Das Sozio-oekonomische Panel

620 2013

SOEPpapers
on Multidisciplinary Panel Data Research

SOEP — The German Socio-Economic Panel Study at DIW Berlin

620-2013

# Examining the Structure of Spatial Health Effects in Germany Using Hierarchical Bayes Models

Peter Eibich and Nicolas R. Ziebarth

ISSN: 1864-6689 (online)

German Socio-Economic Panel Study (SOEP)
DIW Berlin
Mohrenstrasse 58
10117 Berlin, Germany

Contact: Uta Rahmann | soeppapers@diw.de

# Examining the Structure of Spatial Health Effects in Germany Using Hierarchical Bayes Models*

**Peter Eibich[1] and Nicolas R. Ziebarth[2]**

December 19, 2013

**Abstract**

This paper uses Hierarchical Bayes Models to model and estimate spatial health effects in Germany. We combine rich individual-level household panel data from the German SOEP with administrative county–level data to estimate spatial county-level health dependencies. As dependent variable we use the generic, continuous, and quasi-objective SF12 health measure. We find strong and highly significant spatial dependencies and clusters. The strong and systematic county-level impact is equivalent to 0.35 standard deviations in health. Even 20 years after German reunification, we detect a clear spatial East-West health pattern that equals an age impact on health of up to 5 life years for a 40-year old.

**Keywords:** spatial health effects, Hierarchical Bayes Models, Germany, SOEP, SF12

**JEL codes: C21, C11, I12, I14, I18**

---

* Peter Eibich gratefully acknowledges generous support under "BASEII" by the Bundesministerium für Bildung und Forschung (BMBF; "Federal Ministry of Education and Research", 16SV5537). We would like to thank Adam Lederer for excellence in editing this paper. Moreover, we would like to thank Arnab Bhattacharjee, Luke Born, Dean Lillard, Taps Maiti, Sophie Meyer, Ulrich Rendtel, Thomas Siedler, Jonathan Skinner, Daniel Sturm and participants of the 3rd Health Econometrics Workshop in Siena, the 2013 meeting of the Spatial Econometrics Association in Washington D.C., the 2013 meeting of the German Association of Health Economists (dggö) in Essen, the 2013 Conference of the European Society for Population Economics (ESPE) in Aarhus, the 2013 First Bayesian Young Statistician's Meeting in Milan, the 2013 Conference of the German Economic Association (VfS) as well as seminar participants at the German Institute for Economic Research (DIW) Berlin and the Berlin Network of Labour Market Researchers (BeNA) and two anonymous referees for their helpful comments and discussions. In particular, we would like to thank Guiseppe Arbia and Jon Rothbaum for outstanding discussions of this paper. We take responsibility for all errors in and shortcomings of the article. The research reported in this paper is not the result of a for-pay consulting relationship. Our employers do not have a financial interest in the topic of the paper that might constitute a conflict of interest.

[1] Corresponding author: DIW Berlin, Mohrenstrasse 58, 10117 Berlin, Germany, and University of Hamburg, e-mail: peibich@diw.de, Phone: +49-(30) 897 89-223, Fax: +49-(30) 897 89-200.

[2] Cornell University, Policy Analysis and Management (PAM), 106 Martha van Rensselaer Hall, Ithaca, NY 14853, DIW Berlin, and IZA Bonn, e-mail: nrz2@cornell.edu, Phone: +1-(607) 255-1180, Fax: +1-(607) 255-4071.

## 1 Introduction

Applied empirical researchers have long recognized that regional and neighborhood effects may play a crucial role in the analysis of a wide array of relevant outcomes. For example, in the economics literature, Burchardi and Hassan (2013) exploit the natural experiment of the German reunification to demonstrate that social ties have a long-lasting impact on individual and regional economic prosperity. Studies in sociology analyze the relevance of relative deprivation theories by looking at a respondents' relative income position in the neighborhood (Durlauf, 2003; Knies et al., 2008).

When it comes to regional or neighborhood impacts on individuals' health or health care utilization, the literature is fragmented, both within and across fields. One can identify at least two subfields in the health economics literature that are related to this paper. First, a relatively large set of studies—in particular in the epidemiological and public health literature—focuses on geographical variation in health, health care, or health care expenditures as well as their determinants (Di Matteo and Di Matteo, 1998, Di Matteo, 2005 Drewnowski et al. 2007, Lahti-Koski et al., 2008, Michimi and Wimberly, 2010, Voigtländer et al., 2010, Sundmacher et al., 2012, Frakes, 2013). The second subfield emerged from the first older subfield on regional health (care) variation. It focusses on the importance of regional factors and geographic spillover effects in the spatial analysis of health and health care utilization and econometrically models spatial interdependencies explicitly (Browning et al., 2003, Lauridsen at al., 2008, Bech and Lauridsen, 2008, 2009, Filippini et al., 2009, Pouliou and Elliott, 2009, Lauridsen et al. 2010a, 2010b, Andersen et al., 2012, Hajizadeh et al., 2012, Barufi et al., 2012, Filippini et al., 2013, González Ortiz and Masiero, 2013).

This paper contributes to both research fields outlined above. It models and estimates spatial health patterns at the county-level in Germany. First of all, in terms of the empirical modeling techniques, we apply Hierarchical Bayes Models and combine three methodological approaches: *(i)* Hierarchical models are employed to account for the correlation within regions and to disentangle the effect of individual and regional predictors on individual outcomes (Bryk and Raudenbush, 1992; Hox, 2002). The paper uses a three-stage hierarchical model that also accounts for the temporal dimension of the data. *(ii)* Methods from spatial econometrics are employed to model the spatial correlation between regions, e.g., MORAN'S I and Conditional Autoregressive (CAR) models (Arbia, 2006; Cressie, 1993). *(iii)* The analysis is carried out using Bayesian methods, which facilitate the modeling and estimation of regional effects (Banerjee et al., 2004; Cressie and Wikle, 2011).

All three methodological approaches are well-known and are extensively applied. The contribution of this paper lies in the application of Hierarchical Bayes Models, which allows us to combine individual-level panel data with spatial dependence on the regional-level. This is not feasible using standard spatial panel data models, since no distance measure is available at the individual-level and data only provides the residential area. Therefore, spatial correlation cannot be measured and modeled directly in the data. Using Hierarchical Bayes Models, we introduce spatial heterogeneity at a higher level in the hierarchy. Spatial dependence is introduced through the prior distributions for these parameters. Overall, our models make use of temporal and spatio-temporal dynamics and decompose the spatial effects into spatial heterogeneity and spatial dependence (Banerjee et al. (2004)). Hence the paper contributes to the second, relatively young, research field on spatial health econometrics by applying existing methods in a new way.

Second, in terms of the data used, this paper uses rich individual-level panel data and combines them with administrative data on the county level. The latter vary annually. To date, spatial models are almost exclusively used with time-series data on a higher aggregated level. In addition to other advantages, our approach allows to consider a rich array of background information that may predict and explain regional variation in health.

Third, in terms of content, this paper contributes to the literature by focusing on Germany and health *outcomes* instead of health care utilization or expenditure patterns, the latter being the focus of most existing studies. Germany is a particularly interesting case to study since it was divided for 40 years into a communist and a capitalist part. Therefore, one may hypothesize that differences in the economic and environmental conditions in East and West Germany have had a differential long-term impact on health that is identifiable even more than 20 years after 1990's reunification.

As outcome variable we use a continuous generic health measure—the SF12. Public health scientists developed the SF12 in order to *(i)* minimize measurement errors in self-reported health measures; and to *(ii)* develop a single comprehensive and continuous indicator of health. The SF12 is based on 12 health questions that are framed in a way to minimize response biases. Then, a special algorithm converts the responses to these 12 health questions into continuous summary scale measures of mental as well as physical health. The SF12 can be interpreted as a quasi-objective single measure of an individual's health status (RAND, 1995). This paper intentionally focuses on a comprehensive single measure of individual health rather than measures of the health care infrastructure, since the latter do not vary at the individual level.

By taking into account a rich set of socio-demographic, individual information, the paper nets out important health impact factors such as age, gender, employment, or the marital status. Importantly, the models also incorporate measures of health behavior such as the smoking status, alcohol consumption, and BMI. In addition, the empirical approach washes out individual variation in health that is associated with differences in health care utilization. Lastly, the models incorporate 17 county-level indicators that vary across the 401 German counties on an annual basis, such as the unemployment rate or the population density. Hence, the 3-level hierarchical models strive to unravel the conditional spatial structure of population health after having corrected for this rich set of individual and county-level factors just listed. Please note that the paper does not intent to provide any *causal* explanation or interpretation of the coefficients.

The findings show that large regional differences in health exist within Germany. Spatial interdependences and health clusters are of great importance and are not fully explained by observable individual and regional characteristics. Surprisingly, the systematic county-level impact on individual health is the most important driving force of the variable list above. County of residence may have an impact on individual health that equals 0.35 standard deviations in health. Particularly stunning is the clear East-West health divide that seems to be persistent more than 20 years after the fall of the Berlin Wall. It may reflect the 4 decade long impact of communism versus capitalism on health and results in an age impact of up to 5 life years for a 40-year old individual.[1]

The rest of this paper is organized as follows: In Section 2 we give a literature overview. Section 3 provides a short description of our underlying dataset, the German Socio-Economic Panel Study (SOEP). Sections 4.1 and 4.2 explain the econometric models and review details of Bayesian inference. Section 4.3 explains the different statistics used for model diagnostics. Section 5 presents and discusses the results and Section 6 concludes.

## 2 Literature Review

### 2.1 Geographical Variation in Health

Today the so called "Small Area Variation (SAV)" literature is very rich. It emerged from the seminal article, "Small area variations in health care delivery," by Wennberg and Gittelsohn (1973). The SAV Literature is extremely influential, particularly in the US. The DARTMOUTH ATLAS OF HEALTH CARE (2013) provides detailed descriptive information on how the health care infrastructure, health care utilization, and health care spending vary on a disaggregated geographical level, e.g., on the US county level. The high policy comes from the fact that utilization and spending measures are not systematically correlated with better health outcomes. Therefore, understanding the driving forces of regional differences might offer opportunities to limit health care spending while improving health outcomes (Wennberg et al., 2002).

Several studies use multilevel models to estimate associations between measures of health or health expenditures and possible determinants. For example, Michimi and Wimberly (2010) estimate associations between supermarket accessibility and obesity. However, in their synthesis of multilevel studies, Riva et al. (2007) point out that the vast majority of studies account only for within-area correlation and disregard between-area dependency. Not taking spatial patterns in the empirical models into account means that one implicitly assumes the geographical units to be statistically

[1] Note that this does not refer to a change in life expectancy but to the marginal effects of the age polynomial in our regression models.

independent. This might be a strong and misleading assumption since administrative or statistical boundaries might not reflect appropriately underlying ecological, social, and economic processes. Spillover effects are likely to occur.

Therefore, a newer subfield of the health economics literature explicitly focuses on spatial dependence in health, health care, and health care spending. The statistical and econometric spatial modeling approaches differ. Baltagi and Moscone (2010) and Moscone and Tosetti (2010) analyze the relationship between income and health expenditures using the Common Correlated Effects (CCE) Approach and impose a spatial autoregressive (SAR) structure on the error terms (Pesaran, 2006). Moscone et al. (2007) use a similar approach, but analyze cross-municipality variation in mental health expenditures. Felder and Tauchmann (2013) combine a nonparametric efficiency analysis with parametric regressions and model spatial dependence in healthcare provision at the county level in Germany. Bech and Lauridsen (2009) estimate a SUR model relating per capita expenditure for GPs in Denmark to local policies, health care supply as well as regional characteristics. They account for spatial dependencies using spatial autoregressive (SAR-SUR) and spatial autocorrelated (SAC-SUR) models. Filippini et al. (2009) model the demand for antibiotics and account for spatial correlation by introducing a spatial lag of antibiotics consumption. Lauridsen et al. (2010) study spatiotemporal dynamics of pharmaceutical expenditures in Spain. Their model incorporates a spatial autoregressive term as well as spatial lags of the covariates. Similarly, Costa-Font and Pons-Novell (2007) apply methods from spatial econometrics to study health expenditures in Spain and Ocaña-Riola and Mayoral-Cortés (2010) examine spatial mortality in Spain. A combination of spatial econometrics and panel data is applied by, among others, Lauridsen et al. (2008), Lagravinese et al. (2013) and Cohen et al. (2013). Most of the studies cited rely on aggregate regional data, which is reflected in the choice of the modeling techniques.

In contrast, Hierarchical Bayes Models offer a possibility to combine data surveyed on different levels (e.g., individual-level survey data and regional census data). Spatial variants of Hierarchical Bayes Models have been applied, for example, by Zhu et al. (2006), who study the effect of alcohol availability on violent crimes. Best et al. (2000) use spatial Hierarchical Bayes Models to study the health effects of exposure data measured at disparate resolutions. Mueller et al. (2001) and Kazembe and Namangale (2007) use models similar to those employed in this paper to study spatial patterns in child growth and child co-morbidity, respectively. However, they do not take the temporal dimension into account. This paper improves upon their work by introducing temporal and spatio-temporal dynamics, and decomposing the spatial effects into a part capturing spatial heterogeneity and spatial dependence (as suggested by Banerjee et al. (2004)).

### 2.2 Differences in Health within Germany

Several papers study the effect of the separation of Germany on a range of outcomes such as happiness, political attitudes, or trust (Frijters et al., 2004; Frijters et al., 2005; Fuchs-Schündeln and Schündeln, 2005; Alesina and Fuchs-Schündeln, 2007; Fuchs-Schündeln, 2008; Rainer and Siedler, 2009; Brosig-Koch et al. 2011; Burchardi and Hassan, 2013; Heineck and Süssmuth, 2013, Ziebarth and Wagner, 2013).

In addition, there are several papers studying regional differences in health and health care utilization in Germany, although most of these do not link their findings to Germany's division. Felder and Tauchmann (2013), Augurzky et al. (2013), Kopetsch and Schmitz (2013) and Eibich and Ziebarth (2014) study regional variation in health care efficiency and utilization, respectively. Pollack et al. (2004), Dragano et al. (2007, 2009a, 2009b), Breckenkamp et al. (2007), Voigtländer et al. (2010), and Diehl and Schneider (2011) use multilevel models to analyze the impact of specific regional factors on health. Their results indicate that regional deprivation (often measured by the unemployment rate) has a strong effect on individual health. Latzitis et al. (2011) investigate regional differences in mortality and Schipf et al. (2012) in prevalence rates of Type 2 diabetes mellitus. The latter study reports that the highest prevalence are in East Germany. Sundmacher et al. (2012) examine the spatial distribution of avoidable cancer mortality and report a north-south gradient, while Sundmacher and Busse (2011) find that the supply of physicians has a strong impact on avoidable cancer mortality.

Finally, there are a few studies that link their results explicitly to Germany's division. Riphahn and Zimmermann (1998) study the increase in mortality among East German men during the German reunification. Their results indicate that the increase in mortality is caused by an increase in individual stress. Müller-Nordhorn et al. (2004) investigate regional differences in ischaemic heart disease. They report that, although overall mortality has decreased since the 1990s, the East-West gradient remained constant. In contrast, Nolte et al. (2001) analyze infant mortality and report that the East-West gap found in 1991 had disappeared by 1997. Meyer et al. (1998) report that high-risk alcohol drinking was more common in East than in West Germany in 1991, while Nolte et al. (2003) find that mortality attributable to alcohol remained higher in East than in West Germany in 1997.

However, all aforementioned studies either ignore the possibility of spatial dependence (and therefore are prone to bias caused by omitting important regional predictors), or ignore socio-economic differences on the individual-level.

## 3 Data

### 3.1 Dataset

The empirical analysis makes use of individual-level health data and further socioeconomic characteristics provided by the German Socio-Economic Panel Study (SOEP). The SOEP is a representative panel study of private households. Starting in 1984, SOEP interviews subjects annually, and, since 1990, includes residents in former East German states. All respondents answer one main individual questionnaire covering about 150 questions on a range of topics, such as the labor market and family situation, attitudes and perceptions, as well as health. Additionally, a household questionnaire is completed by the head of the household. Since 2000 the survey has reached more than 20,000 individuals across 10,000 households. For further details, see Wagner, Frick and Schupp (2007).

The SOEP provides a variety of health measures. Both the standard 5-categorical Self-Assessed Health (SAH) measure and the 11-categorical health satisfaction measure are included every survey wave.[1] Although widely available and easy to collect, this paper does not make use of them since these simple self-rated health measures are shown to be prone to measurement errors (Juerges, 2007, Bago d'Uva et al., 2008, Ziebarth, 2010). Because of the known issues with these measures, since 2002 in every other year, the continuous quasi-objective SF12 measure and the objective grip strength measure have been included in the SOEP. Furthermore, information on health-related behavior (e.g. alcohol or tobacco consumption) is available since 2006. Therefore, we restrict our analysis to the three years of 2006, 2008 and 2010. We use only observations without item-non-response. In total we obtain 54,734 person-year observations from 23,414 different individuals.

### 3.2 Dependent Variable: SF12

The dependent variable is the generic health measure SF12. A specific algorithm generates the continuous SF12 measure on the basis of 12 health-related questions. More precisely, the algorithm generates eight subscales and two superordinate continuous dimensions, namely physical (*pcs*) and mental health (*mcs*). We average over both components to obtain the dependent variable *SF12*. In the standard SOEP version, by construction, the SF12 takes on continuous values between 0 and 100, has mean 50, and a standard deviation of 7 (see Table 1).[2]

As mentioned, the SF12 was developed by public health scientists to minimize measurement error, which is also called "reporting bias" or "reporting heterogeneity" in the context of self-reported health. To the extent that the SF12 eliminates reporting bias, we can view it as an objective health

[1] See Ayllón and Blanco-Perez (2012) for an application of the SAH measure.
[2] A detailed description of the algorithm and an overview over the differences between the original "SF12v2™ Health Survey" and the SOEP version can be found in Andersen et al. (2007).

measure. To the extent that the SF12 comprehensively captures the complex nature of individual health, we can also view it as a single, but comprehensive measure of health that is at the same time continuous. On the other hand, both assumptions may not entirely hold.

The SF12 also has advantages over purely "objective" physical health measures such as mortality rates, the grip strength, or diagnosis of specific diseases. First of all, by construction, these "objective" health measures define health very narrowly. For example, they totally lack any mental health dimension. Second, analyzing specific disease rates is based on the assumption that all existing diseases are diagnosed and recorded. Particularly when analyzing regional variation in disease rates, this may be a strong assumption since it assumes away regional differences that may affect the diagnoses of diseases, e.g. differences in the health care infrastructure. Finally, purely objective health indicators are mostly not available on a lower regional level, particularly not in combination with rich individual-level background information.

### 3.3 Individual Covariates

By controlling for demographic factors, educational characteristics, labor market participation, health-related behavior as well as health care utilization, we net out all differences in health as measured by the SF12 that can be traced back to these factors.

The ***demographic factors*** are an age polynomial of order three, gender, marital status, and the number of children under 14 in the household. Table 1 shows that average age is 49.3 and that about half of our sample is female. The majority of respondents are married and live in the same household as their spouse. There are no children in the majority of SOEP households.

In terms of ***education and labor market participation,*** we control for whether an individual completed a vocational training or holds a university degree, the working status of the respondent, and the monthly equivalent household income.[3] Less than a quarter hold a university degree, but about 60 percent completed vocational training. Slightly less than half of our sample works and the average equivalent gross household income per capita is €1,800 per month.

Concerning ***health-related behavior***, we control for alcohol and tobacco consumption and Body Mass Index (BMI). Alcohol consumption is captured by dummy variables. In the SOEP questionnaire, the participants are asked to state how often they drink wine and sparkling wine, beer, spirits and mixed drinks. If an individual states that they do not drink any alcohol at all, *no alcohol consumption* is assigned the value "1", otherwise it is "0". In contrast, if an individual states to drink any kind of alcohol on a regular basis, *regular alcohol consumption* takes on the value "1", otherwise

[3] The monthly equivalent household income uses the OECD-modified scale and assigns households a value of 1 to the first adult, 0.5 to each additional adult and 0.3 to each additional child. Further details are provided in *OECD Project on Income Distribution and Poverty (2009).*

“0” (Ziebarth and Grabka, 2009). 17 percent drink alcohol on a regular basis and 13 percent never drink any alcohol. The dummy variable *Smoking* captures whether the respondent consumes tobacco (cigarettes, cigars or pipes). It takes on the value “1” for about 30 percent of the respondents in our sample. *BMI* is measured as body weight in kilogram divided by the squared height in meters (Burkhauser and Cawley, 2008). The average BMI is slightly below 26.

***Health care utilization*** is measured by the number of hospital stays during the previous calendar year and the number of doctor visits in the 12 months prior to the interview.[4] Summary statistics for all variables are given in Table 1.

**[Insert Table 1 about here]**

### 3.4 County-Level Covariates

In order to explain systematic county-level differences in health, we incorporate time-variant county-level information into our analysis. For this purpose, we exploit information provided by the FEDERAL INSTITUTE FOR RESEARCH ON BUILDING, URBAN AFFAIRS, AND SPATIAL DEVELOPMENT (2013) (“*Bundesinistitut für Bau-, Stadt- und Raumforschung*”) in their INKAR (“INDICATORS AND MAPS ON SPATIAL DEVELOPMENT”) database.[5] The data itself comes from different official sources such as the FEDERAL STATISTICAL OFFICE, the statistical offices of the states as well as other governmental agencies. This also implies that this register data is surveyed on the county-level and not simply aggregated individual-level survey data.

The county-level variables net out differences in ***area and population***, ***standard of living and regional labor markets*** as well as the ***health care infrastructure.***

The ***area and population*** of a county are described by the *area size*, the *population density*, the *proportion of the area used for settlement and transport* (i.e. the degree of urbanization), the *proportion of the area suitable for recreation, the ratio of overnight stays to inhabitants* (as a measure of tourism)*, the population share of immigrants* and the *share of the population aged 65*.

Note that there is significant variation in population density across counties, ranging from 39 to 4,355 inhabitants per square kilometer. There is also wide variation in terms of degree of urbanization.

[4] It should be noted that the respondents are asked to state the number of doctor visits during the last three months prior to the interview. The answer is then multiplied by four to generate the annual number of doctor visits. This may overestimate the actual annual number because approximately two-thirds of all interviews are carried out in the first quarter of a year and there is a clear seasonal pattern to doctor visits.

[5] As for the individual level variables, we recode the county codes to the status as of January 1, 2012. However, the data taken from the INKAR database reflects the territorial status as of January 1, 2010. To take the 2011 county mergers into account, we calculate the values for the new counties as a weighted average of the values of the merging counties. The variables are weighted either by population number or by area size, depending on the respective variable (e.g. the *unemployment rate* refers to the population and is thus weighted by population count. In contrast, the *area used for settlement and transport* refers to the area and is therefore weighted by area size).

The values for this variable vary from 5 to 76 percent. Obviously, these indicators capture a great deal of cross-county heterogeneity that is not captured by individual-level indicators and might affect health. For example, the population density is highly correlated with the density of health care providers and the health care infrastructure. On the other hand, low urbanization rates or large recreational facilities may represent low air pollution and low noise, which in turn may affect health (e.g. Coneus and Spieß, 2012).

The ***standard of living*** is characterized by the average *available monthly income per capita*, the *GDP per capita*, the average *price for construction grounds* as well as the *car density*. The ***regional labor market*** is characterized by the annual *county-level unemployment rate* and the *share of minijobs* (i.e. low-paid jobs without social security contributions). Although we also include income and employment status at the individual-level, we hypothesize that the regional standard of living and its labor market may also influence individual health. For example, counties with a higher GDP per capita should be able to devote more resources to enhance public health. Average income per capita may serve as an indicator of regional deprivation and may also capture effects such as crime. The unemployment rate influences both job security and employment perspectives, and hence might influence an individual's health beyond her own employment status. Again, we observe huge variation across counties for all these indicators. For example, average per capita income varies from €1,108 to €2,701 and unemployment rates vary from 1.9 to 26.20 percent.

The ***health care infrastructure*** of the county is measured by the *number of physicians, GPs, hospital beds* and *nursing home places*, all measured per 10,000 inhabitants. In counties with fewer physicians or hospital beds, residents may experience longer travel and waiting times if they want to consult a physician, i.e., a lower degree of access to and utilization of the health care infrastructure and, consequently, health. On the other hand, following the conclusions of Wennberg and Gittelssohn (1973), physician density might be uncorrelated or negatively correlated with individual health. Also note that Germany has basically no provider networks and a free choice of doctors, i.e., gatekeeping does not exist. The uninsured rate is less than 0.5 percent and waiting times are short in an international comparison (OECD, 2011). The benefit package and cost-sharing amounts are regulated at the federal level and do not vary across regions.

### 3.5 Analyzing Variation in Health at the County Level

The SOEP provides information on the place of residence at multiple levels, ranging from federal states to specific postal codes (Knies and Spiess, 2007). The SOEP also provides frequency and probability weights to ensure representativeness for most federal states. The state-level might be too aggregate to detect significant regional effects – especially if these are caused by small scale phenomena and lower-level spillover effects (e.g. structural environmental and economic factors). Our main analysis is conducted at the county level.

One rationale for analyzing county-level clusters of health goes back to the fact that Germany is a federal state and counties are the smallest administrative units. Thus, political and institutional factors at the county-level may determine population health. Also, the institutional setup of the German health care system is partly determined at the county level. While hospital financing is determined by the 16 federal states, the physician density in the outpatient sector is determined by the regional self-governing associations ("*Kassenärztliche Vereinigungen*"). The planning level to determine whether there is an over- or undersupply of physicians and for the issuing of private practice licenses is the county level ("*Bedarfsplanung der Planungsregionen*").

Therefore, this paper focuses on counties ("*Landkreise*" and "*Kreisfreie Städte*"). Nevertheless using SOEP county-level data has some disadvantages. For instance, the data are not representative at the county level and there are no probability weights available to correct for over- and undersampling in certain regions. The number of observations per county varies between 2 (Sömmerda) and 2,025 (Berlin) with a mean of 136 observations and a standard deviation of 148. Thus the findings of this analysis should be interpreted carefully. For some counties the number of observations is too low to draw definite conclusions, particularly when it comes to health development of the county over time.

Therefore, we provide two additional robustness checks: First, we replicate the analysis at the state level, since the SOEP data is representative for most of the 16 German states. However, since the 16 states are relatively large, we expect a lot of regional heterogeneity within states, which would be averaged out in this robustness check. Therefore, second, we also replicate the analysis at the level of the 96 spatial planning regions ("*Raumordnungsregionen*"); the smallest still counts 137 observations. Spatial planning regions are areal units consisting of an urban center and the surrounding rural counties within the same state. They are similar to Metropolitan Statistical Areas (MSA) in the US. The advantages of these units are that the number of observations in each area is higher than at the county-level, while at the same time the units are small enough to reflect heterogeneity within states. However, the spatial planning regions are not administrative units, i.e. these areal units are simply used for reporting purposes.

As of January 1, 2012, Germany consisted of 402 counties.[6] Our dataset contains observations from 401 counties with Memmingen, a county in Bavaria, missing. We ignore this county in our analysis and remove it from the neighborhood matrix for the spatial models.

---

[6] The number of counties, and their respective borders, has evolved since 2006. To ensure consistency of the data with the shapefiles provided by the FEDERAL AGENCY FOR CARTOGRAPHY AND GEODESY (2013) ("*Bundesamt für Kartographie und Geodäsie*"), we recoded the county codes to reflect the boundaries as of January 1, 2012.

## 4 Hierarchical Bayes Models and Methods

### 4.1 Econometric Models

The aim of our analysis is to detect and model spatial patterns in the distribution of regional effects on health. With a few notable exceptions (discussed in Section 2.1), most studies using methods from spatial econometrics or spatial statistics rely on aggregate measures of morbidity (e.g., disease counts) as a dependent variable. In contrast, we model individual health. At the same time, we control for systematic differences between counties. In addition, we control for common time shocks across counties. The resulting models are inherently hierarchical with observations nested within individuals nested within regions. The following model serves as our baseline model throughout the analysis:

$$SF12_{ist} \sim N(\mu_{ist}, \sigma^2)$$

$$\mu_{ist} = \theta + X_{it}\beta + Z_{st}\gamma + c_i + b_s + \delta_t \qquad (1)$$

We assume that the *SF12* indicators for individuals $i = 1,...,23414$ in counties $s = 1,...,401$ and years $t = 2006, 2008, 2010$ are conditionally independent and normally distributed with mean $\mu_{ist}$ and variance $\sigma^2$.[7] The mean $\mu_{ist}$ is a linear function of an intercept $\theta$, the 15 individual regressors $X_{it}$ with parameter vector $\beta$, the 17 county-level covariates $Z_{st}$ with parameter vector $\gamma$, an individual effect $c_i$, a regional effect $b_s$ and aggregate time effects $\delta_t$.

Our interest lies mainly in the regional effects $b_s$ and their spatial distribution. In the following section, we specify three different assumptions on the distribution of these effects and their spatial and temporal dependency, which will lead us to our candidate models.

#### 4.1.1 Model 1: Unstructured Regional Effects (URE)

In our first model, we assume that the county-level effects follow a normal distribution with mean zero and variance $\sigma_b^2$:

$$b_s \sim N(0, \sigma_b^2).$$

[7] The SF12 does not follow a normal distribution by construction. This assumption implies that the idiosyncratic error term is normally distributed around mean zero and with variance $\sigma^2$. Strictly speaking, the SF12 is bounded from above and below. However, to allow for an easier interpretation and estimation we assume that a normal distribution works well for most (if not all) cases.

This implies that, on average, the regional effects are zero, which is reasonable since we include an intercept in our model. The interpretation of this model would be similar to a model with dummy variables for each county, i.e., county-fixed effects. However, it should be noted that these models are not methodologically the same. Given that the number of observations in some counties is relatively small, estimation of the county-level effects via dummy variables would be very inefficient. This problem is mitigated in hierarchical models through "borrowing of strength" across groups (see section 4.2).

We do not specify the type of spatial dependency in this model, which is why we refer to this model as the **Unstructured Regional Effects Model (Model 1).** In this initial stage of our analysis, we seek to obtain estimates of the regional effects whose spatial distribution we would like to analyze. These estimates are not smoothed by assumptions on the spatial dependency; they are used to examine the spatial dependency patterns of the regional effects by maps and test statistics. Then, in the next stage, we impose a structure on the spatial patterns and use models that smooth the regional effects.

### 4.1.2 Models 2 and 3: (Spatiotemporal) Convolution Prior (SCP)

In order to account for spatial dependency between counties, we model the regional effects using an Intrinsic Conditional Autoregressive (ICAR) model as proposed by Besag et al. (1991). However, the assumption that the whole regional effect depends upon its neighbors might be too strict. Instead, we would expect that there are both spatially dependent effects (e.g. spillovers) and random shocks.

Therefore, we decompose the effect

$$b_s = \varphi_s + \omega_s \,, \tag{2}$$

where

$$\varphi_s \mid \varphi_{r \neq s} \sim N(\overline{\varphi}_s, \frac{\sigma_\varphi^2}{m_s}) \,, \text{ and } \omega_s \sim N(0, \sigma_\omega^2) \tag{3}$$

that is, we discriminate between regional heterogeneity captured by $\omega_s$ and regional clustering captured by $\varphi_s$ (Banerjee et al., 2004). The spatially dependent effect $\varphi_s$ is normally distributed around the mean, where $\overline{\varphi}_s = m_s^{-1} \sum_{r \in \eta_s} \varphi_r$ . $m_s$ denotes the number of neighbors and $\eta_s$ is the set of neighbors of county *s*. The mean equals the average effect of the neighboring counties of county *s*. $\sigma_\varphi^2$ is a variance parameter. The variance of the spatially dependent effect depends on this parameter as well as on the number of neighbors. Thus isolated counties with fewer neighbors exhibit a greater variance than counties with more neighbors.

The conditional distribution in equation 3 specifies that the regional effects $\varphi_s$ are Markovian, i.e., they depend only on their neighbors.[8] From equation 3 the joint density can be obtained using Brook's Lemma (Besag, 1974) and the Hammersley-Clifford Theorem (Hammersley and Clifford, 1971). However, Besag (1995) notes that the resulting distributions would not exhibit appreciable correlations unless the parameters are close to the boundaries of the parameter space. Therefore, we use a pair-wise difference density as suggested by Besag et al. (1991), where the parameter determining the strength of the spatial association is fixed to the boundaries of the parameter space. Note that the resulting joint density is not a proper distribution (Banerjee et al., 2004). Nevertheless, if used as prior information about the unknown parameter in Bayesian Analysis, the resulting inference (through the posterior distribution) is proper.

We call the prior distribution imposed by equation 2 a ***Convolution Prior*** since the distribution of $b_s$ is a convolution of the distributions described above (eq. 3) (Besag et al. 1991; Mollié, 1996). Using this structural assumption and the neighborhood matrix based on adjacency, we obtain our **Model 2: the Convolution Prior Model (CP)**.

This model incorporates spatial dependence and aggregate time effects. However, the spatial effects $b_s, \varphi_s$ and $\omega_s$ do not depend on time *t*, i.e., they are assumed to be time-invariant. This may be a plausible assumption if the main driving factors of the regional differences are time-invariant, e.g., structural environmental and economic factors. Nevertheless, we check this assumption by extending the CP-Model to include space-time interaction effects. As such, we obtain our third and final model: the **Spatiotemporal Convolution Prior (SCP)**. Consider the model

$$\mu_{ist} = \theta + X_{it}\beta + Z_{st\gamma} + b_{st} + c_i + \delta_t, \tag{4}$$

where

$$b_{st} = \varphi_{st} + \omega_{st},$$

$$\varphi_{st} \mid \varphi_{rt \neq st} \sim N(\overline{\varphi}_{st}, \frac{\sigma^2_{\varphi t}}{m_s}) \text{ and } \omega_{st} \sim N(0, \sigma^2_{\omega t}).$$

Here, both the spatially dependent part and the random part of the regional effect depend on location *s* and time *t*, i.e., the model produces a separate effect for each county and each year (comparable to an interaction between dummy variables for county *s* and year *t*). This model has the advantage that we can estimate the spatial pattern for each year. However, it requires observations for each county in each year. Therefore, for **Model 3**, we restrict the dataset to counties with observations in each year. The adjusted dataset contains 54,723 out of 54,734 observations and 398 out of 401 counties.

---

[8] The model specified through eq. (3.1.2) is also called the Autonormal or Autogaussian model, i.e., a model for normally distributed data on a Markov Random Field (Arbia, 2006)

### 4.2 Bayesian Inference and Estimation

The models described in Section 4.1 are formulated in a Bayesian framework. Bayesian methods offer some important benefits for this analysis. Probably most important for our analysis, hierarchical and spatiotemporal models can be adapted quite naturally in a Bayesian framework, since our assumptions on spatial dependence can enter the model via the prior distributions (Congdon, 2010). Furthermore, Hierarchical Bayes Models impose a common prior distribution on group-level effects (i.e. individual and regional effects in our models). As noted above, this "borrowing of strength" increases the efficiency of the estimation (Congdon, 2010). Thus, we formulate our models as Hierarchical Bayes Models.

However, this also implies that our results have to be interpreted accordingly. In Bayesian inference, given the data, the main interest lies in learning about the distribution of the unknown parameters (the so-called **posterior distribution**).[9] In our analysis, we obtain a random sample from the posterior distribution through MCMC methods, e.g. Gibbs sampling. This distribution is then summarized to obtain point estimates and Bayesian confidence intervals (also called **credible intervals**). In this paper we use the mean of the distribution for our point estimates.

The confidence intervals in this paper are derived as **Equal Tail (ET) intervals.** The ET interval is given as the interval $[q_L, q_U]$ where:

$$\int_{-\infty}^{q_L} p(\theta \mid y) d\theta = \frac{\alpha}{2} \text{ and } \int_{q_U}^{\infty} p(\theta \mid y) d\theta = 1 - \frac{\alpha}{2}.$$

As usual, $\alpha$ denotes the confidence level. An attractive feature of the credible interval is that the often erroneous interpretation of confidence intervals in frequentist models does actually hold for them, namely the probability that $\theta$ lies in $[q_L, q_U]$ is $(1-\alpha)$ (Banerjee et al., 2004). However, this also implies that standard significance tests are not applicable in this framework. The literature suggests the use of model selection criteria to determine which variables are significant. Nevertheless, this is impractical for assessing the significance of single covariates, since it would require estimating the model twice for each covariate. Therefore, in this paper we use a model selection criterion to assess the significance of certain groups of variables as well as other crucial features of the model. In addition, we provide pseudo p-values to assess the significance of single predictors. These pseudo p-values are calculated as the probability that the parameter has a different sign than the point estimate. A small pseudo p-value implies that it is very likely that the predictor has indeed a positive (or negative) effect.[10]

[9] It is worth noting that the parameters are not regarded as random quantities - they are fixed but unknown. The incomplete knowledge about these parameters is assumed to be random (Gelman and Robert, 2013).
[10] The true effect could still be very small, however, since it is not possible to test whether it is exactly zero.

The prior distributions for our analysis are mostly uninformative. They allocate equal probability mass to all plausible values. In detail, we choose a flat prior $U(-\infty,\infty)$ for the intercept $\theta$ and a normal prior distribution with an inflated variance $N(0,10000)$ for the slopes $\beta_1,...,\beta_{15}$ and $\gamma_1,...,\gamma_{17}$ and the time fixed effects $\delta_t$.

For the regional effects we choose the prior distributions as described in section 4.1. All standard deviation parameters, including hyper-parameters, were assigned a $U(0,100)$ prior distribution.[11] In order to speed up convergence, we standardize all non-dummy variables.

We run three parallel Markov chains with "dispersed"[12] initial values and monitored the values for the intercept, the slope parameters, the regional effects, and the variance parameters using **Trace Plots** and **Gelman-Rubin statistics** (Brooks and Gelman, 1998; Gelman and Rubin, 1992). The length of the burn-in period was determined for each model separately.

For Model 1, we could not detect any departure from convergence after the 5,000$^{th}$ iteration. For Model 2, the burn-in period took 15,000 iterations, after which the sampler stabilized. For Model 3 20,000 iterations were needed. After the burn-in period, we sampled 15,000 draws from the posterior distribution. In order to decrease autocorrelation and speed up "mixing,"[13] we thinned the chain by storing only every 10$^{th}$ draw. The estimation was carried out using WinBUGS. Data handling and post-estimation was done in Stata, R and OpenGeoDa.[14]

### 4.3 Testing for Spatial Dependency and Model Selection

We assess the appropriateness of the models by implementing tests for spatial dependency as well as a model selection criterion. These indicators are described in the following section.

[11] In the CP model we assign a U(0.1,100) prior distribution to the parameter $\sigma_\varphi$ to improve the stability of the sampler. Although this excludes values between 0 and 0.1 from the parameter space, this assumption seems reasonable given the results from the other models as well as previous analyses. Furthermore, if the true value of the parameter would lie in the excluded interval, we would expect that the posterior mean is relatively close to the left boundary of the prior distribution. This is not the case here.

[12] Every chain starts at different initial values. A simple way to monitor convergence of the sampler is to run parallel chains that start at different initial values. If the chains "overlap" each other, i.e., the draws from the chains fluctuate around the same value, this could be interpreted as a sign of convergence. However, if the chains start at the same initial value, any overlap could be pure coincidence. We initialize the first chain at the mean values of the prior distributions and the second and third at the upper and lower end of the range of plausible values respectively.

[13] I.e. the number of iterations needed for the chain to cover the whole support of the posterior distribution.

[14] We especially acknowledge the use of an R-script written by T. Elrod as well as an Ado-file written by J. Thompson, T. Palmer and S. Moreno.

***Testing for Spatial Dependency using Moran's I***

We calculate MORAN'S *I* (Moran, 1950) as a measure of the strength of the spatial association. It is also used to formally test the hypothesis of no spatial autocorrelation.[15]

MORAN'S I can be interpreted as a spatial analogue for the lagged autocorrelation coefficient statistic in time series analysis. It should be noted that we do not test for spatial correlation in the data but for spatial correlation between the regional effects estimated in Model 1. To assess the significance of the spatial dependence, we carry out a Monte Carlo permutation test by drawing a random sample of permutations[16] including the observed one. Then, the observed *I* can be positioned relative to the other permutations and a pseudo *p*-value may be derived.

***Local Indicators of Spatial Association (LISAs)***

MORAN'S I is a measure of global spatial autocorrelation. In contrast, LOCAL INDICATORS OF SPATIAL ASSOCIATION (LISAS) are calculated for each county individually and can be used to test the significance of the spatial association for a specific county. This has the advantage that heterogeneity in the strength or direction of the spatial dependence can be detected. We calculate local *I*'s as proposed by Anselin (1995). Again, we use a Monte Carlo permutation test to obtain pseudo *p*-values.

***Model Selection***

We use the Deviance Information Criterion (DIC) introduced by Spiegelhalter et al. (2002) to compare our candidate models. The DIC is a criterion for model selection similar to the Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC), i.e., it trades off model fit[17] and complexity.[18] As with AIC and BIC, the model with the lowest DIC should be preferred. In contrast to AIC and BIC, the DIC is valid for hierarchical models and can be easily computed as a by-product of MCMC methods. However, the scale of the DIC has no meaning by itself since it includes a term that depends solely on the data. Therefore, only differences between DIC can be interpreted. There is no distinct rule when differences in DIC are considered significant. We use the rule of thumb suggested by Spiegelhalter et al. (2002) who suggest that differences larger or equal to 10 should be considered significant.

---

[15] Arbia (2006) points out that MORAN'S I is not a proper statistical test since it does not consider an explicit alternative hypothesis. Since we use MORAN'S I as an exploratory tool before specifying our spatial models, this can actually be seen as an advantage, since we do not have to consider a single alternative hypothesis and instead explore several possible forms of spatial correlation. (Arbia, 2006).

[16] The observed values of the variable are randomly assigned to the regions.

[17] Model fit is measured through the posterior expected deviance.

[18] Complexity refers to the effective number of parameters calculated as the difference between posterior expected deviance and the deviance at the posterior means.

We also provide the BIC for comparison. However it should be noted that the BIC measures model complexity solely through a function of the nominal number of parameters. This means it does not account for the more efficient estimation of the individual and regional effects, i.e. the “borrowing of strength.” Therefore it is not an appropriate criterion for hierarchical models.

# 5 Results

## 5.1 Regional Differences in Health

First, we estimate and plot the county-level means of the SF12 variable to get an impression of the magnitude and pattern of the regional differences. The results are displayed in the left CHOROPLETH MAP in Figure 1 below. Accordingly, regional deviations from the national mean range from -14 to 7 points in the SF12 health status measure (Figure 1a).

**[Insert Figure 1 about here]**

However, these large differences are not very meaningful, since they could potentially be explained by differences in demographic factors, e.g. age or gender. We account for demographics by calculating age-sex-adjusted county-level means by estimating Model 1 (see Section 3.1) using only age, age squared, cubic age and gender as explanatory variables. The estimated regional effects are shown in Figure 1b. Note that the magnitude of the county-level differences has decreased as compared to Figure 1a. However, regional differences still amount to about 40 percent of the standard deviation of the SF12 measure. A comparison of the two maps in Figure 1 also shows that the estimated patterns are very similar, i.e., gender-age differences might partly explain the magnitude but not the distribution of the regional effects. Furthermore, the map gives first evidence that there are several clusters of positive and negative regional effects.

## 5.2 Model 1: Unstructured Regional Effects (URE)

Since age and gender cannot explain the regional differences observed in the mean values, we estimate the ***Unstructured Regional Effects (URE) Model*** (see Section 4.1.1), which includes the whole set of individual and regional predictors described in Section 3. Table 2 yields the parameter estimates for these variables. Please note that the coefficients are standardized for the non-dummy variables. All **individual-level variables** are statistically significant at conventional levels.

**[Insert Table2 about here]**

Looking at the sign of the coefficients, we note that, in general, ***males*** are healthier than females. This is in line with the stylized fact that women have greater health care expenditures, not only in childbearing years (Owens, 2008). Also, individuals who abstain from ***alcohol*** have a lower health status. This group is likely to contain respondents with current or past serious diseases. On the other hand, regular alcohol consumption is positively correlated with health; this might seem surprising, but is consistent with the literature (Ziebarth and Grabka, 2009). ***Health care utilization*** is negatively correlated with respondents' health. Concerning the size of the effects, gender, alcohol abstinence and ***smoking*** status show the largest associations with individuals' health.

Of the **county-level variables**, the share of elderly people, GDP per capita, the share of minijobs (i.e. low-paid, €400 per month, jobs without social security contributions) as well as the number of physicians show a significant association with health. At first glance, the positive association of ***elderly people*** with health seems counterintuitive; however, a higher share of older people might be connected with a higher longevity of the population, or the fact that communities with a large number of senior citizens provide better health resources (e.g. public health centers). The negative associations of ***average household income*** and ***the number of GPs*** with health might be explained by the fact that we already controlled for GDP per capita and ***the overall number of physicians*** in the model. Since average income is highly correlated with GDP per capita as is the number of physicians and the number of GPs, the clearly positive effects of standard of living and the health care infrastructure are likely captured by these more general predictors. The positive effect of the ***share of minijobs*** might imply that the labor market in a region is more flexible, hence individuals that otherwise would have been unemployed are instead able to take up these low-paid jobs. However, the economic significance of these predictors is rather small. A change of one percentage point in the share of minijobs is associated with a 0.09 point higher health status. Compared with the impact of individual covariates this correlation is rather small. However, the main focus of our analysis lies on the regional effects $b_s$ and their spatial distribution that we discuss in detail below.

***Spatial Pattern of the regional effects***

The "posterior means"—the point estimates of the regional coefficients in the URE-Model—are plotted in form of a quantile map on the left-hand side in Figure 2.

**[Insert Figure 2 about here]**

We observe the following: First, the county level health effects differ between -2.4 and 2.7.[19] This is equivalent to 0.38 standard deviations in health and implies that the average health status varies between 46.4 and 51.5, depending on the county.[20]

[19] Of course not all of these effects are statistically significant. At the 5 percent level, 92 of the 401 regional effects are significant; 152 are significant at a 10 percent level.
[20] Standard Errors and Monte Carlo Errors are excluded from the calculation.

Second, it seems as if the counties in East Germany tend to have lower health values than those in West Germany. Overall, the map is dominated by clusters: We find clusters of high values in the Northwest (around Hanover and Hamburg), West (Cologne) and Southwest (Palatinate region). Clusters of smaller values can be found in the Southeast (Lower Bavaria), Center (Thuringia) and in the Northeast (Mecklenburg-Western Pomerania and Brandenburg).

However, the number of observations in some of the 401 counties is very small. Therefore, as a robustness check we estimate the model on the level of the 96 spatial planning regions (see section 3.5). Even the smallest of the 96 spatial planning regions still count 137 observations. The estimated area effects are shown in the map of Figure 2b, while Figure 2a displays the county-level results. The spatial pattern is very similar to the results on the county-level. The size of the effect ranges from -2.3 to +1.7, i.e. the lower bound is almost the same as on the county-level, whereas the upper bound has shrunken. Most important, we can identify the same clusters as on the county-level. We also notice that most positive clusters are found in West Germany, whereas the clusters of negative values are located in the former East.[21]

These clusters already suggest the presence of spatial dependence between counties. We calculate MORAN'S *I* to formally test the hypothesis of no spatial autocorrelation. We conduct the spatial analysis using a neighborhood matrix based on first-order adjacency. This means that sites are considered to be neighbors if they share a *common border*. This definition is advantageous since it results in a symmetric neighborhood matrix, which is required by the ICAR specifications as described in Section 4.1.2.

Nevertheless, for illustrative purposes, we calculate MORAN'S *I* using eight additional neighborhood matrices based on second- and third-order adjacency, threshold-distances between the county centroids, and Nearest-Neighbor algorithms[22] and test the significance of *I* using 10,000 permutations, including the observed values (Arbia, 2006). Note that all matrices are binary and that the adjacency- and threshold-based definitions lead to symmetric matrices.

The results are given in Table A.2 of the Online Appendix**.** In general, MORAN'S I is not very high, but all values, with the exception of the third-order adjacency matrix, are significant. We also observe that the value of *I* decreases with the size of the neighborhood (as defined by the number of links). The highest value ($I=0.19$) is obtained for the 3-Nearest-Neighbors matrix, whereas the distance-based and higher-order adjacency matrices result in considerably smaller values. Our preferred first-order

[21] We also estimate the model on the level of the 16 Federal states. As expected, we find that the regional effects are considerably smaller. Furthermore, the spatial pattern is quite different, due to the fact that a lot of the spatial heterogeneity occurs within states. However, we still observe a distinct East-West difference. The coefficient estimates for both models are given in Table A.1 in the Online Appendix. The map for the Federal states is available upon request.

[22] The Nearest-Neighbor algorithm determines the neighbors of a region *i* by comparing the distances between the centroids. Then, the *n* regions with the lowest distance are chosen as neighbors.

adjacency matrix shows a comparably high value of $I=0.18$, which is highly significant. This definition also results in comparably small neighborhoods. Thus we conclude that the spatial dependence between county-level health effects occurs on a small, local scale. All in all, we reject the hypothesis of independent regional effects and employ a first-order adjacency matrix for the spatial models as described in Section 3.[23]

***Local Indicators of Spatial Association (LISAs)***

We also calculate the local Moran statistic to investigate in which parts of Germany the spatial correlation is significant on a local scale. The *p*-values of the Local MORAN'S I for each county are depicted in Figure A.1 of the Online Appendix. The map corresponds to an adjacency-based neighborhood matrix. The classes correspond to different types of clusters. They are formed according to the significance of the spatial association, the sign of the regional effect as well as the sign of the spatial correlation. "High-high" stands for counties with positive regional effects and positive spatial correlations, i.e., the regional effects of its neighbors are also positive (Anselin, 2005). We see that for most of the clusters identified above, the spatial association is significant on a local level.

### 5.3 Model 2: Convolution Prior (CP)

Now, we run Model 2 using a convolution prior as described in Section 4.1.2 to obtain smoothed estimates of the spatial pattern. This model has the advantage that it can distinguish between unstructured (i.e., random) regional effects and spatially dependent regional effects. If we plot the overall regional effect, the resulting maps look very similar to URE-Model (Model 1). However, the values of the spatially dependent random effects show a significantly smoother pattern. The map in Figure 3a illustrates the results of Model 2.

The map presents the values for the spatially dependent part of the regional effect $\varphi_s$. First, we observe a prominent cluster of high values in the Northwest (Lower Saxony, Schleswig-Holstein) and a smaller one in the Southwest (Rhineland-Palatinate).

Second, clusters of lower values are found in the (South-)East (Bavaria, Thuringia and Brandenburg). This suggests a pattern in the spatial distribution.

Third, in West Germany, unobserved but systematic county-level effects show a strong positive association with residents' health. Contrarily, in East Germany, county-level effects show a strong negative association with residents' health. Remember that we control for a rich set of socio-economic individual background characteristics such as age, gender, marital, and employment status as well as health-behavior and the degree of health care utilization. In addition, 17 county-level covariates net

[23] The strength of spatial dependence on the level of the spatial planning regions is similar, with a value of $I=0.14$. On the state level, MORAN'S I is much larger with $I=0.31$.

out persistent differences across counties due to unemployment, urbanization and population density. Given this modeling approach, it is surprising and staggering that we still find a clear East-West health pattern, 20 years after the German reunification.[24]

**[Insert Figure 3 about here]**

The map on the right-hand side in Figure 3 shows a LISA cluster map for Model 2. The counties with significant local *I* are color-coded with respect to the type of spatial correlation. We identify three clusters of high values, mainly in the Northwestern part of Germany and clusters of low values in the Eastern part (Bavaria and Thuringia). This again supports the notion of an East-West trend in the spatially dependent component. Deviations from this trend in Western Mecklenburg-West Pomerania are likely due to spillover effects.

The finding of low regional health patterns in regions with former communist governments during the Cold War is consistent with the existing literature. For example Baltagi et al. (2012) and Bonneux et al. (2010) report significantly lower life expectancy in Eastern Europe, specifically Hungary, Czech Republic, Slovak Republic and Poland, when compared to Western Europe. In addition, Baltagi et al. (2012) find significant spillover effects in the healthcare production process across neighboring countries. Disease epidemics may also be an explanation for regional spillover effects of health. Treurniet et al. (2004) report higher avoidable mortality rates for Hungary and the Czech Republic as compared to Western Europe between 1980 and 1997.

### 5.4 Model 3: Spatiotemporal Convolution Prior (SCP)

Finally we estimate the Spatiotemporal Convolution Prior (SCP) Model in order to investigate changes in the spatial pattern over time. The results are shown in Table 3. First, the parameter estimates of the individual covariates are robust when we compare them to the findings from our first model, the URE-Model in Table 2.

---

[24] Note that, theoretically, there could be a plethora of reasons for this distinct pattern. However, all of them could at least be interpreted as an indirect consequence of the 40 year long division of Germany. For example, it is shown that unemployment has a causal impact on individuals' health and well-being (Sullivan and van Wachter, 2009; Kassenboehmer and Haisken-DeNew, 2009). The structurally higher unemployment rate in East Germany could be cited as one potential explanation for the East-West health differential. However, note that we control for the individual employment status as well as the county-level unemployment rate. Even in East Germany, county-level unemployment rates vary widely between 6.9 and 26.2 percent. Another potential explanation might refer to migration. Between 1989 and 2005, 3.4 million, primarily young and healthy, East Germans migrated to West Germany (Hunt, 2006). However, many returned to their home states since the economic conditions have improved considerably. Furthermore, in the SOEP data, only 6.6 percent or 962 individuals living in West Germany in 2006 were born in East Germany. If we subtract these individuals from the West German sample, the average SF12 decreases only slightly from 50.08 to 50.05. Similarly, if we add individuals who migrated from East to West Germany to the East German sample, the average SF12 increases only slightly from 49.01 to 49.2. All these differences are not statistically significant. Thus it is very unlikely that this marked difference has been caused by migration.

Second, among the county-level covariates the variables describing the labor market (***unemployment rate*** and the ***share of minijobs***) show a strong and significant association with health. While technically the ***degree of urbanization*** shows the strongest association with health, the estimate is only significant at the 10 percent level, and hence might be a statistical artifact. As in Model 1, ***GDP per capita*** and ***the share of elderly people*** show a strong and positive association with health. The ***price for construction ground*** is now significant on a 1 percent level (in contrast to the 10 percent level in Model 1) and shows that individuals in wealthier and more attractive counties are on average healthier. The ***number of hospital beds*** shows a significant and negative correlation with individual health. This is not surprising, since it is in line with the results established by Wennberg and Gittelsohn (1973). In fact, this would be expected if the health care infrastructure is planned based on demand. However, as in Model 1 most of the effects are rather small, especially if compared with the effects of individual-level covariates. For example, a ten percent higher unemployment rate is associated with a 0.6 points lower health status, which is only slightly more than half of the effect of male gender and alcohol abstinence. This might suggest that the large county-level differences might be caused by selection on unobserved individual characteristics or important omitted county-level predictors. However, the results also demonstrate that economic factors (such as the labor market) play an important role for population health, even beyond the effect of individual income and employment status. Of course, the estimated coefficients are only correlation and cannot be interpreted as causal effects.

**[Insert Table 3 about here]**

Figure 4 illustrates the spatial patterns for each year. Note that only the spatially dependent part of the regional effect is displayed. First of all, the size of the effects varies across years, but they largely overlap: in 2006, the values fall between -1.8 and 2, in 2008, they fall between -1.3 and 0.9 and in 2010, they fall between -1.7 and 1.2.

**[Insert Figure 4 about here]**

Second, the spatial pattern itself varies slightly across time. However, since the majority of the regional effects are not significant, this is probably just random variation. In those counties with an especially small sample size, the results are also highly sensitive to panel attrition. Even more importantly, the clusters with significant spatial dependence, e.g., the regions in the Northwest (Hanover), Southwest (Rhineland-Palatinate), Center (Thuringia) and Southeast (Bavaria), are stable across the years. Maps of the pseudo *p*-values of the regional effects[25] ($\varphi_{st}$) show that the regional effect in these clusters is significant in all years, whereas most of the effects that change across the years are insignificant.

[25] The figures are available upon request.

Finally, to explicitly test for the East-West-German health pattern, we include an East-West dummy variable in Model 3. The dummy is “1” for counties in pre-1989 communist East Germany and “0” for counties in capitalist West Germany. The highly significant coefficient estimate equals -1.2 and represents one of the strongest health predictors. Its impact is equivalent to a change of 0.17 standard deviations in health. If we compare this effect to the partial effect of age in our model, we find that living in East Germany is equivalent to an age effect of one and a half life years. Furthermore, if we take into account the nonlinearity of age, the effect of living in East Germany is comparable to an age impact of up to 5 life years for a 40-year old (i.e. an individual who was 20 years old at the time of the reunification of Germany).[26]

### 5.5 Model Fit and Robustness

We use the Deviance Information Criterion (as discussed in section 3.3) to compare the model fit of our four main models. The results are given in Table A.3 in the Online Appendix. They indicate that the Convolution Prior-Model 2 performs less well, both in terms of model fit and complexity. This is no surprise: we impose additional structural assumptions; hence we obtain a more complex model. A comparison with Model 3 shows that this Spatiotemporal Model offers a much better model fit than the purely spatial model. This suggests the existence of space-time interactions. The gain in model fit more than compensates for the higher complexity of Model 3. Surprisingly, the model fit ($\overline{D}$) is even better than for the Unstructured Regional Effects Model (Model 1). This implies that our Model 3 should be preferred among all estimated models. Although the DIC for the SCP-Model 3 is larger than the DIC for the URE-Model 1, this effect is solely driven by the complexity of the model ($p_D$).

The BIC shows a different pattern than the DIC; however it should be noted that the BIC measures model complexity solely through a function of the nominal number of parameters. This means it does not account for the “borrowing of strength”. The $R^2$ shows that our preferred SCP-Model 3 explains 46 percent of the observed overall variation in health status.

To assess the robustness of our results, we estimate four additional variants of the URE-Model: the “model without endogenous covariates” variant does not include information on health care utilization, health-related behavior and employment status. The “model without regional covariates” variant includes all individual covariates, but not the regional predictors. The DIC implies that the model without endogenous covariates is much worse than the URE-Model 1, i.e., the excluded individual covariates explain a lot of the observed variation in health.

The model without regional covariates results in a very similar model fit as the URE-Model 1 including regional covariates. In other words, the regional predictors have little to no explanatory

[26] The distinctive nonlinear impact of age on health is partly due to the fact that we account for both physical and mental health. While physical health decreases with age, mental health increases between (roughly) age 50 and age 70.

power. However, it should be noted that this variant includes regional effects, which might have picked up some of the variation explained by the regional covariates. A comparison of the pattern of the regional effects shows that our conclusions about the spatial structure are not affected by the exclusion of the regional covariates. We also estimate a variant of Model 1 without regional covariates and regional effects, i.e., the variation in health is purely modeled on the individual level. Here, the model fit is much worse than in the model with regional information, which indicates that the place of residence explains a significant part of the variance in individual health.[27]

Lastly, we estimate a model with a quintic age polynomial to ensure that the age effect is correctly specified. Since this does not significantly improve the model fit, we conclude that the cubic age polynomial is sufficient.

We also check for multicollinearity. The entries of the correlation matrix for the individual variables are not larger than 0.3. Among the regional covariates, we find several high correlations (e.g. between average income and the unemployment rate). However, since the exclusion of the regional covariates does not affect our conclusions, multicollinearity does not pose a problem.

## 6 Conclusion

This paper combines representative individual-level household panel data and register county-level data to model and estimate spatial health effects on individual health. Methodologically, it makes use of hierarchical models and Bayesian methods.

In a first step, we examine regional effects for spatial associations using adjacency-based, distance-based and Nearest-Neighbor-based definitions of neighborhoods. In all cases, we find a highly significant spatial dependency of individual health. In the next step, we impose structural ICAR assumptions onto our model and re-estimate it, using a convolution prior based on adjacency. For this model we find a trade-off between model fit and smoother estimates. We find the model fit for our Spatiotemporal Model with separate regional effects for each year to be much better than the model fit for our basic Regional Effects Model without structural assumptions.

In terms of content, this paper shows that the regional association with residents' health is systematic and strong. The general county-level predictor has the strongest impact among all individual- and county-level predictors considered. The regional impact is equivalent to 0.35 standard deviations in health. This suggests the presence of important unobserved regional determinants of individual health.

[27] Note that the time-invariant individual effect might pick up some of the information contained in the time-invariant regional effect for individuals who have not moved to another county. Therefore, the amount of variation explained by the place of residence is potentially even larger.

Interestingly and surprisingly, even 20 years after German reunification, we still find a clear East-West spatial health pattern. This finding could be interpreted as the long-lasting health effect of the 40 year long division of Germany into a Communist and a Capitalist Part. The long-lasting legacy of Communism on health equals an age effect of up to 5 life years for a 40-year old.

Lastly, our results show that regional economic factors (e.g. community income or the labor market) have a strong and significant association with individual health, even beyond the impact of individual income and employment status. This demonstrates the need for policymakers to consider the consequences of economic policies for public health. Taken together with the finding of a large heterogeneity within states, it further shows that policymakers should consider place-based public health interventions as demonstrated by Bhattacharjee et al. (2012). This would offer the possibility to improve public health where it is needed the most, and to exploit spillover effects. However, our empirical results are based on correlations and cannot be interpreted as causal effects. A next step would be to establish an identification strategy in order to test whether these regional effects have a causal interpretation or arise due to selection processes or omitted variable bias. Clearly, more research on and applications of modeling techniques for spatial patterns would be fruitful.

## 7 References


Alesina, A. and Fuchs-Schündeln, N., 2007. Good Bye Lenin (Or Not?): The Effect of Communism on People's Preferences. *American Economic Review*, 97(4), 1507-1528.

Andersen, H. H., Mühlbacher, A., Nübling, M., Schupp, J., and Wagner, G. G., 2007. Computation of standard values for physical and mental health scale scores using the SOEP version of SF12v2. *Journal of Applied Social Science Studies (Schmollers Jahrbuch)*, 127(1), 171-182.

Andersen, L. B., Bech, M., and Lauridsen, J., 2012. Political or dental power in private and public service provision: a study of municipal expenditures for child dental care. *Health Economics, Policy and Law*, 7(3), 327-342.

Anselin, L., 1995. Local indicators of spatial association – LISA. *Geographical Analysis*, 27(2), 93-115.

Anselin, L., 2005. Exploring Spatial Data with GeoDa™: A Workbook. Available at http://www2.geog.ucl.ac.uk/~jessletzbichler/GEOG2003/geodaworkbook.pdf, last accessed on May 1, 2012.

Arbia, G., 2006. *Spatial Econometrics: Statistical Foundations and Applications to Regional Convergence.* Springer. Berlin Heidelberg.

Augurzky, B., Kopetsch, T. and Schmitz, H., 2013. What Accounts for the Regional Differences in the Utilisation of Hospitals in Germany? *European Journal of Health Economics*, 14 (4), 615-627.

Ayllón, Sara and Blanco-Perez, Cristina, 2012. State Dependence in Self-Assessed Health in Spain. *Hacienda Pública Española*, 202(3), 9-30.

Bhattacharjee, A., Maiti, T. and Petrie, D., 2012. Spatial structures of health outcomes and health behaviours in Scotland: Evidence from the Scottish Health Survey. Paper presented at the III Health Econometrics Workshop, Siena 2012.

Bago d'Uva, T., O'Donnell, O. and van Doorslaer, E., 2008. Differential health reporting by education level and its impact on the measurement of health inequalities among older Europeans. *International Journal of Epidemiology*. 37(6), 1375-1383.

Baltagi, B. H., and Moscone, F., 2010. Health care expenditure and income in the OECD reconsidered: Evidence from panel data. *Economic Modelling*, 27(4), 804-811.

Baltagi, B.H., Moscone, F. and Tosetti, E., 2012. Medical technology and the production of health care. *Empirical Economics*, 42(2), 395-411.

Banerjee, S., Carlin, B. P. and Gelfand, A. E., 2004. *Hierarchical Modeling and Analysis for Spatial Data*. Chapman & Hall/CRC. Boca Raton.

Barufi, A. M., Haddad, E. and Paez, A., 2012. Infant mortality in Brazil, 1980-2000: A spatial panel data analysis. *BMC Public Health*, 12(1), 1-18.

Bech, M. and Lauridsen, J., 2008. Exploring the spatial pattern in hospital admissions. *Health Policy*, 87(1), 50-62.

Bech, M. and Lauridsen, J., 2009. Exploring spatial patterns in general practice expenditure, *The European Journal of Health Economics*. 10(3), 243-254.

Besag, J., 1974. Spatial interaction and the statistical analysis of lattice systems (with Discussion). *Journal of the Royal Statistical Society*,Series B 36, 192-236.

Besag, J. and Kooperberg, C., 1995. On conditional and intrinsic autoregressions. *Biometrika*, 82(4),733-746.

Besag, J., York, J.C. and Mollié, A., 1991. Bayesian image restoration with two applications in spatial statistics (with discussion). *Annals of the Institute of Statistical Mathematics*, 43(1), 1-59.

Best, N.G., Ickstadt, K., and Wolpert, R. L. (2000). Spatial Poisson Regression for Health and Exposure Data Measured at Disparate Resolutions, *Journal of the American Statistical Association*. 95(452): 1076-1088.

Bonneux, L.G., Huisman, C. and de Beer, J.A., 2010. Mortality in 272 European regions, 2002–2004. An update. *European Journal of Epidemiology*, 25(2), 77-85.

Breckenkamp, J., Mielck, A., Razum, O., 2007. Health inequalities in Germany: do regional-level variables explain differentials in cardiovascular risk? *BMC Public Health*, 7: 132.

Brooks, S.P. and Gelman, A., 1998. General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics*, 7(4), 434-455.

Brosig-Koch, J., Helbach, C., Ockenfels, A. and Weimann, J., 2011. Still Different After All These Years: Solidarity Behavior in East and West Germany. *Journal of Public Economics*, 95(11), 1373-1376.

Browning, C.R., Cagney, K.A. and Wen, M., 2003. Explaining variation in health status across space and time: implications for racial and ethnic disparities in self-rated health. *Social Science & Medicine*, 57(7), 1221-1235.

Bryk, A.S. and Raudenbush, S.W., 1992. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Sage Publications. Newbury Park. London. New Delhi.

Burchardi, K.B. and Hassan, T.A., 2013. The economic impact of social ties: evidence from German reunification. *Quarterly Journal of Economics*, 128 (3), 1219-1271.

Burkhauser, R.V. and Cawley, J., 2008. Beyond BMI: The value of more accurate measures of fatness and obesity in social science research. *Journal of Health Economics*, 27(2), 519-529.

Cohen, J. P., Osleeb, J. P. and Yang, K. (2013). Semi-parametric regression models and Connecticut's Hospital cost. Paper presented at the annual meeting of the Spatial Econometrics Association 2013.

Coneus, K. and Spieß, K.C., 2012. Pollution exposure and child health: Evidence for infants and toddlers in Germany. *Journal of Health Economics*, 31(1), 180-196.

Congdon, P.D., 2010. *Applied Bayesian Hierarchical Methods*. Chapman & Hall/CRC. Boca Raton. London. New York.

Costa-Font, J. and Pons-Novell, J., 2007. Public health expenditures and spatial interactions in a decentralized national health system. *Health Economics*, 16(3), 291–306.

Cressie, N. A.C., 1993. *Statistics for Spatial Data*. Wiley. New York.

Cressie, N. A.C. and Wikle, C. K., 2011. *Statistics for Spatio-Temporal Data*. Wiley, Hoboken.

Dartmouth Atlas of Health Care, 2013. Available at www.dartmouthatlas.org, last accessed December 12, 2013.

Diehl, K. and Schneider, S., 2011. How relevant are district characteristics in explaining subjective health in Germany? – A multilevel analysis. *Social Science & Medicine*, 72(7), 1205-1210.

Di Matteo, L. and Di Matteo, R., 1998. Evidence on the determinants of Canadian provincial government health expenditures: 1965-1991. *Journal of Health Economics*, 17(2), 211-228.

Di Matteo, L., 2005. The macro determinants of health expenditure in the United States and Canada: assessing the impact of income, age distribution and time. *Health Policy*, 71(1), 23-42.

Dragano, N., Bobak, M., Wege, N., Peasey, A., Verde, P.E., Kubinova, R., Weyers, S., Moebus, S., Möhlenkamp, S., Stang, A., Erbel, R., Jöckel, K.-H., Siegrist, J. and Pikhart, H., 2007. Neighbourhood socioeconomic status and cardiovascular risk factors: a multilevel analysis of nine cities in the Czech Republic and Germany. *BMC Public Health*, 7:255.

Dragano, N., Hoffmann, B., Stang, A., Moebus, S., Verde, P. E., Weyers, S., Möhlenkamp, S., Schmermund, A., Mann, K., Jöckel, K.-H., Erbel, R., Siegrist, J., Heinz Nixdorf Recall Study Investigative Group, 2009. Subclinical coronary atherosclerosis and neighbourhood deprivation in an urban region. *European Journal of Epidemiology*, 24(1): 25-35.

Dragano, N., Hoffmann, B., Moebus, S., Möhlenkamp, S., Stang, A., Verde, P. E., Jöckel, K.-H., Erbel, R., Siegrist, J., Heinz Nixdorf Recall Study Investigative Group, 2009. Traffic exposure and subclinical cardiovascular disease: is the association modified by socioeconomic characteristics of individuals and neighbourhoods? Results from a multilevel study in an urban region. *Occupational and Environmental Medicine*, 66(9): 628-635.

Drewnowski, A., Rehm, C. D. and Solet, D. , 2007. Disparities in obesity rates: analysis by ZIP code area. *Social Science & Medicine*, 65(12), 2458-2463.

Durlauf, S. N., 2003. Neighbourhood effects, in: J. V. Henderson / J.-F. Thisse(eds.), *Handbook of Regional and Urban Economics*, edition 1, volume 4, chapter 50, 2173-2242.

Eibich, P. and Ziebarth, N.R., 2014. Analyzing regional variation in health care utilization using (rich) household microdata. *Health Policy*, 114(1), 41-53.

Federal Agency for Cartography and Geodesy, 2013. Maps of the Administrative Units of Germany, available at http://www.geodatenzentrum.de/geodaten/, last accessed December 12, 2013.

Federal Institute for Research on Building, Urban Affairs and Spatial Development, 2013. INKAR 2012. Indicators and Maps on Spatial Development, CD-ROM.

Felder, S. and Tauchmann, H., 2013. Federal state differentials in the efficiency of health production in Germany: an artifact of spatial dependence? *European Journal of Health Economics*, 14(1), 21-39.

Filippini, M., Masiero, G. and Moschetti, K., 2009. Regional consumption of antibiotics: A demand system approach. *Economic Modelling*, 26(6), 1389-1397.

Filippini, M., González Ortiz, L. and Masiero, G., 2013. Assessing the impact of national antibiotic campaigns in Europe. *European Journal of Health Economics*, 14(4), 587-599.

Frakes, M., 2013. The Impact of Medical Liability Standards on Regional Variations in Physician Behavior: Evidence from the Adoption of National-Standard Rules. *American Economic Review*, 103(1), 257-76.

Frijters, P., Haisken-DeNew, J. P. and Shields, M. A., 2004. Money Does Matter! Evidence from Increasing Real Income and Life Satisfaction in East Germany Following Reunification. *American Economic Review*, 94(3), 730-740.

Frijters, P., Haisken-DeNew, J. P. and Shields, M. A., 2005. The Causal Effect of Income on Health: Evidence from German Reunification. *Journal of Health Economics*, 24(5), 997-1017.

Fuchs-Schündeln, N., 2008. The Response of Household Saving to the Large Shock of German Reunification. *American Economic Review*, 98(5), 1798-1828.

Fuchs-Schündeln, N. and Schündeln, M., 2005. Precautionary Savings and Self-Selection: Evidence from the German Reunification Experiment. *Quarterly Journal of Economics*, 120(3), 1085-1120.

Gelman, A. and Robert, C.P., 2013. "Not only defended but also applied": The perceived absurdity of Bayesian inference. *The American Statistician*, 67(1), 1-5.

Gelman, A. and Rubin, D.B., 1992. Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, 7(4), 457-511.

González Ortiz, L. G. and Masiero, G., 2013. Disentangling spillover effects of antibiotic consumption: a spatial panel approach. *Applied Economics*, 45(8), 1041-1054.

Hajizadeh, M., Campbell, M. K., and Sarma, S., 2012. Does Spatial Heterogeneity Influence Adult Obesity in Canada? Spatial Panel Data Analysis, *mimeo*, available at http://www.chesg.ca/images/30000102/ConferencePapers/2012/Hajizadeh_Campbell_Sarma.pdf, last accessed September 9, 2013.

Hammersley, J. M. and Clifford, P., 1971. Markov field on finite graphs and lattices. Unpublished, available at http://www.statslab.cam.ac.uk/~grg/books/hammfest/hamm-cliff.pdf, last accessed on August 7, 2012.

Heineck, G. and Süssmuth, B., 2013. A Different Look at Lenin's Legacy: Social Capital and Risk Taking in the two Germanies. *Journal of Comparative Economics*, 41(3), 789-803.

Hox, J., 2002. *Multilevel Analysis. Techniques and Applications*. Lawrence Erlbaum Associates, Publishers. Mahwah (New Jersey), London.

Hunt, J. (2006): "Staunching Emigration from East Germany: Age and the Determinants of Migration," *Journal of the European Economic Associatio*n, 4(5): 1014-1037.

Jürges, H., 2007. True health vs. response styles: exploring cross-country differences in self-reported health. *Health Economics*, 16(2), 163-178.

Kassenboehmer, S. C. and Haisken-DeNew, J.P., 2009.You're Fired! The Causal Negative Effect of Entry Unemployment on Life Satisfaction. *Economic Journal*, 119(536), 448-462.

Kazembe, L. N. and Namangale, J. J. (2007). A Bayesian multinomial model to analyse spatial patterns of childhood co-morbidity in Malawi. *European Journal of Epidemiology*, 22(8): 545-556.

Knies, G. and Spieß, C.K., 2007. Regional data in the German Socio-Economic Panel Study (SOEP), Data Documentation 17, DIW Berlin, German Institute for Economic Research, available at http://ideas.repec.org/e/pkn24.html, last accessed May 2, 2012.

Knies, G., Burgess, S. and Propper, C., 2008. Keeping up with the Schmidt's – an empirical test of relative deprivation theory in the neighbourhood context. *Journal of Applied Social Science Studies (Schmollers Jahrbuch)*, 128(1), 75-108.

Kopetsch, T. and Schmitz, H., 2013. Regional variation in the utilisation of ambulatory services in Germany. *Health Economics*., forthcoming.

Lagravinese, R., Lee, H., Moscone, F. and Tosetti, E. (2013). The impact of air pollutions on hospital admissions: Evidence from Italy. Departmental Working Papers of Economics - University 'Roma Tre' 170.

Lahti-Koski, M., Taskinen, O., Similä, M., Männistö, S., Laatikainen, T., Knekt, P., and Valsta, L. M., 2008. Mapping geographical variation in obesity in Finland. *European Journal of Public Health*, 18(6), 637-643.

Latzitis, N., Sundmacher, L. and Busse, R., 2011. Regional differences in life expectancy in Germany at county levels and their possible determinants. *Gesundheitswesen*, 73(4), 217-228.

Lauridsen, J., Bech, M., Lopez, F. and Mate Sanchez, M., 2008. Geographic and temporal heterogeneity in public prescription pharmaceutical expenditures in Spain. *The Review of Regional Studies*, 38(1), 89-103.

Lauridsen, J., Bech, M., López, F., and Sánchez, M., 2010. A spatiotemporal analysis of public pharmaceutical expenditure. *The Annals of Regional Science*, 44(2), 299-314.

Lauridsen, J., Sánchez, M. and Bech, M., 2010. Public pharmaceutical expenditure: identification of spatial effects. *Journal of Geographical Systems*, 12(2), 175-188.

Meyer, C., Rumpf, H. J. , Hapke, U. and John, U., 1998. Regional differences in the prevalence of high-risk alcohol drinking: secondary analysis of the East-West health survey. *Gesundheitswesen*, 60(8-9), 486-492.

Michimi, A. and Wimberly, M. C., 2010. Spatial patterns of obesity and associated risk factors in the conterminous U.S. *American Journal of Preventive Medicine*, 39(2), e1-12.

Mollié, A., 1996. Bayesian mapping of disease. In A. Lawson, A. Biggeri, D. Bohning, E.Lesaffre, J.F. Viel, and R. Bertollini, (editors): *Markov Chain Monte Carlo in Practice: Interdisciplinary Statistics*. Chapman and Hall. London. 359-379.

Moran, P.A.P., 1950. Notes on continuous stochastic phenomena. *Biometrika*, 37(1), 17-23.

Moscone, F. ,Knapp, M. and Tosetti, E., 2007. Mental health expenditure in England: A spatial panel approach. *Journal of Health Economics*, 26(4), 842-864.

Moscone, F. and Tosetti, E., 2010. Health expenditure and income in the United States. *Health Economics*, 19(12), 385-1403.

Mueller, I., Vounatsou, P., Allen, B.J. and Smith, T., 2001. Spatial patterns of child growth in Papua New Guinea and their relation to environment, diet, socio-economic status and subsistence activities. *Annals of Human Biology*, 28(3), 263-280.

Müller-Nordhorn, J., Rossnagel, K., Mey, W. and Willich, S. N., 2004. Regional variation and time trends in mortality from ischaemic heart disease: East and West Germany 10 years after reunification. *Journal of Epidemiology & Community Health*, 58(6): 481-485.

Nolte, E., Kupilová, I. and McKee, M.,2011. Temporal and spatial pattern of infant mortality in Germany after unification. *Sozial- und Praventivmedizin*, 46(5), 303-310.

Nolte, E., Britton, A. and McKee, M.,2003. Trends in mortality attributable to current alcohol consumption in east and west Germany. *Social Science & Medicine*, 56(7), 1385-1395.

Ocaña-Riola, R. and Mayoral-Cortés, J.M., 2010. Spatio-temporal trends of mortality in small areas of Southern Spain. *BMC Public Health*, 10:26.

OECD Project on Income Distribution and Poverty, 2009. What are equivalence scales?. Availabe online at http://www.oecd.org/dataoecd/61/52/35411111.pdf, last accessed on April 30, 2012.

OECD, 2011. Waiting times, *in* OECD, *Health at a Glance 2011: OECD Indicators*, OECD Publishing.

Owens, G. M., 2008. Gender differences in health care expenditures, resource utilization, and quality of Care. *Journal of Managed Care Pharmacy*, 14(3), 2-6.

Pesaran, M.H., 2006. Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, 74(4), 967-1012.

Pollack, C., Knesebeck, O. v.d. and Siegrist, J., 2004. Housing and health in Germany, *Journal of Epidemiology & Community Health*, 58(3), 216-222.

Pouliou T. and Elliott S. J., 2009. An exploratory spatial analysis of overweight and obesity in Canada. *Preventive Medicine*, 48(4), 362-367.

RAND, 1995. RAND. User's Manual for the Medical Outcomes Study (MOS) Core Measures of Health-Related Quality of Life. (by Hays, R. D., Sherbourne, C. D, Mazel, R.). RAND Monograph Report, Santa Monica, Available at: http://www.rand.org/pubs/monograph_reports/MR162.html, last accessed December 12, 2013.

Rainer, H. and Siedler, T., 2009. Does Democracy Foster Trust? *Journal of Comparative Economics*, 37(2), 251-269.

Riphahn, R. T. and Zimmermann, K. F., 1998. The mortality crisis in East Germany. IZA Discussion Papers 6.

Riva, M., Gauvin, L. and Barnett, T.A., 2007. Toward the next generation of research into small area effects on health: a synthesis of multilevel investigations published since July 1998. *Journal of Epidemiology & Community Health*, 61(10), 853-861.

Schipf, S., Werner, A., Tamayo, T., Holle, R., Schunk, M., Maier, W., Meisinger, C., Thorand, B., Berger, K., Mueller, G., Moebus, S., Bokhof, B., Kluttig, A., Greiser, K.-H., Neuhauser, H., Ellert, U., Icks, A., Rathmann, W. and Völzke, H., 2012. Regional differences in the prevalence of known Type 2 diabetes mellitus in 45–74 years old individuals: Results from six population-based studies in Germany (DIAB-CORE Consortium). *Diabetic Medicine*, 29(7): e88-e95.

Spiegelhalter, D.J., Best, N., Carlin, B.P. and van der Linde, A., 2002. Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society*, Series B, 64(4), 583-639.

Sullivan, D. and Wachter, T.v., 2009. Job Displacement and Mortality: An Analysis Using Administrative Data. *Quarterly Journal of Economics*, 124(3): 1265-1306.

Sundmacher, L. and Busse, R., 2011. The impact of physician supply on avoidable cancer deaths in Germany. A spatial analysis. *Health Policy*, 103(1), 53-62.

Sundmacher, L., Gaskins, M., Hoffman, K, and Busse R., 2012. Spatial distribution of avoidable cancer mortality. *Journal of Public Health*, 20(3), 279-288.

Treurniet, H.F., Boshuizen, H.C. and Harteloh, P.P.M., 2004. Avoidable mortality in Europe (1980-1997): a comparison of trends. *Journal of Epidemiology & Community Health*, 58(4), 290-295.

Voigtländer, S., Berger, U. and Razum, O., 2010. The impact of regional and neighbourhood deprivation on physical health in Germany: a multilevel study. *BMC Public Health*, 10:403.

Wennberg, J. and Gittelsohn, A., 1973. Small area variations in health care delivery. *Science*, 182 (4117), 1102-1108.

Wennberg, J.,Fisher, E. and Skinner, J.S., 2002. Geography and the debate over Medicare reform. *Health Affair* 21 (2, Suppl. Web Exclusive): W96–114.

Wagner, G. G., Frick, J. R., and Schupp, J., 2007. The German Socio-Economic Panel study (SOEP) - evolution, scope and enhancements. *Journal of Applied Social Science Studies (Schmollers Jahrbuch)*, 127(1), 139-169.

Ziebarth, N.R.and M.M Grabka, 2009. In Vino Pecunia? The Association Between Beverage-Specific Drinking Behavior and Wages. *Journal of Labor Research*, 30(3), 219-244.

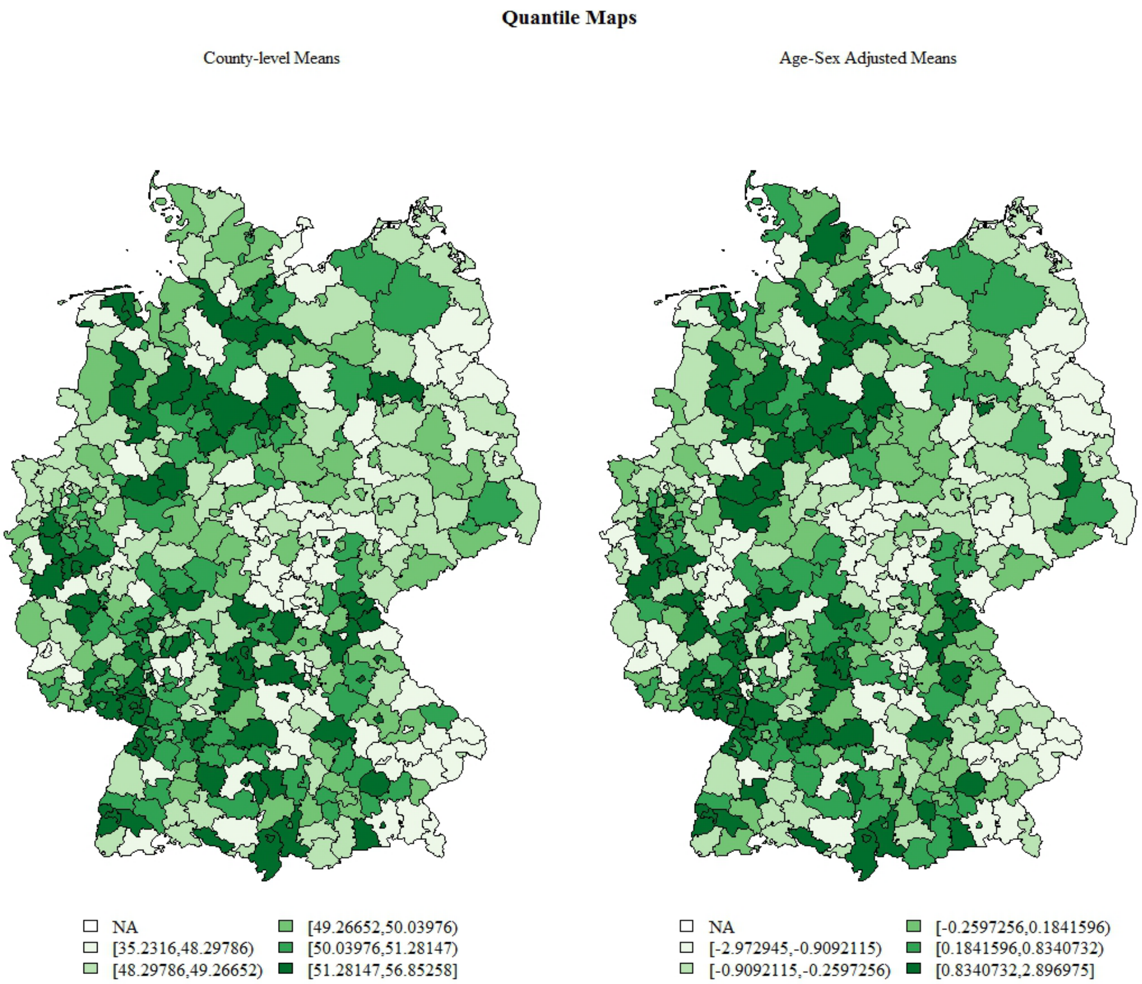
Ziebarth, N.R., 2010. Measurement of Health, Health Inequality, and Reporting Heterogeneity. *Social Science & Medicine*. 71(1), 116–124.

Ziebarth, N.R. and Wagner, G. G., 2013. How attitudes toward risk behavior are shaped: evidence from performance-enhancement drugs and divided Germany, www.human.cornell.edu/pam/people/upload/2013-7-10-Attitudes-risky-behavior.pdf, last accessed on December 13, 2013.

Zhu, L., Gorman, D. N. and Horel, S. (2006). Hierarchical Bayesian spatial models for alcohol availability, drug "hot spots" and violent crime. *International Journal of Health Geographics*, 5:54.

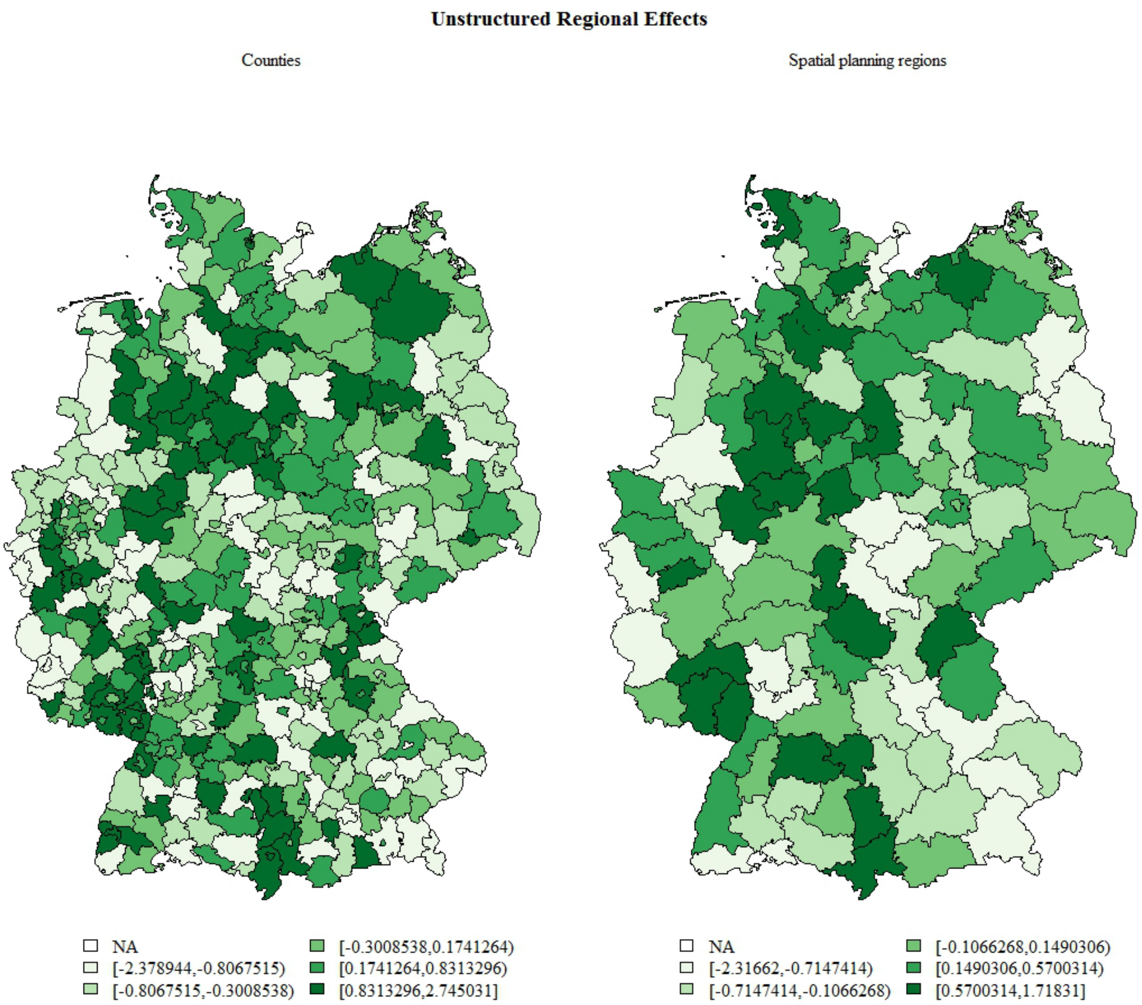
## Tables and Figures

### Figure 1




Source: SOEP v28, own calculations. Means of SF12 by county are displayed. The map on the left-hand side shows the overall means, whereas the map on the right-hand side shows age-sex adjusted means. The county borders reflect the territorial statuses as of January 1, 2012. The values of the variables are divided into five classes; the quintiles of the SF12 distribution serve as cutoff points. Each county is colored in a shade according to the class of the respective value of the variable. Lighter shades stand for lower values and darker shades for higher values. Areas without observations (1) are depicted in white.

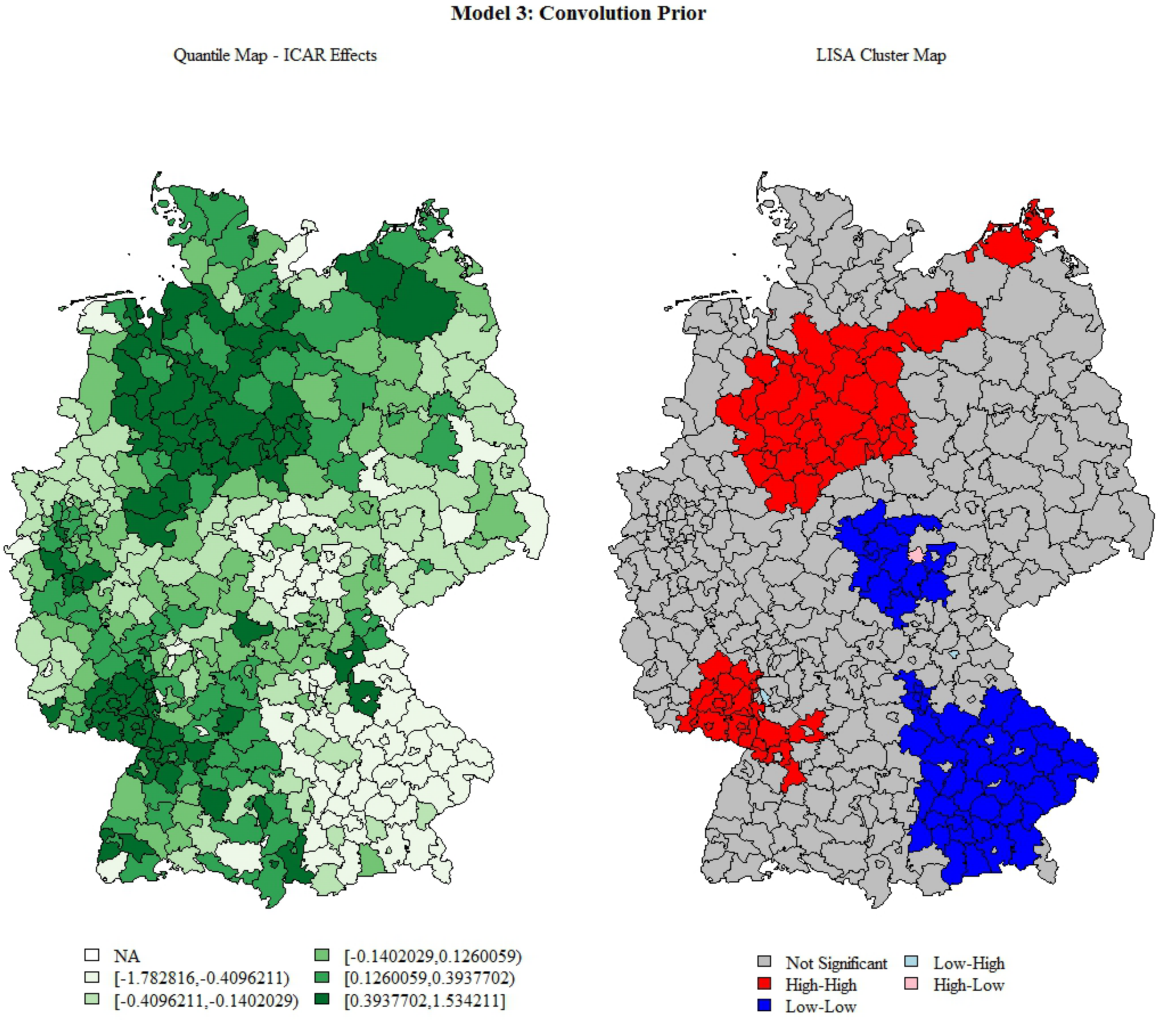
**Figure 2**




S

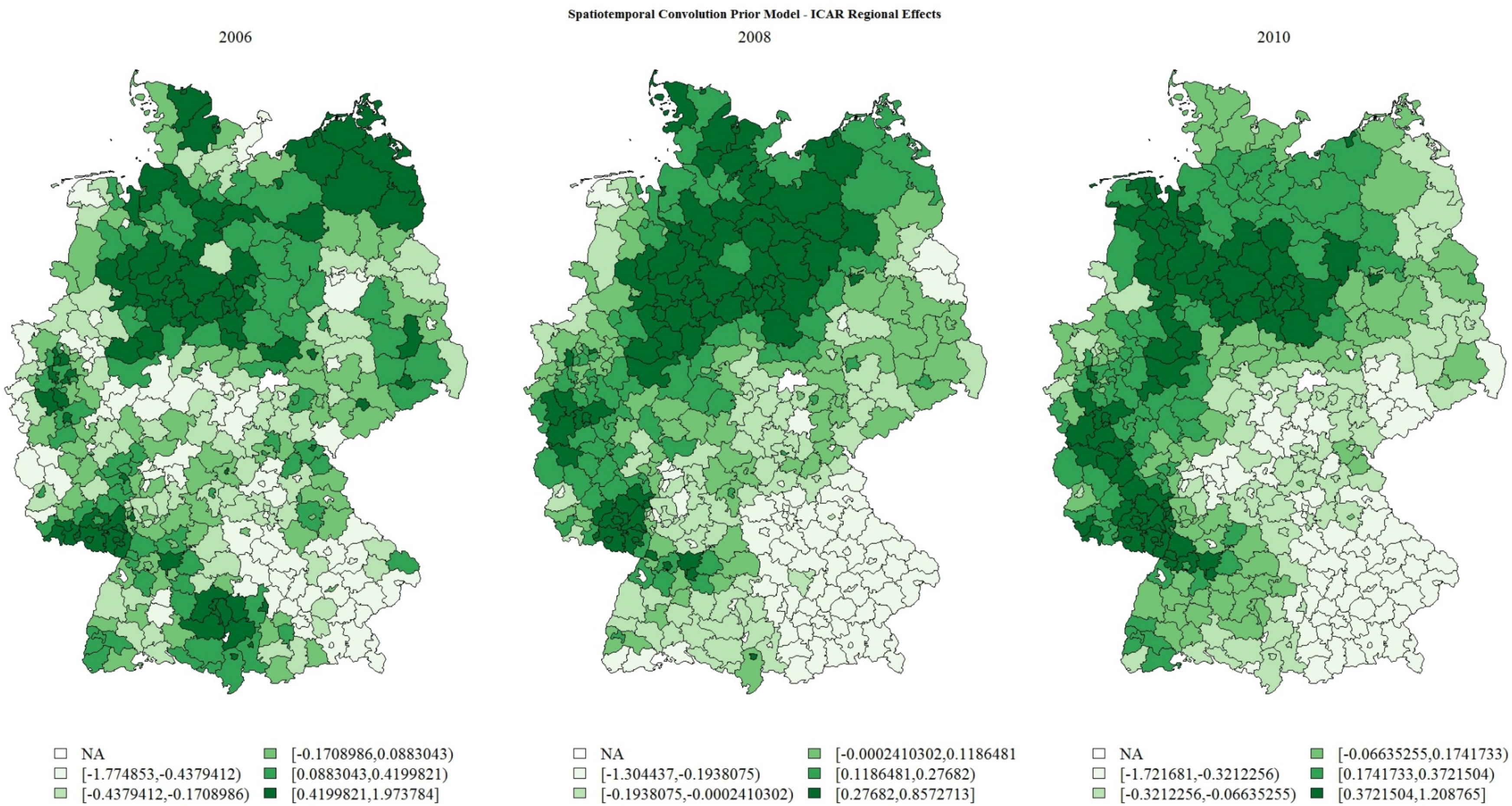
Source: SOEP v28, own calculations. Estimated county-level effects according to Model 1 (URE-Model) are displayed. The map on the left-hand side shows the estimated county-level effects. The county borders reflect the territorial statuses as of January 1, 2012. The map on the right-hand side shows the results for the model using the spatial planning regions as areal units. One county without observations is depicted in white.

**Figure 3**




Source: SOEP v28, own calculations. This graph represents the spatial correlation according to Model 3 (CP-Model). The map on the left shows the quantile map for the spatially dependent part $\varphi_s$. The map on the right shows the results for the local Moran's I, estimated using a first-order adjacency matrix. The areas are color-coded with respect to the sign of their county-level effect and the sign and significance of the spatial autocorrelation. One county without observations is depicted in white on the left and in grey on the right map.

**Figure 4**




Source: SOEP v28, own calculations.This graph represents the spatial correlation according to Model 3 (SCP-Model) for each year separately. Displayed is only the spatially dependent part $\varphi_{st}$ . The county borders reflect the territorial statuses as of January 1, 2012. Five counties without observations are depicted in white.

**Table 1:** Summary Statistics

| Variable | Mean | Standard Deviation | Min | Max | N |
|---|---|---|---|---|---|
| **Dependent Variable** | | | | | |
| SF12 - generic health measure | 49.73 | 7.12 | 22 | 62 | 54,734 |
| **A: Individual-Level Characteristics** | | | | | |
| ***Demographics*** | | | | | |
| Age | 49.31 | 17.58 | 17 | 100 | 54,734 |
| Number of children under 14 | 0.36 | 0.75 | 0 | 8 | 54,734 |
| Male gender | 0.48 | 0.50 | 0 | 1 | 54,734 |
| Married and cohabiting | 0.60 | 0.49 | 0 | 1 | 54,734 |
| ***Education & Labor Market Participation*** | | | | | |
| Monthly household income - equivalence scale | 1,776.74 | 1,263.56 | 0 | 66,666 | 54,734 |
| Not working | 0.43 | 0.50 | 0 | 1 | 54,734 |
| University degree | 0.21 | 0.41 | 0 | 1 | 54,734 |
| No vocational training | 0.23 | 0.42 | 0 | 1 | 54,734 |
| ***Health-related Behavior*** | | | | | |
| No alcohol consumption | 0.13 | 0.34 | 0 | 1 | 54,734 |
| Regular alcohol consumption | 0.17 | 0.38 | 0 | 1 | 54,734 |
| Smoking | 0.28 | 0.45 | 0 | 1 | 54,734 |
| Body Mass Index (BMI) | 25.96 | 4.60 | 13 | 73 | 54,734 |
| ***Health Care Utilization*** | | | | | |
| Number of stays in hospital during the past year | 0.16 | 0.62 | 0 | 42 | 54,734 |
| Number of doctors visits during the past year | 9.80 | 15.18 | 0 | 396 | 54,734 |
| **B: County-level Characteristics** | | | | | |
| ***Area and Population*** | | | | | |
| Area size | 891.75 | 721.61 | 36 | 5,812 | 1,200 |
| Population density | 521.48 | 676.08 | 39 | 4,355 | 1,200 |
| Proportion of area used for settlement and transport | 20.72 | 15.47 | 2 | 77 | 1,200 |
| Proportion of area suitable for recreation | 1.96 | 2.37 | 0 | 15 | 1,200 |
| Share of the population aged 65 and above | 20.65 | 2.23 | 14 | 28 | 1,200 |
| Percentage of residents with foreign nationality | 7.25 | 4.50 | 1 | 26 | 1,200 |
| Number of overnight stays per inhabitant | 4.74 | 5.69 | 0 | 43 | 1,200 |
| ***Labor Market and Standard of Living*** | | | | | |
| Unemployment quota | 8.34 | 4.12 | 2 | 24 | 1,200 |
| Share of mini jobs | 21.33 | 4.88 | 8 | 36 | 1,200 |
| GDP per capita | 27.71 | 10.55 | 13 | 85 | 1,200 |
| Average available income per month | 1,534.28 | 205.58 | 1,109 | 2,702 | 1,200 |
| Average price for construction grounds in €/m² | 129.41 | 116.62 | 6 | 1,067 | 1,200 |
| Cars per 1000 inhabitants | 547.04 | 66.17 | 318 | 1,073 | 1,200 |
| ***Health Care Supply*** | | | | | |
| Physicians per 10,000 inhabitants | 160.62 | 55.80 | 82 | 393 | 1,200 |
| GPs per 10,000 inhabitants | 51.96 | 8.41 | 21 | 80 | 1,200 |
| Number of hospital beds per 10,000 inhabitants | 63.93 | 37.46 | 0 | 227 | 1,200 |
| Nursing home places per 10,000 inhabitants | 102.28 | 28.71 | 45 | 257 | 1,200 |

*Source: SOEP v28, INCAR 2012*

**Table 2:** Coefficient Estimates for Model 1: Unstructured Regional Effects

| Variable | Mean | Rescaled | Significance | S.E | 2.5% Percentile | 97.5% Percentile |
|---|---|---|---|---|---|---|
| ***Individual-Level Variables*** | | | | | | |
| Age | -16.242 | -0.924 | *** | 0.999 | -18.225 | -14.290 |
| Age squared | 15.317 | 0.009 | *** | 1.071 | 13.230 | 17.440 |
| Age cubed | -2.645 | 0.000 | *** | 0.173 | -2.987 | -2.315 |
| Number of children | 0.174 | 0.233 | *** | 0.034 | 0.106 | 0.241 |
| Male gender | 1.083 | - | *** | 0.073 | 0.937 | 1.228 |
| Married | 0.416 | - | *** | 0.076 | 0.265 | 0.568 |
| Household income | 0.380 | 0.000 | *** | 0.030 | 0.321 | 0.437 |
| Not working | -0.465 | - | *** | 0.068 | -0.600 | -0.334 |
| University Degree | 0.939 | - | *** | 0.091 | 0.754 | 1.117 |
| No vocational training | -0.717 | - | *** | 0.087 | -0.886 | -0.548 |
| No alcohol consumption | -1.125 | - | *** | 0.079 | -1.283 | -0.969 |
| Regular alcohol consumption | 0.255 | - | *** | 0.070 | 0.119 | 0.389 |
| Smoking | -0.881 | - | *** | 0.069 | -1.015 | -0.746 |
| BMI | -0.585 | -0.127 | *** | 0.034 | -0.651 | -0.520 |
| Number of hospital stays | -0.392 | -0.628 | *** | 0.022 | -0.435 | -0.348 |
| Number of doctors visits | -1.549 | -0.102 | *** | 0.024 | -1.596 | -1.501 |
| ***County-Level Variables*** | | | | | | |
| Area size | 0.055 | 0.000 | | 0.105 | -0.148 | 0.262 |
| Population density | 0.087 | 0.000 | | 0.359 | -0.602 | 0.789 |
| Area used for settlement and transport | 0.007 | 0.000 | | 0.331 | -0.656 | 0.641 |
| Area used for recreation | 0.003 | 0.001 | | 0.203 | -0.393 | 0.400 |
| Share of elderly | 0.241 | 0.110 | *** | 0.089 | 0.064 | 0.414 |
| Percentage of migrants | -0.017 | -0.003 | | 0.168 | -0.345 | 0.311 |
| Prices for construction grounds | 0.155 | 0.001 | * | 0.110 | -0.065 | 0.369 |
| Cars per 1,000 Inhabitants | 0.101 | 0.001 | | 0.105 | -0.103 | 0.306 |
| Overnight stays per inhabitant | -0.004 | -0.001 | | 0.070 | -0.138 | 0.138 |
| Unemployment quota | -0.115 | -0.026 | * | 0.086 | -0.284 | 0.050 |
| Share of minijobs | 0.416 | 0.084 | *** | 0.099 | 0.226 | 0.616 |
| GDP per capita | 0.238 | 0.021 | ** | 0.128 | -0.016 | 0.493 |
| Average household income | -0.173 | -0.001 | * | 0.107 | -0.380 | 0.039 |
| Physicians per 10,000 population | 0.254 | 0.005 | ** | 0.139 | -0.019 | 0.520 |
| GPs per 10,000 population | -0.117 | -0.015 | * | 0.075 | -0.260 | 0.030 |
| Hospital beds per 10,000 population | -0.110 | -0.004 | | 0.089 | -0.283 | 0.065 |
| Nursing home places per 10,000 population | -0.071 | -0.003 | | 0.066 | -0.198 | 0.058 |
| Constant | 51.612 | - | | 0.218 | 51.195 | 52.040 |
| $\sigma_C$ | 4.622 | - | | 0.030 | 4.563 | 4.681 |
| $\sigma_B$ | 1.133 | - | | 0.063 | 1.015 | 1.259 |
| $\sigma$ | 3.873 | - | | 0.016 | 3.843 | 3.903 |

Source: SOEP v28, INKAR 2012, own calculations. N=54,734. 4,500 draws from the posterior distribution. Column 1 gives the posterior mean (i.e., the point estimate) for the 15 individual predictors, the 17 county-level covariates, the aggregate time effects, the constant and the standard deviations of the dependent variable, ($\sigma$), the individual effects, ($\sigma_C$) , and the regional effects, ($\sigma_B$). Note that the continuous variables are standardized and therefore the measurement units are standard deviations. Column 2 gives a rescaled coefficient estimate, i.e., the effect size corresponding to the original measurement unit (left blank for dummy variables). Column 3 gives the posterior probability that the parameter has a different sign than the point estimate: ***=<0.1%, **=<1%, *<5% and .=<10%. Column 4 gives the standard errors. Columns 5 and 6 give the 95%-Equal Tail Credible Interval around the median.

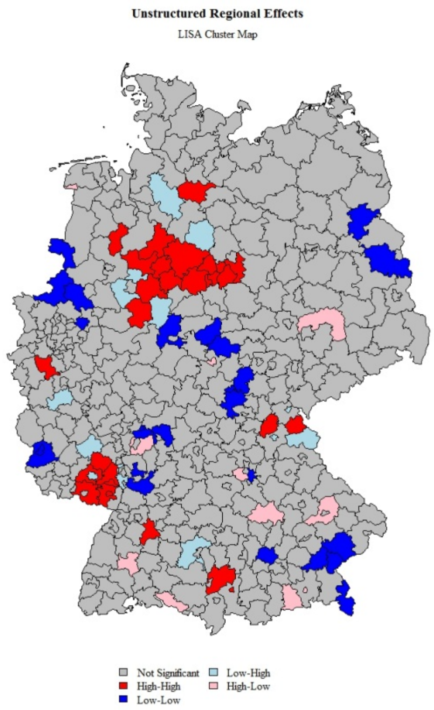
**Table 3: Coefficient estimates for Model 3: Spatiotemporal Convolution Prior**

| Variable | Mean | Rescaled | Significance | S.E | 2.5% Percentile | 97.5% Percentile |
|---|---|---|---|---|---|---|
| ***Individual-Level Variables*** | | | | | | |
| Age | -12.305 | -0.700 | *** | 0.739 | -13.720 | -10.865 |
| Age squared | 22.965 | 0.013 | *** | 1.542 | 19.955 | 25.900 |
| Age cubed | -12.922 | 0.000 | *** | 0.834 | -14.535 | -11.280 |
| Number of children | 0.185 | 0.247 | *** | 0.035 | 0.117 | 0.253 |
| Male gender | 1.078 | - | *** | 0.074 | 0.936 | 1.225 |
| Married | 0.408 | - | *** | 0.076 | 0.260 | 0.559 |
| Household income | 0.384 | 0.000 | *** | 0.031 | 0.325 | 0.444 |
| Not working | -0.465 | - | *** | 0.067 | -0.596 | -0.335 |
| University Degree | 0.957 | - | *** | 0.093 | 0.774 | 1.141 |
| No vocational training | -0.709 | - | *** | 0.087 | -0.877 | -0.536 |
| No alcohol consumption | -1.112 | - | *** | 0.080 | -1.263 | -0.960 |
| Regular alcohol consumption | 0.280 | - | *** | 0.071 | 0.139 | 0.414 |
| Smoking | -0.879 | - | *** | 0.069 | -1.014 | -0.745 |
| BMI | -0.592 | -0.129 | *** | 0.035 | -0.660 | -0.525 |
| Number of hospital stays | -0.392 | -0.628 | *** | 0.022 | -0.435 | -0.349 |
| Number of doctors visits | -1.538 | -0.101 | *** | 0.024 | -1.585 | -1.489 |
| ***County-Level Variables*** | **Mean** | **Rescaled** | **Significance** | **S.E** | **2.5% Percentile** | **97.5% Percentile** |
| Area size | 0.066 | 0.000 | | 0.072 | -0.077 | 0.213 |
| Population density | -0.271 | 0.000 | | 0.263 | -0.771 | 0.240 |
| Area used for settlement and transport | 0.423 | 0.021 | * | 0.271 | -0.104 | 0.950 |
| Area used for recreation | -0.036 | -0.011 | | 0.138 | -0.311 | 0.227 |
| Share of elderly | 0.227 | 0.104 | *** | 0.077 | 0.074 | 0.378 |
| Percentage of migrants | -0.144 | -0.027 | | 0.126 | -0.394 | 0.100 |
| Prices for construction grounds | 0.280 | 0.002 | *** | 0.114 | 0.057 | 0.511 |
| Cars per 1,000 Inhabitants | -0.022 | 0.000 | | 0.103 | -0.222 | 0.182 |
| Overnight stays per inhabitant | 0.010 | 0.002 | | 0.049 | -0.086 | 0.109 |
| Unemployment quota | -0.265 | -0.061 | ** | 0.115 | -0.490 | -0.041 |
| Share of minijobs | 0.271 | 0.055 | *** | 0.086 | 0.105 | 0.439 |
| GDP per capita | 0.252 | 0.022 | *** | 0.098 | 0.065 | 0.445 |
| Average household income | -0.044 | 0.000 | | 0.087 | -0.216 | 0.126 |
| Physicians per 10,000 population | 0.128 | 0.002 | | 0.108 | -0.083 | 0.340 |
| GPs per 10,000 population | -0.001 | 0.000 | | 0.065 | -0.130 | 0.125 |
| Hospital beds per 10,000 population | -0.160 | -0.005 | *** | 0.069 | -0.296 | -0.029 |
| Nursing home places per 10,000 population | -0.041 | -0.002 | | 0.055 | -0.151 | 0.067 |
| Constant | 48.923 | - | *** | 0.109 | 48.700 | 49.140 |
| $\sigma_C$ | 4.674 | - | *** | 0.030 | 4.616 | 4.735 |
| $\sigma_B / \sigma_\omega / \sigma_\varphi$ | | | *Results for each year available upon request* | | | |
| $\sigma$ | 3.848 | - | *** | 0.016 | 3.818 | 3.878 |

Source: SOEP v28, INKAR 2012, own calculations. Number of observations N=54,734. Number of draws from the Posterior Distribution D=4,500. Column 1 gives the posterior mean (i.e., the point estimate) for the 15 individual predictors, the 17 county-level covariates, the aggregate time effects, the constant and the standard deviations of the dependent variable, (σ), the individual effects, (σC) , and the regional effects, (σB). Note that the continuous variables are standardized and therefore the measurement units are standard deviations. Column 2 gives a rescaled coefficient estimate, i.e., the effect size corresponding to the original measurement unit (left blank for dummy variables). Column 3 gives the posterior probability that the parameter has a different sign than the point estimate: ***=<0.1%, **=<1%, *<5% and .=<10%. Columns 5 and 6 give the 95%-Equal Tail Credibility Interval around the median.

# Online Appendix

## Figure A.1




Source: SOEP v28, own calculations. This graph represents the spatial correlation according to Model 1 (URE-Model), estimated by local Moran's I with a first-order adjacency matrix. The areas are color-coded with respect to the sign of their county-level effect and the sign and significance of the spatial autocorrelation. One county without observations is depicted in grey.

**Table A.1: Coefficient estimates for the federal states and spatial planning regions for Model 1 (URE)**

| | *Federal states* | | | | | | *Spatial planning regions* | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Variable** | **Mean** | **Rescaled** | **Significance** | **S.E** | **2.5% Percentile** | **97.5% Percentile** | **Mean** | **Rescaled** | **Significance** | **S.E** | **2.5% Percentile** | **97.5% Percentile** |
| ***Individual-Level Variables*** | | | | | | | | | | | | |
| Age | -12.411 | -0.706 | *** | 0.734 | -13.845 | -10.910 | -12.326 | -0.701 | *** | 0.736 | -13.750 | -10.895 |
| Age squared | 23.232 | 0.013 | *** | 1.537 | 20.160 | 26.205 | 23.062 | 0.013 | *** | 1.526 | 20.160 | 26.050 |
| Age cubed | -13.076 | 0.000 | *** | 0.833 | -14.710 | -11.410 | -12.995 | 0.000 | *** | 0.822 | -14.590 | -11.405 |
| Number of children | 0.183 | 0.245 | *** | 0.035 | 0.113 | 0.251 | 0.175 | 0.234 | *** | 0.034 | 0.109 | 0.244 |
| Male gender | 1.083 | | *** | 0.074 | 0.937 | 1.226 | 1.081 | | *** | 0.073 | 0.935 | 1.224 |
| Married | 0.392 | | *** | 0.076 | 0.241 | 0.540 | 0.396 | | *** | 0.075 | 0.248 | 0.543 |
| Household income | 0.390 | 0.000 | *** | 0.030 | 0.333 | 0.448 | 0.384 | 0.000 | *** | 0.030 | 0.326 | 0.442 |
| Not working | -0.452 | | *** | 0.068 | -0.585 | -0.320 | -0.452 | | *** | 0.068 | -0.585 | -0.320 |
| University Degree | 1.027 | | *** | 0.090 | 0.852 | 1.199 | 0.993 | | *** | 0.091 | 0.814 | 1.178 |
| No vocational training | -0.730 | | *** | 0.087 | -0.901 | -0.557 | -0.724 | | *** | 0.086 | -0.891 | -0.562 |
| No alcohol consumption | -1.107 | | *** | 0.079 | -1.263 | -0.953 | -1.112 | | *** | 0.080 | -1.272 | -0.954 |
| Regular alcohol consumption | 0.268 | | *** | 0.071 | 0.129 | 0.408 | 0.273 | | *** | 0.070 | 0.135 | 0.410 |
| Smoking | -0.888 | | *** | 0.070 | -1.025 | -0.753 | -0.878 | | *** | 0.070 | -1.011 | -0.741 |
| BMI | -0.606 | -0.132 | *** | 0.034 | -0.673 | -0.539 | -0.599 | -0.130 | *** | 0.034 | -0.666 | -0.533 |
| Number of hospital stays | -0.395 | -0.633 | *** | 0.022 | -0.438 | -0.352 | -0.393 | -0.630 | *** | 0.022 | -0.436 | -0.350 |
| Number of doctors visits | -1.540 | -0.101 | *** | 0.024 | -1.589 | -1.492 | -1.542 | -0.102 | *** | 0.024 | -1.590 | -1.495 |
| ***County-Level Variables*** | **Mean** | **Rescaled** | **Significance** | **S.E** | **2.5% Percentile** | **97.5% Percentile** | **Mean** | **Rescaled** | **Significance** | **S.E** | **2.5% Percentile** | **97.5% Percentile** |
| Area size | -0.060 | 0.000 | | 0.386 | -0.874 | 0.639 | -0.284 | 0.000 | ** | 0.148 | -0.573 | 0.006 |
| Population density | 0.528 | 0.001 | | 0.940 | -1.196 | 2.601 | 0.034 | 0.000 | | 0.268 | -0.501 | 0.563 |
| Area used for settlement and transport | 0.631 | 0.049 | | 1.228 | -1.823 | 3.080 | -0.387 | -0.028 | ** | 0.190 | -0.762 | -0.020 |
| Area used for recreation | -0.208 | -0.093 | | 0.520 | -1.261 | 0.762 | 0.179 | 0.078 | | 0.186 | -0.188 | 0.538 |
| Share of elderly | -0.386 | -0.249 | | 0.314 | -1.003 | 0.240 | -0.131 | -0.071 | | 0.141 | -0.411 | 0.143 |

*…continued on next page*

*…Table A.1 continued*

| ***County-Level Variables*** | **Mean** | **Rescaled** | **Significance** | **S.E** | **2.5% Percentile** | **97.5% Percentile** | **Mean** | **Rescaled** | **Significance** | **S.E** | **2.5% Percentile** | **97.5% Percentile** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Percentage of migrants | -0.949 | -0.257 | *** | 0.571 | -2.155 | 0.106 | -0.281 | -0.062 | | 0.248 | -0.771 | 0.207 |
| Prices for construction grounds | 0.043 | 0.001 | | 0.129 | -0.211 | 0.292 | 0.104 | 0.001 | * | 0.070 | -0.029 | 0.241 |
| Cars per 1,000 Inhabitants | 1.232 | 0.023 | *** | 0.454 | 0.402 | 2.155 | 0.118 | 0.002 | * | 0.090 | -0.062 | 0.290 |
| Overnight stays per inhabitant | -0.045 | -0.021 | | 0.159 | -0.383 | 0.239 | 0.047 | 0.013 | | 0.084 | -0.123 | 0.212 |
| Unemployment quota | 0.056 | 0.015 | | 0.248 | -0.455 | 0.526 | -0.104 | -0.026 | | 0.107 | -0.309 | 0.108 |
| Share of minijobs | 0.183 | 0.053 | | 0.319 | -0.451 | 0.811 | 0.185 | 0.047 | * | 0.146 | -0.094 | 0.490 |
| GDP per capita | 0.183 | 0.033 | | 0.443 | -0.708 | 1.050 | 0.118 | 0.017 | | 0.203 | -0.280 | 0.515 |
| Average household income | 0.148 | 0.001 | | 0.422 | -0.688 | 0.976 | -0.015 | 0.000 | | 0.192 | -0.391 | 0.360 |
| Physicians per 10,000 population | 0.322 | 0.015 | | 0.447 | -0.540 | 1.187 | 0.297 | 0.010 | ** | 0.174 | -0.036 | 0.638 |
| GPs per 10,000 population | -0.465 | -0.091 | * | 0.316 | -1.088 | 0.163 | -0.306 | -0.047 | *** | 0.120 | -0.539 | -0.074 |
| Hospital beds per 10,000 population | 0.011 | 0.002 | | 0.094 | -0.174 | 0.192 | 0.028 | 0.003 | | 0.061 | -0.090 | 0.150 |
| Nursing home places per 10,000 population | 0.173 | 0.015 | | 0.182 | -0.184 | 0.535 | 0.046 | 0.003 | | 0.100 | -0.152 | 0.238 |
| ***Time effects*** | | | | | | | | | | | | |
| Dummy 2006 | -0.433 | | | 0.438 | -1.338 | 0.369 | 0.526 | | *** | 0.149 | 0.239 | 0.819 |
| Dummy 2008 | 0.752 | | *** | 0.173 | 0.417 | 1.086 | 0.506 | | *** | 0.070 | 0.371 | 0.641 |
| Constant | 49.125 | | | 0.378 | 48.160 | 49.750 | 48.973 | | *** | 0.152 | 48.670 | 49.270 |
| $\sigma_C$ | 4.728 | | *** | 0.030 | 4.668 | 4.785 | 4.684 | | *** | 0.030 | 4.624 | 4.742 |
| $\sigma_B$ | 0.953 | | | 0.404 | 0.429 | 1.951 | 0.902 | | *** | 0.090 | 0.743 | 1.094 |
| $\sigma$ | 3.876 | | | 0.016 | 3.845 | 3.907 | 3.874 | | *** | 0.016 | 3.843 | 3.905 |

Source: SOEP v28, INKAR 2012, own calculations. Number of observations N=54,734. Number of draws from the Posterior Distribution D=4,500. The left panel gives the results for Model 1 estimated on the level of the federal states. The right panel gives the coefficients estimated on the level of the spatial planning regions. Column 1 gives the posterior mean (i.e., the point estimate) for the 15 individual predictors, the 17 county-level covariates, the aggregate time effects, the constant and the standard deviations of the dependent variable, (σ), the individual effects, (σC) , and the regional effects, (σB). Note that the continuous variables are standardized and therefore the measurement units are standard deviations. Column 2 gives a rescaled coefficient estimate, i.e., the effect size corresponding to the original measurement unit (left blank for dummy variables). Column 3 gives the posterior probability that the parameter has a different sign than the point estimate: ***=<0.1%, **=<1%, *<5% and .=<10%. Columns 5 and 6 give the 95%-Equal Tail Credible Interval around the median.

**Table A.2: Moran's I for Different Definitions of Neighborhood**

| **Weights** | **Links** | **I** | **p-value** |
|---|---|---|---|
| first-order adjacency | 2084 | 0.18 | 0.00 |
| second-order adjacency | 4880 | 0.03 | 0.08 |
| third-order adjacency | 7580 | 0.00 | 0.51 |
| threshold distance - 50 km | 3694 | 0.13 | 0.00 |
| threshold distance - 68.2 km | 27092 | 0.09 | 0.00 |
| threshold distance - 150 km | 8410 | 0.03 | 0.00 |
| 3 nearest neighbors | 1201 | 0.19 | 0.00 |
| 5 nearest neighbors | 2000 | 0.16 | 0.00 |
| 7 nearest neighbors | 2798 | 0.15 | 0.00 |

Source: SOEP v28, INKAR 2012, own calculations; Moran's I for nine different weight matrices are displayed. Column 1 gives the number of non-zero links, i.e., spatial connections between sites. Column 2 gives the estimate of Moran's I and Column 3 the corresponding empirical *p*-value, which is derived by a Monte-Carlo permutation test with 10,000 permutations.

**Table A.3: Model Selection Criteria for all Candidate Models**

| Model | Dbar | pD | DIC | Deviance | BIC | $R^2$ |
|---|---|---|---|---|---|---|
| Model 1: Unstructured Regional Effects | 303,537 | 17,563 | 321,101 | 303,538 | 416,550 | 0.456 |
| Model 2: Convolution Prior Model | 304,472 | 18,469 | 322,941 | 304,475 | 419,388 | 0.456 |
| Model 3: Spatiotemporal Convolution Prior Model | 303,129 | 18,419 | 321,548 | 303,128 | 425,536 | 0.460 |
| ***Robustness*** | | | | | | |
| Age-sex adjusted regional means | 305,876 | 18,460 | 324,336 | 305,878 | 418,753 | 0.444 |
| Model 1 with quintic age polynomial | 303,538 | 17,557 | 321,095 | 303,537 | 416,559 | 0.456 |
| Model 1 without regional covariates | 303,543 | 17,561 | 321,104 | 303,547 | 416,479 | 0.456 |
| Model 1 without regional effects | 303,600 | 17,733 | 321,333 | 303,589 | 414,620 | 0.456 |
| Model 1 without endogenous covariates | 305,876 | 18,460 | 324,336 | 306,122 | 419,101 | 0.443 |
| Federal states - Model 1 | 303,624 | 17,675 | 321,299 | 303,628 | 414,816 | 0.456 |
| Spatial planning regions - Model 1 | 303,577 | 17,617 | 321,193 | 303,573 | 415,140 | 0.456 |

Source: SOEP v28, INKAR 2012, own calculations. Column 1 gives the measure of model fit used for the calculation of the DIC, i.e., a function of the deviance and the data. Column 2 gives the measure of model complexity used for the DIC, i.e., the number of effective parameters. Column 3 states the Deviance Information Criterion (DIC). Column 4 gives the estimated deviance for each model and Column 5 states the Bayesian Information Criterion (BIC). The last column gives the unadjusted $R^2$.