes and Access to Education: A Cross-country Analysis (1960-95), Labour, Vol. 17(2), p. 153-201.

Choe J., 2008, Income Inequality and Crime in the United States, Economic Letters, Vol. 101, p. 31-33.

Deaton A., 2003, Health, Inequality, and Economic Development, Journal of Economic Literature, Vol. 41(1), p. 113-158.

D'Hombres B., Weber A. and L. Elia, 2012, Literature review on income inequality and the effects on social outcomes, Joint Research Centre, JRC scientific and policy reports.

De Vogli R., Mistry R., Gnesotto R. and G.A. Cornia, 2005, Has the relation between income inequality and life expectancy disappeared? Evidence from Italy and top industrialized countries, Journal of Epidemiology and Community Health, Vol. 59(2), p. 158 – 162.

Dorling D., Mitchell R. and J. Pearce, 2007, The global impact of income inequality on health by age: an observational study, British Medical Journal, Vol. 335(7625), p. 873.

Elia L., d'Hombres B., Weber A. and A. Saltelli, 2013, Income Inequality and Social Outcomes: Bivariate Correlations at NUTS1 Level, Joint Research Centre, JRC scientific and policy reports.

Entorf H. and H. Spengler, 2000, Socioeconomic and demographic factors of crime in Germany: evidence from panel data of the German states, International Review of Law and Economics, Vol. 20, p. 75-106.

Fajnzylber P., Lederman D. and L. Norman, 2002a, What causes violent crime? European Economic Review, Vol. 46(7), p. 1323-1357.

Fajnzylber P., Lederman D. and L. Norman, 2002b, Inequality and Violent Crime, Journal of Law and Economics, Vol. 45(1), p. 1-40.

Aakvik A., Salvanes K.G. and K. Vaage, 2005, Educational Attainment and Family Background, German Economic Review, Vol. 6 (3), p. 377–394.

Flug K., Spilimbergo A. and E. Wachtenheim ,1998, Investment in Education: Do Economic Volatility and Credit Constraints Matter? Journal of Development Economics, Vol. 55(2), p. 465-481.

Galor O. and J. Zeira , 1993, Income Distribution and Macroeconomics, Review of Economic Studies, Vol. 60(1), p. 35-52.

Gravelle H., 1998, How much of the relation between population mortality and unequal distribution of income is a statistical artefact?, British medical journal, Vol. 316(7128), p. 382-385.

Gravelle H., Wildman J. and M. Sutton, 2002, Income, income inequality and health: what can we learn from aggregate data?, Social Science & Medicine, Vol. 54(4), p. 577-589.

Heerink N., Mulatu A. and E. Bulte, 2001, Income inequality and the environment: aggregation bias in environmental Kuznets curves, Ecological Economics, Vol. 38(3), p. 359-367.

Holzner M. and S. Leitner, 2008, Economic Inequality in Central, East and Southeast Europe, Intervention: European Journal of Economics and Economic Policies, Vol. 5(1), p. 155-88.

Kelly M., 2000, Inequality and crime, The Review of Economics and Statistics, Vol. 82(4), p. 530-539.

Leigh A., Jencks C. and T.M. Smeeding, 2009, Health and Economic Inequality, in: Salverda W., Nolan B. and T. Smeeding, The Oxford Handbook of Economic Inequality, Oxford University Press, Oxford, p. 384-405.

Lynch J., Smith G.D., Kaplan G.A. and J.S. House, 2000, Income inequality and mortality: importance to health of individual income, psychosocial environment, or material conditions, British Medical Journal, Vol. 320(7243), p. 1200–1204.

Lynch J., Smith G.D., Hillemeier M., Shaw M., Raghunathan T. and G. Kaplan, 2001, Income inequality, the psychosocial environment, and health: comparisons of wealthy nations, The Lancet, Vol. 358, p. 194-200.

Lynch J., Smith G.D., Harper S., Hillemeier M., Ross N., Kaplan G.A. and M. Wolfson, 2004, Is income inequality a determinant of population health? Part 1. A systematic review, The Milbank Quarterly, Vol. 82(1), p. 5 –99.

Machin S. and C. Meghir, 2004, Crime and Economic Incentives, Journal of Human Resources, University of Wisconsin Press, Vol. 39(4), p. 958-979.

Macinko J.A., Shi L., Starfield B. and J.T. Wulu Jr., 2003, Income inequality and health: a critical review of the literature, Medical Care Research and Review, Vol. 60(4), p. 407–452.

Messner S.F., Raffalovich L.F. and P. Shrock, 2002, Reassessing the cross-national relationship between income inequality and homicide rates: implications of data quality control in the measurement of income distribution, Journal of Quantitative Criminology, Vol. 18(4), p. 377-395.

Neumayer E., 2005, Inequality and violent crime: evidence from data on robbery and violent theft, Journal of peace research, Vol. 42(1), p. 101-112.

Perotti R., 1993, Growth, Income Distribution and Democracy: What the Data Say, Vol. 1(2), p. 149-187.

Porta M., Borrell C. and J.L. Copete, 2002, Commentary: Theory in the fabric of evidence on the health effects of inequalities in income distribution, International Journal of Epidemiology, Vol. 31(3), p. 543-546.

Preston S.H., 1975, The Changing Relation between Mortality and Level of Economic Development, Population Studies, Vol. 29(2), p. 231-248.

Torre R. and M. Myrskylä, 2014, Income inequality and population health: an analysis of panel data for 21 developed countries, 1975–2006. Population Studies, Vol. 68(1), p. 1-13.

Wilkinson R.G., 1992, Income Distribution And Life Expectancy, British Medical Journal, Vol. 304(6820), p. 165-168.

Wilkinson R.G. and K.E. Pickett, 2009a, Income Inequality and Social Dysfunction, Annual review of sociology, Vol. 35(1), p 493-511.

Wilkinson R.G. and K.E. Pickett, 2009b, The Spirit Level: Why More Equal Societies Almost Always Do Better, Allen Lane.

Appendix

Table A.1. Data sources and time periods of dependent variables and GDP p.c.

Variables	Data source	Time period (average of years)	Missing countries
Population health and social outcome indicators			
Life expectancy	Eurostat Database	2009-2011	
Infant mortality (< 1 year)	Eurostat Database	2009-2011	
Deathrate assault	Eurostat Database	2008-2010	DK
Deathrate drug dependence, toxicomania	Eurostat Database	2008-2010	DK, IE
Deathrate diseases of the circulatory system	Eurostat Database	2008-2010	DK
Deathrate mental deseases	Eurostat Database	2008-2010	DK
Homicide	Eurostat Database	2008-2010, except: DE (09-10), IE (2010)	EL, NL, UK
Robbery	Eurostat Database	2008-2010, except: DE (09-10)	EL, UK
Domestic burglary	Eurostat Database	2008-2010, except: DE (09-10)	EL, UK
Theft of motor vehicles	Eurostat Database	2008-2010, except: DE (09-10)	EL, UK
Youngsters (age 15-24) not in employment, education or training	Eurostat Database	2009-2011	
Early leavers from education (age 18-24)	Eurostat Database	2009-2011	
GDP per capita at PPP	Eurostat Database	2009-2011	

Source: Eurostat, OECD regional well-being dataset, EU SILC, own calculations.

Table A.2. Data sources and time periods of inequality indices: Gini index, (At-risk-of) poverty rate and Income quintile share ratio (S80/S20)

Country	Source and time period (average of years)	Regional level	Number of regions
Austria	OECD: EU SILC, 3 year averages for 2008-2010	NUTS 2	9
Belgium	OECD: EU SILC, 2011 wave (2010 reference income)	NUTS 1	3
Cyprus	calculated from EU SILC, 2011 wave (2010 reference income)	NUTS 2	1
Denmark	OECD: Danish Law Model System, register data, 2010	NUTS 2	5
Finland	OECD: EU SILC, 2012 wave (2011 reference income)	NUTS 2	4
France	OECD: ERFS - Tax and Social Incomes Survey, 2010 reference income	NUTS 2	21
Germany	OECD: SOEP, 2011 wave (2010 reference income)	NUTS 1	16
Greece	OECD: EU SILC, 2011 wave (2010 reference income)	NUTS 1	4
Ireland	calculated from EU SILC, 2011 wave (2010 reference income)	NUTS 1	1
Italy	OECD: UDB IT-SILC, 2012 wave (2011 reference income)	NUTS 2	21
Luxembourg	calculated from EU SILC, 2011 wave (2010 reference income)	NUTS 2	1
Malta	calculated from EU SILC, 2011 wave (2010 reference income)	NUTS 2	1
Netherlands	OECD: Income Panel Survey, 2010	NUTS 1	4
Spain	OECD: EU SILC, 3 year averages for 2008-2010	NUTS 2	16
Sweden	OECD: Income Distribution Survey, 2011 reference income	NUTS 2	8
United Kingdom	calculated from EU SILC, 2011 wave (2010 reference income)	NUTS 2	36
Bulgaria	calculated from EU SILC, 2011 wave (2010 reference income)	NUTS 1	2
Czech Republic	OECD: EU SILC, 2011 wave (2010 reference income)	NUTS 2	8
Estonia	calculated from EU SILC, 2011 wave (2010 reference income)	NUTS 2	1
Hungary	OECD: EU SILC, 2011 wave (2010 reference income)	NUTS 1	3
Poland	OECD: EU SILC, 2011 wave (2010 reference income)	NUTS 1	6
Latvia	calculated from EU SILC, 2011 wave (2010 reference income)	NUTS 2	1
Lithuania	calculated from EU SILC, 2011 wave (2010 reference income)	NUTS 2	1
Romania	calculated from EU SILC, 2011 wave (2010 reference income)	NUTS 1	4
Slovak Republic	OECD: EU SILC, 2011 wave (2010 reference income)	NUTS 2	4
Slovenia	OECD: EU SILC, 2011 wave (2010 reference income)	NUTS 2	2

Source: Eurostat, OECD regional well-being dataset, EU SILC, own calculations.

Table A.3. Regression results for life expectancy (in logs) – fixed and random effects models

VARIABLES	EU regions						EU regions excluding CEE NMS						CEE NMS regions					
ln_gdppcpp	0.017***	0.010***	0.019***	0.021***	0.015***	0.022***	0.014***	0.009*	0.016***	0.011***	0.006	0.012***	0.025***	0.008	0.025***	0.029***	0.010	0.027***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.007)	(0.005)	(0.005)	(0.008)	(0.005)
ln_gini	0.008			0.001			0.010			0.021**			-0.023			-0.061**		
	(0.010)			(0.010)			(0.010)			(0.010)			(0.035)			(0.026)		
ln_poverty		-0.011***			-0.009**			-0.010***			-0.010***			-0.024**			-0.023**	
		(0.003)			(0.003)			(0.004)			(0.004)			(0.009)			(0.010)	
ln_S8020			-0.005			-0.008			-0.004			0.004			-0.028			-0.044***
			(0.005)			(0.005)			(0.006)			(0.005)			(0.021)			(0.015)
Constant	4.188***	4.313***	4.206***	4.156***	4.240***	4.162***	4.217***	4.340***	4.238***	4.215***	4.357***	4.265***	4.164***	4.321***	4.128***	4.251***	4.295***	4.133***
	(0.039)	(0.044)	(0.032)	(0.041)	(0.045)	(0.033)	(0.045)	(0.051)	(0.039)	(0.046)	(0.050)	(0.039)	(0.105)	(0.091)	(0.049)	(0.089)	(0.098)	(0.049)
Observations	181	181	180	181	181	180	149	149	148	149	149	148	32	32	32	32	32	32
Number of countries	26	26	25	26	26	25	16	16	15	16	16	15	10	10	10	10	10	10
R2 within	0.173	0.228	0.175	0.170	0.222	0.174	0.117	0.164	0.113	0.107	0.163	0.0996	0.545	0.657	0.575	0.524	0.656	0.566
R2 between	0.584	0.415	0.580	0.592	0.504	0.575	0.00598	0.0479	0.0493	0.0133	0.0489	0.00726	0.657	0.329	0.661	0.658	0.348	0.656
R2 overall	0.338	0.141	0.301	0.330	0.211	0.286	0.0332	0.0343	0.00392	0.0649	0.0361	0.0293	0.557	0.427	0.615	0.607	0.434	0.626
model	fixed effects			random effects			fixed effects			random effects			fixed effects			random effects		

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Eurostat, OECD regional well-being dataset, EU SILC, own calculations.

Table A.4. Regression results for infant mortality rates (in logs) – fixed and random effects models

VARIABLES	EU regions						EU regions excluding CEE NMS						CEE NMS regions					
ln_gdppcpp	-0.238***	-0.002	-0.232***	-0.291***	-0.092	-0.280***	-0.139**	0.097	-0.134*	-0.129**	0.082	-0.124*	-0.505***	-0.208	-0.490***	-0.606***	-0.316	-0.548***
	(0.059)	(0.067)	(0.058)	(0.054)	(0.064)	(0.054)	(0.068)	(0.074)	(0.068)	(0.064)	(0.070)	(0.064)	(0.115)	(0.171)	(0.102)	(0.111)	(0.203)	(0.104)
ln_gini	0.509***			0.631***			0.464***			0.454***			0.877			1.383***		
	(0.171)			(0.159)			(0.173)			(0.163)			(0.738)			(0.433)		
ln_poverty		0.277***			0.243***			0.273***			0.262***			0.365			0.402*	
		(0.058)			(0.058)			(0.059)			(0.057)			(0.215)			(0.238)	
ln_S8020			0.294***			0.353***			0.246**			0.227**			0.781*			0.932***
			(0.097)			(0.091)			(0.098)			(0.093)			(0.435)			(0.258)
Constant	1.939***	0.511	3.150***	2.126***	1.575**	3.602***	1.059	-0.542	2.197***	0.985	-0.374	2.119***	3.471	2.545	5.100***	2.750	3.502	5.416***
	(0.698)	(0.778)	(0.568)	(0.683)	(0.739)	(0.537)	(0.770)	(0.852)	(0.668)	(0.751)	(0.816)	(0.643)	(2.207)	(2.150)	(1.036)	(1.788)	(2.518)	(1.104)
Observations	183	183	182	183	183	182	151	151	150	151	151	150	32	32	32	32	32	32
Number of countries	26	26	25	26	26	25	16	16	15	16	16	15	10	10	10	10	10	10
R2 within	0.113	0.182	0.115	0.113	0.174	0.115	0.0633	0.150	0.0572	0.0633	0.150	0.0572	0.499	0.531	0.538	0.495	0.530	0.538
R2 between	0.486	0.159	0.473	0.486	0.364	0.473	0.0457	0.00152	0.0123	0.0463	0.000839	0.0123	0.657	0.572	0.649	0.626	0.609	0.642
R2 overall	0.321	0.114	0.287	0.320	0.183	0.288	0.0912	0.0695	0.0521	0.0915	0.0714	0.0521	0.684	0.563	0.727	0.693	0.564	0.728
model	fixed effects			random effects			fixed effects			random effects			fixed effects			random effects		

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Eurostat, OECD regional well-being dataset, EU SILC, own calculations.

Table A.5. Regression results for standardised (age adjusted) death rates: Assault (in logs) – fixed and random effects models

VARIABLES	EU regions						EU regions excluding CEE NMS						CEE NMS regions					
ln_gdppcpp	-0.203	0.287*	-0.128	-0.248**	0.155	-0.161	-0.396**	0.120	-0.302*	-0.301*	0.170	-0.241	0.063	0.470**	0.112	0.014	0.475***	0.096
	(0.124)	(0.155)	(0.119)	(0.114)	(0.144)	(0.112)	(0.167)	(0.207)	(0.163)	(0.160)	(0.191)	(0.158)	(0.115)	(0.188)	(0.099)	(0.116)	(0.184)	(0.100)
ln_gini	1.113***			1.295***			1.039**			0.805*			1.856**			2.546***		
	(0.394)			(0.366)			(0.436)			(0.422)			(0.739)			(0.652)		
ln_poverty		0.502***			0.446***			0.477***			0.499***			0.428*			0.452**	
		(0.122)			(0.119)			(0.139)			(0.134)			(0.236)			(0.230)	
ln_S8020			0.810***			0.882***			0.724***			0.592**			1.362***			1.632***
			(0.227)			(0.214)			(0.254)			(0.247)			(0.421)			(0.376)
Constant	0.651	-1.934	2.423*	0.768	-0.148	2.934**	2.776	-0.317	4.215**	2.776	-0.749	3.981**	-4.116*	-3.002	-0.408	-5.773***	-2.839	-0.489
	(1.618)	(1.799)	(1.245)	(1.571)	(1.672)	(1.190)	(2.048)	(2.361)	(1.718)	(2.028)	(2.203)	(1.688)	(2.207)	(2.364)	(1.003)	(2.128)	(2.342)	(1.056)
Observations	139	139	138	139	139	138	107	107	106	107	107	106	32	32	32	32	32	32
Number of countries	23	23	22	23	23	22	13	13	12	13	13	12	10	10	10	10	10	10
R2 within	0.0742	0.137	0.109	0.0742	0.132	0.108	0.0958	0.150	0.118	0.0957	0.149	0.118	0.329	0.242	0.420	0.322	0.241	0.418
R2 between	0.318	0.0116	0.308	0.317	0.00173	0.307	0.0576	0.0826	0.0963	0.0575	0.0960	0.0966	0.582	0.0158	0.546	0.567	0.0223	0.540
R2 overall	0.0754	0.00315	0.0450	0.0772	0.0167	0.0481	0.000803	0.0779	1.77e-05	0.000764	0.0786	0.0000	0.455	0.0347	0.457	0.452	0.0412	0.454
model	fixed effects			random effects			fixed effects			random effects			fixed effects			random effects		

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Eurostat, OECD regional well-being dataset, EU SILC, own calculations.

Table A.6. Regression results for standardised (age adjusted) death rates: Drug dependence, toxicomania (in logs) – fixed and random effects models

VARIABLES	EU regions						EU regions excluding CEE NMS						CEE NMS regions					
ln_gdppcpp	-0.203	0.287*	-0.128	-0.248**	0.155	-0.161	-0.396**	0.120	-0.302*	-0.301*	0.170	-0.241	0.063	0.470**	0.112	0.014	0.475***	0.096
	(0.124)	(0.155)	(0.119)	(0.114)	(0.144)	(0.112)	(0.167)	(0.207)	(0.163)	(0.160)	(0.191)	(0.158)	(0.115)	(0.188)	(0.099)	(0.116)	(0.184)	(0.100)
ln_gini	1.113***			1.295***			1.039**			0.805*			1.856**			2.546***		
	(0.394)			(0.366)			(0.436)			(0.422)			(0.739)			(0.652)		
ln_poverty		0.502***			0.446***			0.477***			0.499***			0.428*			0.452**	
		(0.122)			(0.119)			(0.139)			(0.134)			(0.236)			(0.230)	
ln_S8020			0.810***			0.882***			0.724***			0.592**			1.362***			1.632***
			(0.227)			(0.214)			(0.254)			(0.247)			(0.421)			(0.376)
Constant	0.651	-1.934	2.423*	0.768	-0.148	2.934**	2.776	-0.317	4.215**	2.776	-0.749	3.981**	-4.116*	-3.002	-0.408	-5.773***	-2.839	-0.489
	(1.618)	(1.799)	(1.245)	(1.571)	(1.672)	(1.190)	(2.048)	(2.361)	(1.718)	(2.028)	(2.203)	(1.688)	(2.207)	(2.364)	(1.003)	(2.128)	(2.342)	(1.056)
Observations	139	139	138	139	139	138	107	107	106	107	107	106	32	32	32	32	32	32
Number of countries	23	23	22	23	23	22	13	13	12	13	13	12	10	10	10	10	10	10
R2 within	0.0742	0.137	0.109	0.0742	0.132	0.108	0.0958	0.150	0.118	0.0957	0.149	0.118	0.329	0.242	0.420	0.322	0.241	0.418
R2 between	0.318	0.0116	0.308	0.317	0.00173	0.307	0.0576	0.0826	0.0963	0.0575	0.0960	0.0966	0.582	0.0158	0.546	0.567	0.0223	0.540
R2 overall	0.0754	0.00315	0.0450	0.0772	0.0167	0.0481	0.000803	0.0779	1.77e-05	0.000764	0.0786	0.0000	0.455	0.0347	0.457	0.452	0.0412	0.454
model	fixed effects			random effects			fixed effects			random effects			fixed effects			random effects		

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Eurostat, OECD regional well-being dataset, EU SILC, own calculations.

Table A.7. Regression results for standardised (age adjusted) death rates: Diseases of the circulatory system (in logs) – fixed and random effects models

VARIABLES	EU regions						EU regions excluding CEE NMS						CEE NMS regions					
ln_gdppcpp	-0.203	0.287*	-0.128	-0.248**	0.155	-0.161	-0.396**	0.120	-0.302*	-0.301*	0.170	-0.241	0.063	0.470**	0.112	0.014	0.475***	0.096
	(0.124)	(0.155)	(0.119)	(0.114)	(0.144)	(0.112)	(0.167)	(0.207)	(0.163)	(0.160)	(0.191)	(0.158)	(0.115)	(0.188)	(0.099)	(0.116)	(0.184)	(0.100)
ln_gini	1.113***			1.295***			1.039**			0.805*			1.856**			2.546***		
	(0.394)			(0.366)			(0.436)			(0.422)			(0.739)			(0.652)		
ln_poverty		0.502***			0.446***			0.477***			0.499***			0.428*			0.452**	
		(0.122)			(0.119)			(0.139)			(0.134)			(0.236)			(0.230)	
ln_S8020			0.810***			0.882***			0.724***			0.592**			1.362***			1.632***
			(0.227)			(0.214)			(0.254)			(0.247)			(0.421)			(0.376)
Constant	0.651	-1.934	2.423*	0.768	-0.148	2.934**	2.776	-0.317	4.215**	2.776	-0.749	3.981**	-4.116*	-3.002	-0.408	-5.773***	-2.839	-0.489
	(1.618)	(1.799)	(1.245)	(1.571)	(1.672)	(1.190)	(2.048)	(2.361)	(1.718)	(2.028)	(2.203)	(1.688)	(2.207)	(2.364)	(1.003)	(2.128)	(2.342)	(1.056)
Observations	139	139	138	139	139	138	107	107	106	107	107	106	32	32	32	32	32	32
Number of countries	23	23	22	23	23	22	13	13	12	13	13	12	10	10	10	10	10	10
R2 within	0.0742	0.137	0.109	0.0742	0.132	0.108	0.0958	0.150	0.118	0.0957	0.149	0.118	0.329	0.242	0.420	0.322	0.241	0.418
R2 between	0.318	0.0116	0.308	0.317	0.00173	0.307	0.0576	0.0826	0.0963	0.0575	0.0960	0.0966	0.582	0.0158	0.546	0.567	0.0223	0.540
R2 overall	0.0754	0.00315	0.0450	0.0772	0.0167	0.0481	0.000803	0.0779	1.77e-05	0.000764	0.0786	0.0000	0.455	0.0347	0.457	0.452	0.0412	0.454
model	fixed effects			random effects			fixed effects			random effects			fixed effects			random effects		

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Eurostat, OECD regional well-being dataset, EU SILC, own calculations.

Table A.8. Regression results for standardised (age adjusted) death rates: Mental diseases (in logs) – fixed and random effects models

VARIABLES	EU regions						EU regions excluding CEE NMS						CEE NMS regions					
ln_gdppcpp	-0.203	0.287*	-0.128	-0.248**	0.155	-0.161	-0.396**	0.120	-0.302*	-0.301*	0.170	-0.241	0.063	0.470**	0.112	0.014	0.475***	0.096
	(0.124)	(0.155)	(0.119)	(0.114)	(0.144)	(0.112)	(0.167)	(0.207)	(0.163)	(0.160)	(0.191)	(0.158)	(0.115)	(0.188)	(0.099)	(0.116)	(0.184)	(0.100)
ln_gini	1.113***			1.295***			1.039**			0.805*			1.856**			2.546***		
	(0.394)			(0.366)			(0.436)			(0.422)			(0.739)			(0.652)		
ln_poverty		0.502***			0.446***			0.477***			0.499***			0.428*			0.452**	
		(0.122)			(0.119)			(0.139)			(0.134)			(0.236)			(0.230)	
ln_S8020			0.810***			0.882***			0.724***			0.592**			1.362***			1.632***
			(0.227)			(0.214)			(0.254)			(0.247)			(0.421)			(0.376)
Constant	0.651	-1.934	2.423*	0.768	-0.148	2.934**	2.776	-0.317	4.215**	2.776	-0.749	3.981**	-4.116*	-3.002	-0.408	-5.773***	-2.839	-0.489
	(1.618)	(1.799)	(1.245)	(1.571)	(1.672)	(1.190)	(2.048)	(2.361)	(1.718)	(2.028)	(2.203)	(1.688)	(2.207)	(2.364)	(1.003)	(2.128)	(2.342)	(1.056)
Observations	139	139	138	139	139	138	107	107	106	107	107	106	32	32	32	32	32	32
Number of countries	23	23	22	23	23	22	13	13	12	13	13	12	10	10	10	10	10	10
R2 within	0.0742	0.137	0.109	0.0742	0.132	0.108	0.0958	0.150	0.118	0.0957	0.149	0.118	0.329	0.242	0.420	0.322	0.241	0.418
R2 between	0.318	0.0116	0.308	0.317	0.00173	0.307	0.0576	0.0826	0.0963	0.0575	0.0960	0.0966	0.582	0.0158	0.546	0.567	0.0223	0.540
R2 overall	0.0754	0.00315	0.0450	0.0772	0.0167	0.0481	0.000803	0.0779	1.77e-05	0.000764	0.0786	0.0000	0.455	0.0347	0.457	0.452	0.0412	0.454
model	fixed effects			random effects			fixed effects			random effects			fixed effects			random effects		

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Eurostat, OECD regional well-being dataset, EU SILC, own calculations.

Table A.9. Regression results for homicide rates (in logs) – fixed and random effects models

VARIABLES	EU regions						EU regions excluding CEE NMS						CEE NMS regions					
ln_gdppcpp	-0.203	0.287*	-0.128	-0.248**	0.155	-0.161	-0.396**	0.120	-0.302*	-0.301*	0.170	-0.241	0.063	0.470**	0.112	0.014	0.475***	0.096
	(0.124)	(0.155)	(0.119)	(0.114)	(0.144)	(0.112)	(0.167)	(0.207)	(0.163)	(0.160)	(0.191)	(0.158)	(0.115)	(0.188)	(0.099)	(0.116)	(0.184)	(0.100)
ln_gini	1.113***			1.295***			1.039**			0.805*			1.856**			2.546***		
	(0.394)			(0.366)			(0.436)			(0.422)			(0.739)			(0.652)		
ln_poverty		0.502***			0.446***			0.477***			0.499***			0.428*			0.452**	
		(0.122)			(0.119)			(0.139)			(0.134)			(0.236)			(0.230)	
ln_S8020			0.810***			0.882***			0.724***			0.592**			1.362***			1.632***
			(0.227)			(0.214)			(0.254)			(0.247)			(0.421)			(0.376)
Constant	0.651	-1.934	2.423*	0.768	-0.148	2.934**	2.776	-0.317	4.215**	2.776	-0.749	3.981**	-4.116*	-3.002	-0.408	-5.773***	-2.839	-0.489
	(1.618)	(1.799)	(1.245)	(1.571)	(1.672)	(1.190)	(2.048)	(2.361)	(1.718)	(2.028)	(2.203)	(1.688)	(2.207)	(2.364)	(1.003)	(2.128)	(2.342)	(1.056)
Observations	139	139	138	139	139	138	107	107	106	107	107	106	32	32	32	32	32	32
Number of countries	23	23	22	23	23	22	13	13	12	13	13	12	10	10	10	10	10	10
R2 within	0.0742	0.137	0.109	0.0742	0.132	0.108	0.0958	0.150	0.118	0.0957	0.149	0.118	0.329	0.242	0.420	0.322	0.241	0.418
R2 between	0.318	0.0116	0.308	0.317	0.00173	0.307	0.0576	0.0826	0.0963	0.0575	0.0960	0.0966	0.582	0.0158	0.546	0.567	0.0223	0.540
R2 overall	0.0754	0.00315	0.0450	0.0772	0.0167	0.0481	0.000803	0.0779	1.77e-05	0.000764	0.0786	0.0000	0.455	0.0347	0.457	0.452	0.0412	0.454
model	fixed effects			random effects			fixed effects			random effects			fixed effects			random effects		

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Eurostat, OECD regional well-being dataset, EU SILC, own calculations.

Table A.10. Regression results for robbery rates (in logs) – fixed and random effects models

VARIABLES	EU regions						EU regions excluding CEE NMS						CEE NMS regions					
ln_gdppcpp	0.696***	1.934***	0.931***	0.811***	1.632***	0.997***	0.767***	2.261***	1.088***	0.781***	2.085***	1.065***	0.697***	1.056***	0.736***	0.693***	1.080***	0.752***
	(0.205)	(0.259)	(0.198)	(0.180)	(0.221)	(0.179)	(0.280)	(0.344)	(0.273)	(0.269)	(0.325)	(0.262)	(0.182)	(0.285)	(0.163)	(0.167)	(0.272)	(0.154)
ln_gini	3.528***			3.068***			3.690***			3.396***			1.643			2.077***		
	(0.649)			(0.578)			(0.731)			(0.709)			(1.167)			(0.793)		
ln_poverty		1.216***			1.081***			1.344***			1.282***			0.377			0.438	
		(0.204)			(0.191)			(0.231)			(0.225)			(0.358)			(0.335)	
ln_S8020			2.172***			1.988***			2.278***			2.219***			1.278*			1.275**
			(0.377)			(0.340)			(0.424)			(0.409)			(0.693)			(0.504)
Constant	-12.47***	-16.46***	-6.256***	-12.18***	-13.11***	-6.71***	-13.66***	-20.23***	-7.96***	-12.89***	-18.48***	-7.67***	-6.425*	-5.417	-3.209*	-7.82***	-5.730*	-3.342**
	(2.665)	(3.004)	(2.072)	(2.546)	(2.568)	(1.912)	(3.429)	(3.925)	(2.879)	(3.382)	(3.740)	(2.791)	(3.487)	(3.585)	(1.653)	(2.857)	(3.425)	(1.624)
Observations	143	143	142	143	143	142	111	111	110	111	111	110	32	32	32	32	32	32
Number of countries	24	24	23	24	24	23	14	14	13	14	14	13	10	10	10	10	10	10
R2 within	0.303	0.330	0.320	0.300	0.329	0.318	0.289	0.336	0.308	0.289	0.335	0.308	0.572	0.555	0.598	0.570	0.554	0.598
R2 between	0.0613	0.141	0.147	0.0910	0.143	0.166	0.0145	0.0542	0.0528	0.0112	0.0551	0.0528	0.503	0.00954	0.255	0.431	0.0135	0.259
R2 overall	0.190	0.236	0.206	0.210	0.237	0.219	0.132	0.139	0.157	0.132	0.141	0.157	0.292	0.227	0.170	0.267	0.231	0.174
model	fixed effects			random effects			fixed effects			random effects			fixed effects			random effects		

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Eurostat, OECD regional well-being dataset, EU SILC, own calculations.

Table A.11. Regression results for rates of domestic burglary (in logs) – fixed and random effects models

VARIABLES	EU regions						EU regions excluding CEE NMS						CEE NMS regions					
ln_gdppcpp	0.665***	1.131***	0.774***	0.749***	1.137***	0.824***	0.664***	1.261***	0.805***	0.695***	1.228***	0.812***	0.690***	0.535**	0.723***	0.635***	0.600**	0.698***
	(0.161)	(0.210)	(0.158)	(0.145)	(0.183)	(0.146)	(0.220)	(0.280)	(0.219)	(0.210)	(0.263)	(0.210)	(0.148)	(0.226)	(0.134)	(0.136)	(0.239)	(0.127)
ln_gini	1.642***			1.606***			1.677***			1.586***			1.227			1.828**		
	(0.507)			(0.464)			(0.573)			(0.553)			(0.945)			(0.751)		
ln_poverty		0.433***			0.472***			0.523***			0.507***			-0.371			-0.213	
		(0.165)			(0.156)			(0.188)			(0.182)			(0.284)			(0.296)	
ln_S8020			0.888***			0.869***			0.895***			0.834**			0.893			1.122**
			(0.301)			(0.278)			(0.340)			(0.329)			(0.569)			(0.460)
Constant	-4.730**	-5.117**	-1.659	-5.345***	-5.156**	-2.036	-4.750*	-6.656**	-1.905	-4.559*	-6.141**	-1.708	-3.924	2.665	-1.467	-5.224**	1.894	-1.394
	(2.084)	(2.430)	(1.651)	(2.021)	(2.124)	(1.554)	(2.687)	(3.195)	(2.306)	(2.634)	(3.036)	(2.245)	(2.823)	(2.843)	(1.357)	(2.472)	(3.018)	(1.338)
Observations	143	143	142	143	143	142	111	111	110	111	111	110	32	32	32	32	32	32
Number of countries	24	24	23	24	24	23	14	14	13	14	14	13	10	10	10	10	10	10
R2 within	0.229	0.206	0.218	0.228	0.206	0.218	0.183	0.176	0.170	0.183	0.176	0.170	0.644	0.644	0.656	0.637	0.640	0.654
R2 between	0.335	0.272	0.248	0.337	0.277	0.249	0.0827	0.104	0.0528	0.0876	0.104	0.0562	0.151	0.206	0.111	0.270	0.177	0.157
R2 overall	0.207	0.222	0.192	0.215	0.222	0.197	0.0689	0.0776	0.0583	0.0722	0.0777	0.0612	0.184	0.0133	0.172	0.237	0.0201	0.189
model	fixed effects			random effects			fixed effects			random effects			fixed effects			random effects		

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Eurostat, OECD regional well-being dataset, EU SILC, own calculations.

Table A.12. Regression results for rates of motor vehicle theft (in logs) – fixed and random effects models

VARIABLES	EU regions						EU regions excluding CEE NMS						CEE NMS regions					
ln_gdppcpp	0.264	1.301***	0.452	0.564	1.377***	0.670**	-0.460**	0.706**	-0.202	-0.443**	0.551**	-0.236*	1.554***	1.345***	1.563***	1.709***	1.225***	1.694***
	(0.377)	(0.306)	(0.296)	(0.383)	(0.276)	(0.306)	(0.197)	(0.257)	(0.138)	(0.196)	(0.248)	(0.140)	(0.304)	(0.275)	(0.263)	(0.153)	(0.332)	(0.158)
ln_gini	2.826**			1.948*			2.957***			2.650***			0.930			-0.266		
	(1.061)			(1.180)			(0.643)			(0.656)			(1.607)			(1.005)		
ln_poverty		1.029***			1.015***			1.043***			0.978***			-0.422			-0.717**	
		(0.306)			(0.320)			(0.186)			(0.176)			(0.324)			(0.344)	
ln_S8020			1.899***			1.411**			1.820***			1.621***			0.944			-0.036
			(0.550)			(0.660)			(0.420)			(0.410)			(0.876)			(0.469)
Constant	-4.976	-8.784**	-0.234	-5.327*	-9.742***	-1.985	2.197	-2.645	6.772***	3.140	-0.928	7.484***	-12.0***	-5.754	-10.3***	-9.59***	-3.832	-10.2***
	(3.507)	(3.553)	(2.990)	(3.089)	(3.285)	(2.834)	(1.941)	(3.000)	(1.647)	(1.952)	(2.908)	(1.601)	(3.280)	(3.282)	(1.501)	(3.633)	(4.047)	(1.688)
Observations	143	143	142	143	143	142	111	111	110	111	111	110	32	32	32	32	32	32
Number of countries	24	24	23	24	24	23	14	14	13	14	14	13	10	10	10	10	10	10
R2 within	0.174	0.214	0.216	0.150	0.213	0.196	0.206	0.238	0.224	0.206	0.236	0.224	0.634	0.639	0.641	0.628	0.636	0.629
R2 between	0.0413	0.245	0.0625	0.0515	0.255	0.00502	0.00274	0.00981	0.00674	0.00364	0.00373	0.00350	0.512	0.510	0.394	0.433	0.529	0.449
R2 overall	0.00360	0.208	0.00556	0.0670	0.216	0.0471	0.0246	0.0558	0.0135	0.0254	0.0627	0.0148	0.442	0.604	0.334	0.557	0.630	0.544
model		fixed effects			random effects			fixed effects			random effects			fixed effects			random effects	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Eurostat, OECD regional well-being dataset, EU SILC, own calculations.

Table A.13. Regression results for share of youngsters (age 15-24) not in employment, education or training - NEET (in logs) – fixed and random effects models

VARIABLES	EU regions						EU regions excluding CEE NMS						CEE NMS regions					
ln_gdppcpp	-0.54***	-0.15**	-0.54***	-0.55***	-0.19***	-0.55***	-0.49***	-0.073	-0.50***	-0.52***	-0.093	-0.54***	-0.62***	-0.48**	-0.62***	-0.72***	-0.52***	-0.69***
	(0.063)	(0.062)	(0.061)	(0.058)	(0.059)	(0.056)	(0.074)	(0.067)	(0.073)	(0.073)	(0.067)	(0.071)	(0.129)	(0.194)	(0.119)	(0.110)	(0.179)	(0.112)
ln_gini	0.577***			0.717***			0.598***			0.706***			-0.091			0.849***		
	(0.182)			(0.170)			(0.186)			(0.185)			(0.824)			(0.307)		
ln_poverty		0.493***			0.489***			0.516***			0.516***			0.222			0.316	
		(0.054)			(0.053)			(0.053)			(0.053)			(0.243)			(0.195)	
ln_S8020			0.411***			0.489***			0.422***			0.493***			-0.079			0.486**
			(0.101)			(0.094)			(0.103)			(0.102)			(0.505)			(0.191)
Constant	5.940***	2.578***	7.311***	5.484***	2.917***	7.206***	5.434***	1.738**	6.942***	5.288***	1.818**	7.119***	8.728***	6.497**	8.558***	6.552***	6.675***	8.397***
	(0.747)	(0.720)	(0.598)	(0.736)	(0.686)	(0.559)	(0.836)	(0.771)	(0.709)	(0.846)	(0.781)	(0.709)	(2.464)	(2.432)	(1.205)	(1.686)	(2.177)	(1.210)
Observations	182	182	181	182	182	181	150	150	149	150	150	149	32	32	32	32	32	32
Number of countries	26	26	25	26	26	25	16	16	15	16	16	15	10	10	10	10	10	10
R2 within	0.324	0.536	0.349	0.321	0.535	0.347	0.259	0.536	0.291	0.258	0.536	0.290	0.598	0.614	0.598	0.575	0.613	0.576
R2 between	0.460	0.532	0.584	0.479	0.540	0.596	0.448	0.502	0.594	0.461	0.510	0.603	0.760	0.782	0.750	0.842	0.772	0.823
R2 overall	0.352	0.422	0.412	0.373	0.426	0.430	0.437	0.366	0.511	0.446	0.371	0.518	0.598	0.661	0.587	0.703	0.663	0.698
model	fixed effects			random effects			fixed effects			random effects			fixed effects			random effects		

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Eurostat, OECD regional well-being dataset, EU SILC, own calculations.

Table A.14. Regression results for share of early leavers from education (age 18-24) (in logs) – fixed and random effects models

VARIABLES	EU regions						EU regions excluding CEE NMS						CEE NMS regions					
ln_gdppcpp	-0.31***	-0.113	-0.30***	-0.29***	-0.078	-0.28***	-0.25***	-0.053	-0.24***	-0.31***	-0.111	-0.30***	-0.420*	-0.141	-0.454**	-0.655***	-0.296	-0.594***
	(0.076)	(0.089)	(0.076)	(0.072)	(0.084)	(0.072)	(0.083)	(0.094)	(0.083)	(0.080)	(0.092)	(0.081)	(0.213)	(0.317)	(0.197)	(0.188)	(0.319)	(0.186)
ln_gini	0.489**			0.564***			0.512**			0.658***			-0.290			1.411**		
	(0.218)			(0.210)			(0.207)			(0.203)			(1.364)			(0.643)		
ln_poverty		0.229***			0.273***			0.214***			0.208***			0.462			0.531	
		(0.077)			(0.076)			(0.074)			(0.074)			(0.397)			(0.361)	
ln_S8020			0.226*			0.276**			0.209*			0.318***			0.243			1.015***
			(0.124)			(0.120)			(0.119)			(0.116)			(0.835)			(0.376)
Constant	4.031***	2.992***	5.192***	3.377***	2.373**	4.750***	3.491***	2.547**	4.726***	3.512***	3.063***	5.155***	6.974	2.071	5.981***	3.615	3.446	6.225***
	(0.896)	(1.030)	(0.736)	(0.895)	(0.976)	(0.714)	(0.930)	(1.075)	(0.815)	(0.934)	(1.070)	(0.811)	(4.077)	(3.980)	(1.992)	(2.958)	(3.915)	(2.005)
Observations	182	182	181	182	182	181	150	150	149	150	150	149	32	32	32	32	32	32
Number of countries	26	26	25	26	26	25	16	16	15	16	16	15	10	10	10	10	10	10
R2 within	0.108	0.129	0.0978	0.106	0.128	0.0960	0.0863	0.102	0.0658	0.0862	0.0999	0.0651	0.212	0.260	0.213	0.174	0.258	0.196
R2 between	0.0598	0.168	0.0539	0.0734	0.226	0.0683	0.540	0.431	0.572	0.541	0.492	0.585	0.526	0.542	0.706	0.650	0.582	0.665
R2 overall	0.0690	0.121	0.0669	0.0897	0.146	0.0931	0.335	0.149	0.348	0.339	0.160	0.369	0.253	0.389	0.475	0.479	0.399	0.526
model	fixed effects			random effects			fixed effects			random effects			fixed effects			random effects		

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Source: Eurostat, OECD regional well-being dataset, EU SILC, own calculations.

Table A.15. Results of Hausman tests for the choice between fixed effects and random effects model specification for individual regressions

VARIABLES (in logs)	EU regions			EU regions excluding CEE NMS			CEE NMS regions		
	Gini	Poverty	S80/S20	Gini	Poverty	S80/S20	Gini	Poverty	S80/S20
	(in logarithms)			(in logarithms)			(in logarithms)		
Life expectancy	0.000	0.000	0.001	0.040	0.115	0.009	0.256	0.047	0.438
Infant mortality (< 1 year)	0.083	0.002	0.107	0.854	0.619	0.804	0.023	0.002	0.046
Deathrate assault	0.276	0.025	0.371	0.227	0.798	0.227	0.096	0.590	0.138
Deathrate drug dependence, toxicomania	0.057	0.029	0.101	0.939	0.738	0.905	0.374	0.432	0.331
Deathrate diseases of the circulatory system	0.000	0.000	0.001	0.374	0.643	0.228	0.079	0.029	0.092
Deathrate mental deseases	0.135	0.022	0.193	0.354	0.654	0.418	0.567	0.739	0.521
Homicide	0.452	0.067	0.566	0.143	0.818	0.117	0.062	0.519	0.139
Robbery	0.304	0.086	0.508	0.244	0.265	0.765	0.194	0.757	0.609
Domestic burglary	0.339	0.352	0.684	0.794	0.929	0.741	0.556	0.062	0.720
Theft of motor vehicles	0.001	0.097	0.001	0.126	0.300	0.076	0.300	0.335	0.290
Youngsters (age 15-24) not in employment, education or training	0.032	0.018	0.061	0.033	0.026	0.034	0.509	0.143	0.479
Early leavers from education (age 18-24)	0.015	0.020	0.010	0.035	0.014	0.009	0.095	0.027	0.203

Random effects model preferable

Fixed effects model preferable