SUPPLEMENT

Supplement 1: Weighting of the domains and use of indicators from the 'German Index of Multiple Deprivation' (GIMD).

Domains	Domain weight (%)	Indicators (reference)
Income deprivation	25	- Total earnings (Number of Taxpayers)
Employment deprivation	25	- Total number of unemployed (Population, 15 to 65 years)
Educational deprivation	15	- Persons without vocational training (Employees subject to social security contributions at the place of residence)
Municipal revenue deprivation	15	- Tax revenue of municipalities (Total population)
Social capital deprivation	10	- Migration balance* (Total population) - Electoral participation in % (Federal parliament)
Environment deprivation	5	- Commercial, industrial and traffic areas ** (Total area)
Security deprivation	5	- Number of road accidents(Total population) - Number of crimes(Total population)

* People moving into a municipality or a district minus people leaving a municipality or a district.

** Indicator for soil sealing.

Supplement 2: Calculation of the standardized mortality rates (SMR):

1. SMR ‘total mortality’ per district = total deaths per district / expected total deaths per district

2. Expected total deaths per district = total population size per district * total mortality rate per 100,000 per district / 100,000

3. Total mortality rate per 100,000 per district = total deaths per district/total population size per district*100,000

4. SMR ‘premature mortaltiy’ per district = premature (before 65 years) deaths per district / expected premature (before 65 years) deaths per district

5. Expected premature deaths per district = premature population size per district * premature mortality rate per 100,000 per district / 100,000

6. Premature mortality rate per 100,000 per district = premature deaths per district/premature population size per district*100,000

Supplement 3: Working steps of the greedy weighting algorithm.

- The vector P containing the greedy solution of the non-normalized weighted sum in each step is initialized with zero elements.
- All column weights and the total number of weights are also initialized to zero. In each iteration, first, the total number of weights is incremented.
- Then, all sums of P with a column of X are normalized by the total number of weights and evaluated separately on the evaluation metric (correlation).
- The column corresponding to the highest value is assigned one weight factor and added to P. This procedure is repeated 100 times.
- The algorithm returns a vector of length N, with the number of columns of X, containing weights for each column, summing to 1.

Supplement 4: Results of the linear regression: Outcome: deprivation proxy, Covariables: domains of the GIMD10.

Deprivation of living space = Income + Employment + Education + Municipal income + Social capital + Environment + Security

Coefficients (Robust Standard Errors in Parenthesis)	
Income	0.014
	(0.051)
Employment	0.067*
	(0.029)
Education	-0.048*
	(0.023)
Municipal income	-0.094***
	(0.025)
Social capital	0.035**
	(0.014)
Environment	0.045***
	(0.011)
Security	-0.006
	(0.008)

Model

R-squared = 0.34
adj. R-squared = 0.33
F = 30.01
$p < 0.001$
Log-likelihood = -1050.76
Deviance = 3959.74
AIC = 2117.52
BIC = 2149.69
N = 412

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; n.s. not significant

Source: R-Output, regression results conducted with R-package 'stargazer'

(Hlavac M. stargazer: Well-Formatted Regression and Summary Statistics Tables. R package version 5.2. 2015.)

→ All domains have a significant effect on the proxy, except security and Income

→ Overall model explains the variance of living space deprivation significantly, $R^2 = 0.34$

Test of the assumptions of the linear regression model:

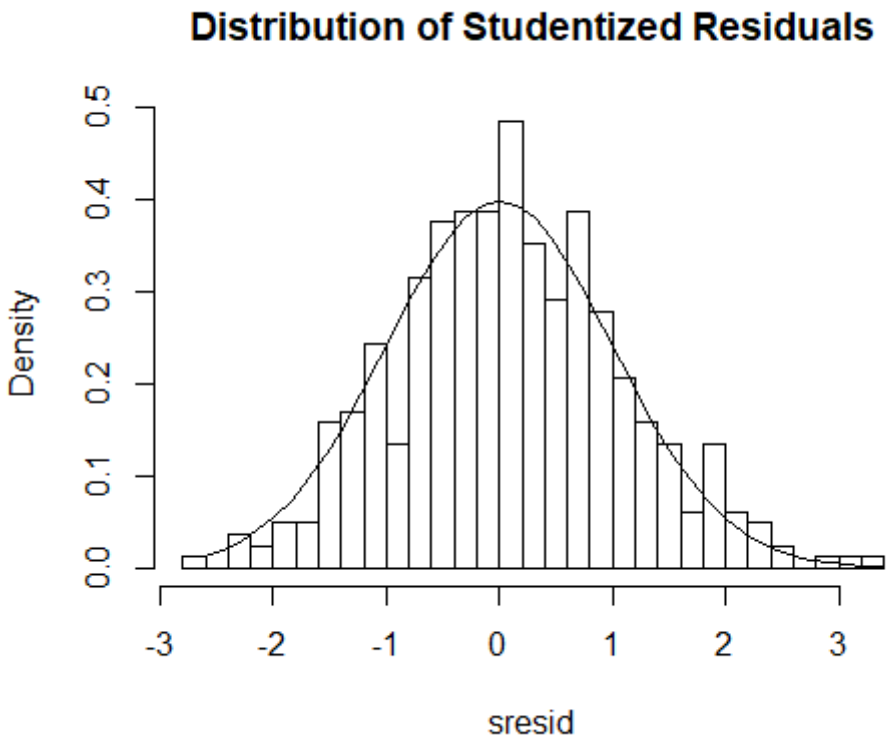
1. Normality of the residuals

Shapiro-Wilk normality test of the residuals of the model:

W = 0.99668, p-value = 0.5588

→ Distribution of residuals of the model differ not significantly from normal distribution

Figure 1: Histogram of the distribution of the residuals:



2.

→ Distribution of residuals of the model differ not significantly from normal distribution

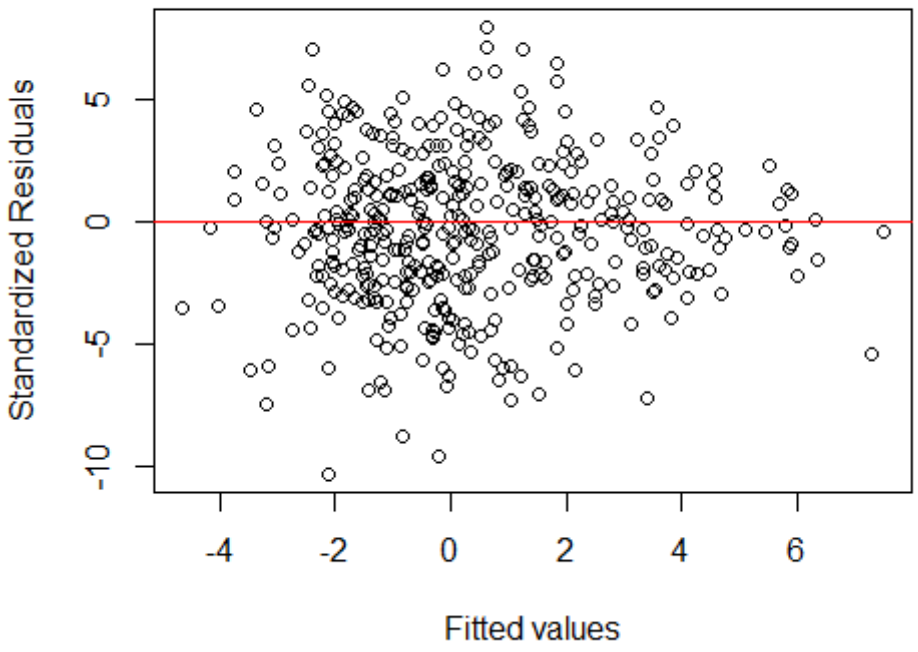
2.Homoscedasticy

Non-constant Variance Score Test

Chi-square = 5.910324 Df = 1 p = 0.0150524

→ assumption of constant variance violated

Figure 2: Plot of the standardized residuals vs. fitted values



→ Visually no violation of the homoscedasticy assumption

→ Due to the results of the Non-constant Variance Score Test, we conducted robust standard errors for the model

3. Multicollinearity

Variance inflation factors of the independent variables:

Income: 10.01, Employment: 8.79, Education: 2.49, Municipal Income: 5.75, Social Capital: 7.86, Environment: 3.96, Security: 2.88

→ Only Income has a value above the critical value of 10

→ Some minor multicollinearity regarding Income

4. Autocorrelation

Durbin Watson Test

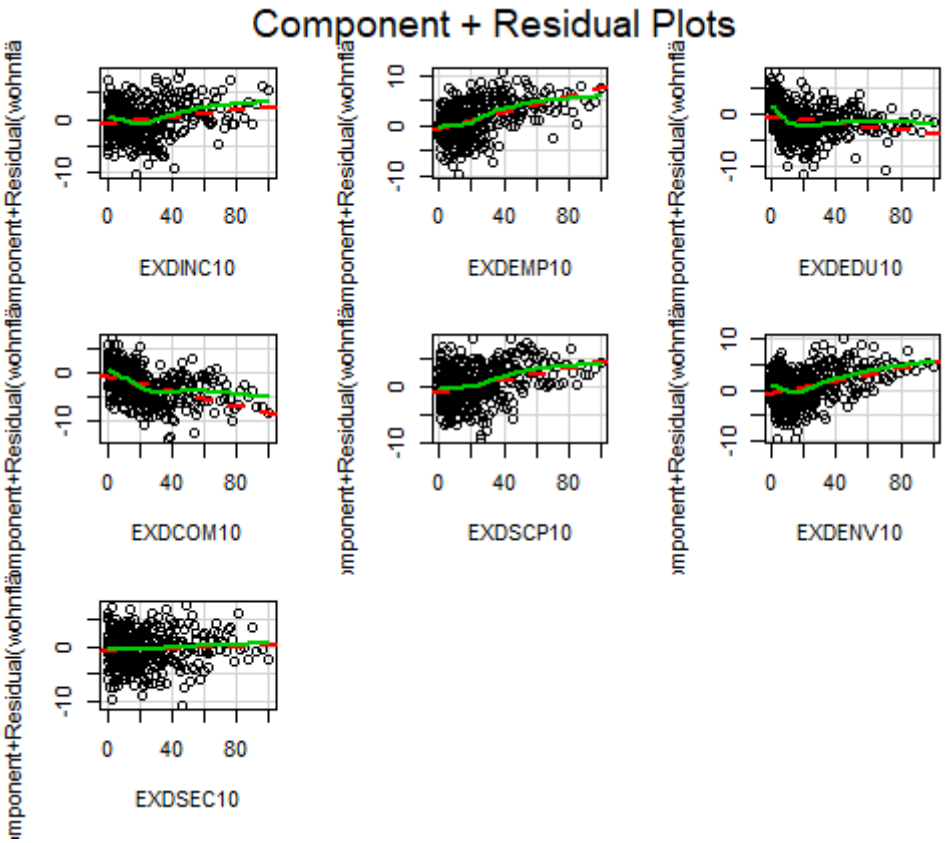
Autocorrelation	D-W Statistic	p-value
0.5036934	0.9876657	<0.001

Alternative hypothesis: rho != 0

→ Autocorrelation is present

5. Nonlinearity

Figure 3: Partial residuals plot



→Linearity assumption violated for domains of education(EXDEDU10), municipal income (EXDCOM10) and Environment (EXDENV10)

→ But domain weights should be obtained, so we use the untransformed variables

Supplement 5: Results of an exploratory factor analysis of the deprivation domains with the extraction of one factor, method = principal axis factor analysis.

Standardized loadings	Factor 1	Communality (u^2)	Specific variance ($1– u^2$)
Income	0.92	0.85	0.15
Employment	0.76	0.58	0.42
Education	–0.36	0.13	0.87
Municipal income	0.87	0.75	0.25
Social capital	0.80	0.64	0.36
Environment	-0.23	0.06	0.94
Security	-0.05	0.01	0.99

Model	Factor 1
Variance, explained by the factor (SS loadings)	3.01
Proportion of total variance	0.43

Model fit measures	
Root mean square of the residuals (RMSR)	0.17
Likelihood chi square	584.65 ($p < 0.001$)
Tucker–Lewis index of fact oring reliability	0.50
RMSEA index (confidence interval)	0.32 ([0.30; 0.34])
BIC	500.35

Source: Tables output from R and own presentation

Supplement 6: Test[1] of the differences in the Spearman correlation coefficients for the relationship of the GIMD[2] versions and both total and premature mortality

Total mortality	**Original weighting**	**Equal weighting**	**Linear regression**	**Maximization algorithm**	**Factor analysis**
Original weighting	0	0.043** [0.015, 0.074]	0.014$^{n.s.}$ [-0.006, 0.034]	–0.037*** [-0.060, -0.016]	–0.020** [-0.038, -0.005]
Equal weighting	–0.043** [-0.074, -0.015]	0	–0.029* [-0.059, -0.001]	–0.080*** [-0.122, -0.041]	–0.063*** [-0.105, -0.025]
Linear regression	–0.014$^{n.s.}$ [-0.034, 0.006]	0.029* [0.001, 0.059]	0	–0.051*** [-0.080, -0.024]	-0.034** [-0.059, -0.011]
Maximization algorithm	0.037*** [0.016, 0.060]	0.080*** [0.041, 0.122]	0.051*** [0.024, 0.080]	0	0.016** [0.003, 0.031]
Factor analysis	0.020** [0.005, 0.038]	0.063*** [0.025 , 0.105]	0.034** [0.011, 0.059]	–0.016 ** [-0.031, -0.003]	0
Premature mortality					
Original weighting	0	0.068*** [0.044, 0.097]	0.028*** [0.012, 0.049]	–0.065*** [-0.093, - 0.043]	–0.005$^{n.s.}$ [-0.021, 0.019]
Equal weighting	–0.068*** [-0.097, -0.044]	0	–0.040***	–0.133*** [-0.174, -0.098]	–0.073*** [0.110, -0.040]
Linear regression	–0.028*** [-0.049, -0.012]	–0.040***	0	–0.094 *** [-0.128, -0.066]	–0.034*** [-0.014, -0.057]
Maximization algorithm	0.065*** [0.043, 0.093]	0.133*** [0.098, 0.174]	0.094*** [0.066, 0.128]	0	0.060*** [0.037, 0.088]
Factor analysis	0.005$^{n.s.}$ [-0.019, 0.021]	0.073*** [0.040, 0.110]	0.034*** [0.014, 0.057]	–0. 060*** [-0.088, -0.037]	0

*** p < 0.001; ** p < 0.01; * p < 0.05; n.s. not significant, 95% confidence intervals in square brackets

[1] Test of the significance of the differences with Williams's t-test for paired correlations

[2] **GIMD:** German Index of Multiple Deprivation

Original weighting: Spearman correlation between GIMD (weighting according to Maier et al. [8]) and both total and premature mortality;

Equal weighting: Spearman correlation between GIMD (domains equally weighted) and both total and premature mortality;

Linear regression: Spearman correlation between GIMD (weighting of the domains with regression coefficients with a deprivation proxy as dependent and domains as independent variables) and both total and

premature mortality;

Maximization algorithm: Spearman correlation between GIMD (weighting of the domains for the maximum Spearman correlation between overall index and mortality) and both total and premature mortality;

Factor analysis: Spearman correlation between GIMD (weighting of the domains with loadings from principal axis factoring) and both total and premature mortality.

Supplement 7: Corrected Test[1] of the differences in the Spearman correlation coefficients for the relationship of the GIMD[2] versions and both total and premature mortality.

Total mortality (all age groups)	**Original weighting**	**Equal weighting**	**Linear regression**	**Maximization algorithm**	**Factor analysis**
Original weighting	0	0.043*	0.014$^{n.s.}$	–0.037**	–0.020*
Equal weighting	–0.043*	0	–0.029*	–0.080***	–0.064**
Linear regression	–0.014$^{n.s.}$	0.029*	0	–0.051 ***	–0.034*
Maximization algorithm	0.037**	0.080**	0.051**	0	0.016*
Factor analysis	0.020*	0.064**	0.034*	–0.016 *	0
Premature mortality (< 65 years)					
Original weighting	0	0.068***	0.028***	–0.065***	–0.005$^{n.s.}$
Equal weighting	–0.068***	0	–0.040**	–0.133***	–0.073***
Linear regression	–0.028***	–0.040**	0	–0.094 ***	–0.034***
Maximization algorithm	0.065***	0.133***	0.094 ***	0	0.060***
Factor analysis	0.005$^{n.s.}$	0.073***	0.034***	–0. 060***	0

*** p < 0.001; ** p < 0.01; * p < 0.05; n.s. not significant

[1] Test of the significance of the differences with Williams's t-test for paired correlations

[2] **GIMD:** German Index of Multiple Deprivation

Original weighting: Spearman correlation between GIMD (weighting according to Maier et al. [8]) and both total and premature mortality;

Equal weighting: Spearman correlation between GIMD (domains equally weighted) and both total and premature mortality;

Linear regression: Spearman correlation between GIMD (weighting of the domains with regression

coefficients with a deprivation proxy as dependent and domains as independent variables) and both total and premature mortality;

Maximization algorithm: Spearman correlation between GIMD (weighting of the domains for the maximum Spearman correlation between overall index and mortality) and both total and premature mortality;

Factor analysis: Spearman correlation between GIMD (weighting of the domains with loadings from principal axis factoring) and both total and premature mortality.