EUROPEAN SPATIAL RESEARCH AND POLICY

Volume 26 2019 Number 1

http://dx.doi.org/10.18778/1231-1952.26.1.10

Agnieszka ORWAT-ACEDAŃSKA*

DYNAMIC SPATIAL PANEL DATA MODELS IN IDENTIFYING SOCIO-ECONOMIC FACTORS AFFECTING THE LEVEL OF HEALTH IN SELECTED EUROPEAN COUNTRIES

Abstract. The aim of the paper is to investigate the relationship between socio-economic factors and the level of health of citizens of selected European countries. Disability-adjusted life years (DALYs) were used as the measure of health. The author applied dynamic spatial panel data models with fixed effects and spatial autocorrelation of the error term. The models were estimated using a novel, modified quasi maximum likelihood method based on M-estimators. The approach is resistant to deviations from the assumptions on the distribution of initial observations. The estimation of initial observations is a severe weakness of standard methods based on the maximization of the quasi-likelihood function in the case of short panels. M-estimators are consistent and asymptotically normally distributed. The empirical analysis covers the specification, estimation, and verification of the models.

Key words: dynamic spatial panel data models, M-estimation, fixed effects, short panels, DALYs – disability-adjusted life years, the level of health, socio-economic factors.

1. INTRODUCTION

Public health affects the productivity of labour and human capital as well as public spending. Investing in health is one of the priorities of the "Global Europe 2050" strategy, which aims at securing sustainable economic growth in Europe (Eurostat's Report for the European Commission, 2017). The implementation of programs promoting a healthy lifestyle and an efficient allocation of funds for balanced healthcare systems at national and international levels should be supported

* Agnieszka ORWAT-ACEDAŃSKA, University of Economics in Katowice, Department of Demography and Economic Statistics, ul. Bogucicka 3, 40-287 Katowice, Poland; e-mail: agnieszka.orwat@ue.katowice.pl

by the identification of factors affecting the health of a population. Statistical modelling of the relationship between socio-economic factors and the levels of health in time and space also facilitates designing an adequate public policy in this area.

The aim of the paper is to investigate the relationship between socio-economic factors and the level of health of citizens of selected European countries. The level of health can be assessed using either summary health measures or disease burden indicators (Murray *et al.*, 2002; Wróblewska, 2008; Robine, 2006). Healthy life years (HLY) (Trzpiot, Orwat-Acedańska, 2016) is one of the most popular measures from the former group. The paper uses a disease burden measure, i.e. disability-adjusted life years – DALYs. The DALY is a measure that expresses the total life years lost due to premature death or damage to health caused by a disease. Series of DALYs are characterized by significant autocorrelations in both time and space.

At the level of individual countries, these dependencies can be investigated using spatial panel data (SPD) econometric models. Spatial panel data has the cross-sectional and time dimensions and, additionally, the time dimension includes the spatial context. In other words, spatial models account for spatial interaction effects and dependencies between neighbouring objects, which allows for more informative and precise analysis of various phenomena.

The paper proposes spatial dynamic panel data (SDPD) models with fixed effects and spatial autocorrelation of the error term. These models extend the spatial panel framework with time dynamics that represents a trend in a dependent variable. In particular, the paper focuses on Dynamic Spatial Autoregressive Fixed Effects Models (DSAR-FEM).

In contrast to static models, most of the estimators of dynamic spatial panel models are biased and inefficient. This problem is particularly severe in the case of a short panel (with a small number of periods). The shortcomings of the standard estimation methods led to the development of several sublime alternative techniques. Elhorst (Elhorst, 2010c) and Su, Yang (Su, Yang, 2015) proposed the quasi-ML (QML) estimation of the SDPD model with short panels. The main difficulty in using the ML or QML method to estimate SDPS models with short panels is the modelling of initial observations (the data generating process for the pre-sample period)[1] because statistical properties of the ML estimators largely depend on the assumptions of initial observations (Dańska-Borsiak, 2011). The model for initial differences involves an unknown process starting time. Moreover, its predictability typically requires that time-varying regressors be trend or first-difference stationary. When there are many time-varying regressors in the model, the modelling of the initial difference may introduce too many additional parameters, causing an efficiency decline (Yang, 2018). It is highly desirable to have a method that is free from the specification of initial observations and pos-

[1] Alternative estimation methods, such as the IV-GMM approach, despite being free from the assumption on initial observations, deliver inefficient estimates (Dańska-Borsiak, 2011).

sesses good statistical properties, especially in the case of short panels. It should be noted that ML and QML estimators are considered more efficient than GMM estimators for dynamic panel data models (Hsiao *et al.*, 2002; Binder *et al.*, 2005; Bum, Carce, 2005; Gourieroux *et al.*, 2010; Kruiniger, 2013).

In the paper, the models are estimated using a novel, modified quasi maximum likelihood (QML) method with M-estimators proposed recently by Yang (Yang, 2018). The approach is free from assumptions on the distribution of initial observations. Moreover, M-estimators are consistent and asymptotically normally distributed. The empirical analysis covers the specification, estimation, and verification of the models.

The paper is organized as follows. First, the concept of the DALY is presented in the context of the brief review of literature related to the subject. The evolution of how the DALY measure has been applied in the European countries since 1990 is discussed, and the rationale behind the choice of a model class and the criteria for the selection of countries is presented. The third section consists of three subsections. First, based on literature, the general idea of spatial models and their estimation methods are discussed. The second subsection is concerned with the analytical form of the models used in the study, while the third one describes the M-estimation procedure. The fourth section contains an empirical analysis. It consists of two subsections. First, endogenous and explanatory variables are presented together with the main assumptions used in the empirical study, which is followed by the discussion of the results. The last section concludes the paper.

2. DISABILITY-ADJUSTED LIFE YEARS (DALYS)

2.1. The essence of measurement and literature review

Disability-adjusted life years (DALYs) represent an increasingly popular population health metric. DALYs belong to a family of population health summary indicators and measure the Global Burden of Disease (GBD). The indicator was introduced by Murray and Lopez (Murray, 1994; Murray, 1996; Murray, Lopez, 1996a, 1996b), who conducted a GBD study (World Bank, 1993) for the World Bank and the World Health Organization. DALYs are based on measuring health gaps, as opposed to measuring health expectancies (Murray, Lopez, 1994), and as such the indicators measure the difference between current conditions and a selected target, for example an ideal health state.

The DALYs are the sum of the Years Lived with Disability (YLD) and the Years of Life Lost to premature death (YLL). One DALY is thus one lost year of healthy life. DALYs are calculated as follows (Murray, Lopez, 1994):

$$\text{DALYs} = \text{YLD} + \text{YLL} \quad (1)$$

where: YLD – the adjusted number of years lived with disability; YLL – the number of years of life lost due to premature mortality.

Years lived with disability (YLD) refer to years lived in health worse than ideal. To estimate the YLD at the level of a population, the number of disability cases is multiplied by an average duration of a disease and the weight factor that reflects the severity of a disease on a scale from 0 (perfect health) to 1 (dead). The basic formula for one disabling event is:

$$\text{YLD} = \text{I x DW x L} \quad (2)$$

where: I – the number of incidents; DW – a disability weight; L – an average duration of a disability (years). Disability weights for 22 diseases are calculated using a person trade-off method. The weights are also used to define 7 classes of disability and a distribution of severity of a few hundreds of treated and untreated diseases (Murray, Lopez, 1994). The weights also account for a person's age. The paper proposes higher weights to the same diseases and disabilities in the case of young and middle-aged people and lower ones for infants and elderly people. Moreover, the use of discounting related to the level of health and health benefits results in higher weights for the current levels of health and lower for the expected ones.

The YLL metric essentially corresponds to the number of deaths multiplied by the standard life expectancy at the age at which death occurs, and it can be rated according to social preferences (see below). The basic formula for calculating the YLL for a given cause, age or sex is: (Murray, 1996)

$$\text{YLL} = \text{N x L} \quad (3)$$

where: N – the number of deaths; L – standard life expectancy at the age of death (in years). To estimate the YLL at the level of a population, age-specific mortality rates must be combined with life expectancy for fatal cases, had they not developed into the disease. If mortality affects the population in a random fashion, life expectancy can be derived from standard life tables. Murray (Murray, 1996) proposed a table based on the highest observed national life expectancy (for Japanese women), taking into account differences in life expectancy between men and women. If mortality affects a susceptible sub-population, the use of standard life expectancy would lead to a gross overestimation of the YLL. In this case, disease-specific information is necessary to estimate the additional loss of life years by the disease under consideration. For estimating the YLL for more than 100 causes of death, the following information sources are used: the death registration system based on International Classification of Diseases (ICD-9), sample data, epidemiological estimates, and mortality models for selected causes of death.

The philosophical and methodological aspects of the DALY calculation were examined in detail in Murray, 1994; Murray, Acharya, 1997. The measure is also widely discussed in literature (Anand, Hanson, 1997; Anand, Hanson, 1998; Laurell, Arellano, 1996; Barker, Green, 1996; Berman, 1995; Desjarlais *et al.*, 1995; Lozano *et al.*, 1995; Martens *et al.*, 1995). The steps preceding the actual calculation, however, are still under-researched. Devleesschauwer (Devleesschauwer, 2014) attempted to address this gap by presenting a stepwise approach towards the DALY calculation (Devleesschauwer, 2014).

The paper identified the socio-economic factors that are related to DALYs in selected European countries. To the author's knowledge, this issue has not been studied with the use of the spatial statistical models. Furthermore, the paper posited that it is of crucial importance to go beyond standard statistical and econometric procedures based on dynamic spatial panel data models analysis. Instead, the application of M-estimation of fixed effects spatial dynamic models with short panels was proposed for investigating the relationships between selected socio-economic factors and the DALY measure in Europe in space and time, which can also constitute an important contribution to literature. The identification of the factors affecting a population's health should contribute to the creation of reliable and effective policies within national and internationals public health management strategies.

2.2. DALYs in European countries in the last 25 years

Currently, the estimated years of life lost due to premature mortality, YLL, and the years of life lived with disability, YLD, account for almost 300 diseases belonging to three major groups and twenty subgroups of causes. The values of DALYs for almost every country of the world are published by the Institute for Health Metrics and Evaluation (IHME), an independent population health centre at the University of Washington.

European countries are heterogeneous in terms of the DALY measure (observed in the last 25 years). This is primarily because of differences in the geopolitical conditions that affect the distributions of life expectancy, epidemiological estimates, and mortality models.

Figure 1 depicts mean values of DALYs for selected 26 European countries divided into three groups in the years 1990–2015:

a) **Group A**: Lithuania, Latvia, Estonia, Hungary;

b) **Group B**: Austria, Belgium, Switzerland, the Czech Republic, Germany, Denmark, Spain, Finland, France, the United Kingdom, Greece, Ireland, Italy, Luxembourg, the Netherlands, Norway, Poland, Portugal, Slovakia, Slovenia, Sweden;

c) **Group C**: Iceland.

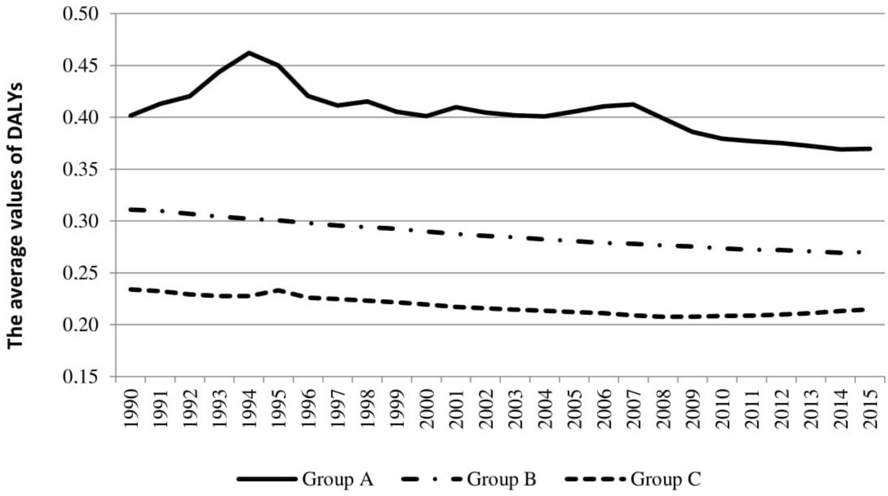

Fig 1. Average values of DALYs for countries for the three groups (A, B, and C) in the period 1990–2015

Source: Own work based on the Institute for Health Metrics and Evaluation (IHME) data.

The DALY time series are characterised by a downward trend, which is related to a global increase in life expectancy and changes in the burden of diseases.[2] Moreover, spatial heterogeneity among European countries can also be observed. The countries classified in Group A – Lithuania, Latvia, Estonia, and Hungary – are characterised with the highest values of DALYs. The countries from Group B are characterised by lower DALYs compared to the previous group. Clearly, the lowest values of the studied measure are observed in Iceland and, therefore, the country forms a separate group in the figure. The differences reflect the geopolitical locations of the countries, which are correlated with the burden of diseases.

The spatial heterogeneity of the DALY measure is shown in Figure 2. The DALY values for the countries shown there are average values for the years 1990–2015. Apart from the four countries in Group A, high values of DALYs are observed in other Central and Eastern Europe countries, such as Poland or the Czech Republic. Then, the measure takes considerably lower values in the original EU Member States.

The DALY measure is calculated taking into account the entire population of a country. It is also possible to estimate the measure for selected subpopulations defined by age, e.g. people more than *x* years old. The study also examines the DALYs_70+ measure, which is calculated for people over 70 years of age. This subpopulation is important from the point of view of public health management

---

[2] In 1990s, major diseases comprised respiratory system diseases and perinatal diseases, whereas in the last decade the considerable increase in civilizational diseases such as cardiovascular diseases and neoplasms has been observed. In the future, a growing role of car accidents and mental diseases (mainly depression) is expected.

and social security spending. Similarly to DALYs, the downward trend can be observed for DALYs_70+. Fig. 3 depicts these values for the years 1990–2015.

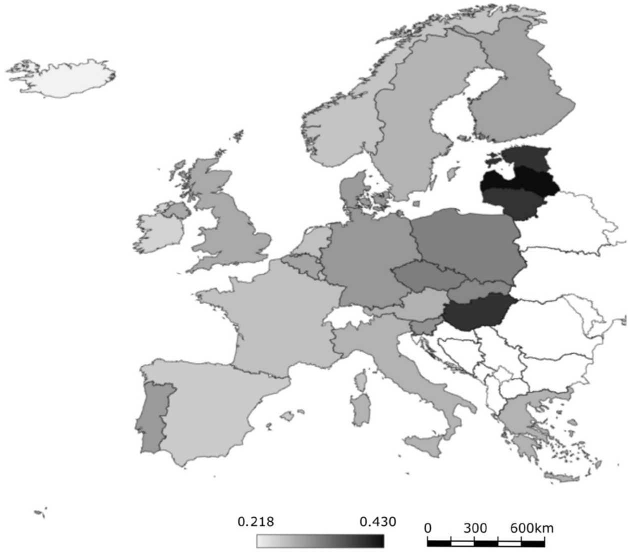

Fig. 2. Average values of DALYs in the European countries in the years 1990–2015.
Source: own work based on the Institute for Health Metrics and Evaluation (IHME) data.

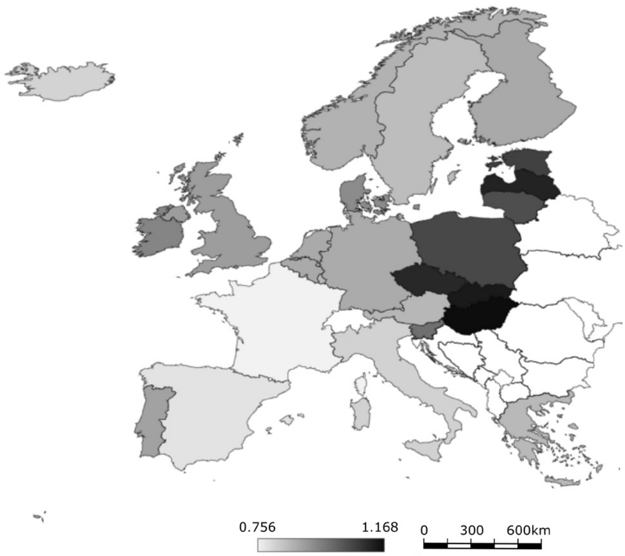

Fig. 3. Average values of DALYs_70+ s for countries in groups A, B and C in the years 1990–2015
Source: own work based on the Institute for Health Metrics and Evaluation (IHME) data.

3. THE THEORETICAL BASIS OF SDPD MODELS

In the standard panel data (SPD) methods, two additional effects can be taken into account in addition to the usual effect of explanatory variables. One is an unobserved object-specific effect, independent of time, also referred to as a group effect. The other is also an unobserved, time-specific effect, constant across objects, which sometimes is shortened as a time effect. Among the models with object or time-specific effects, one can identify varying-parameter models (e.g. Random Coefficient Models (RCM), Seemingly Unrelated Regressions (SUR)), and error decomposition models (e.g. Fixed Effects Models (FEM), Random Effects Models (REM)).

The plethora of the model types is accompanied by a variety of estimation procedures, such as the Maximum Likelihood (ML) method, the Generalized Method of Moments (GMM), and the Method of Instrumental Variables (IVM). Literature on spatial panel data was hugely inspired by the seminal paper of (Anselin, 1988), where the ML estimation techniques were studied. The typology of the spatial panel data models and modifications of the ML estimators suitable for certain model types are presented in Elhorst 2005. The effective estimation techniques of various types of the spatial panel models are discussed in Elhorst (2010a, pp. 377–407; 2010b, pp. 9–28) and Lee, Yu (2010a, pp. 165–185; 2010b; 2010c, pp. 255–271). Spatial interaction in panel models can be modelled as the processes of: spatial autoregression for a dependent variable (Spatial Autoregressive – SAR), spatial autocorrelation of the error term (Spatial Error – SEM), moving average spatial autocorrelation of the error term (Spatial Moving Average – SMA), and the spatial lags of regressors (Spatial Crossregressive – SCM).

Additional difficulties related to spatial panel data modelling occur if one wants to include the time dynamics of a dependent variable. The dynamics in a linear model can be accounted for in many ways. One is to add the time lags of a dependent variable.[3]

Extensive literature, summarized by Anselin, 2001; Anselin *et al.*, 2008, exists on spatial dynamic panel data (SDPD) models. Spatial effects in dynamic panel data models may appear in the form of the spatial lags of a response variable (Yu *et al.*, 2008; Yu, Lee, 2010; Lee, Yu, 2010d; Korniotis, 2010; Elhorst, 2010c), spatial errors (Elhorst, 2010c; Yang *et al.*, 2006; Su, Yang, 2015) or space-time lag.

3.1. Dynamic Spatial Error Fixed Effects Model (DSE_FEM)

As already mentioned, in order to investigate the relationship between socio-economic factors and the DALYs in EU Member States, the study used dynamic spatial panel data models with fixed effects and spatial autocorrelation of the error

[3] Alternative approaches include: 1) adding a time variable to a set of regressors; 2) using a two-way model; 3) using a time-varying parameters model.

term. The dependent variable and k regressors were observed for N spatial units and T periods. Dynamic Spatial Error Fixed Effects Model (DSE_FEM) has the following form:

$$y_{it} = \rho y_{it-1} + \sum_{j=1}^{k} \beta_j x_{ijt} + \mu_i + \varepsilon_{it} \tag{4}$$

$$\varepsilon_{it} = \lambda \sum_{l=1}^{N} w_{il} \varepsilon_{lt} + v_{it} \tag{5}$$

$$i = 1, ..., N;\ t = 1, ..., T;\ l = 1, ..., k,$$

where: y_{it} – a dependent variable; – an element of the spatial weight matrix; ρ – a time autoregression parameter; – a fixed effects parameter; x_{ijt} – a regressor; β_j – a parameter representing an impact of regressors on a dependent variable; $\boldsymbol{\varepsilon}_{it,}$ v_{it} – the error term; λ – a spatial autoregression parameter. Random variables v_{it} are normally, independently and identically distributed with the expected value equal to 0.

3.2. The idea of unified M-estimation of the Dynamic Spatial Error Fixed Effects Model (DSE_FEM)

The main difficulty in using the ML or QML method to estimate SDPS models with short panels is the modelling of initial observations (the data generating process for the pre-sample period). The model for initial differences involves an unknown process starting time. Moreover, its predictability typically requires that the time-varying regressors be trend or first-difference stationary. When there are many time-varying regressors in the model, the modelling of an initial difference may introduce too many additional parameters, causing an efficiency decline (Yang, 2018).

Yang (2018) proposed a unified initial-condition free approach to estimate the SDPD models with fixed effects. The method generates estimators that are consistent and asymptotically normal. Corrections on the conditional quasi-scores are totally free from the specification of the distribution of initial differences. The proposed estimator is simply referred to in this paper as the *M*-estimator in view of Huber (1981) or van der Vaart (1998). The method proposed by Yang (2018) for the covariance matrix estimation is valid only when T is small, but when T is large, a standard alternative (plug-in method) based on the conditional variance of the adjusted quasi-scores, treating initial differences as exogenous, can be used. The exact steps leading to the calculation of estimates are complicated and, therefore, they are not presented here. The details can be found in Yang (2018).

4. EMPIRICAL ANALYSIS

4.1. Data, variables and empirical procedure

In empirical analysis, two groups of DSPD models were used. One consisted of models (4-5), in which DALYs were an endogenous variable, and the other consisted of models (4-5), in which DALYs_70+ were an endogenous variable. Both indicators are expressed in per capita terms. The selection of potential exogenous variables focused on health and lifestyle determinants as well as socio-economic factors. Eight exogenous variables were examined for each country under study (Table 1).

Table 1. Exogenous variables used in the first stage of the estimation procedure for models (4–5) in the specification procedure

Variable	Symbol	Description of the variable
X_1	GDP	GDP per capita in constant US dollars
X_2	ΔGDP	Real GDP growth rate
X_3	ALCOH	Yearly alcohol consumption per capita in litres
X_4	AIR_POL	Yearly nitric oxide emissions per capita in kilograms
X_5	EDUC	Average education years
X_6	H_CARE	Healthcare spending as a percentage of GDP
X_7	N_BEDS	Number of hospital beds per 1,000 inhabitants
X_8	SOCIAL	Social spending as a percentage of GDP

Source: own work.

The final set of analysed countries (objects in models) was determined using two criteria. First, the countries that are relatively homogeneous in terms of the measure were selected out of the 26 countries for which the average DALY measure is depicted in Fig. 1. It was assumed that it would be incorrect to include all the 26 countries mentioned in Section 2.2 in one model because of significant differences in DALY values. Second, the analysis only included those countries for which the complete set of data relating to the eight exogenous variables was available.

As a result, the sample consisted of 17 countries, mostly of the "original EU"[4] (without Luxemburg, yet including Iceland, Norway and the Czech Republic):

[4] The term refers to the countries that constituted the European Union prior to the accession of the new members in 2004.

Portugal, Spain, France, Belgium, the Netherlands, the Czech Republic, Italy, Greece, Austria, Germany, Denmark, the United Kingdom, Ireland, Iceland, Norway, Sweden, and Finland.

The final investigation period of 2003–2013 was a compromise between the criteria for the selection of countries and availability of exogenous variables for each country.

Exogenous variables were obtained from the OECD database. Endogenous variables (DALYs, DALYs_70+) came from the database of the Institute for Health Metrics and Evaluation. The spatial matrix **W** was created using the common border criterion. All the computations were carried out in Matlab using the procedures written by Yang.

4.2. Results

Table 2 presents the estimates of models (4-5) and their statistical significance for the DALY measure. Five out of eight exogenous variables are statistically significant at the level of significance equal to 0.1. These are: **ΔGDP** (GDP growth rate), **ALCOH** (alcohol consumption), **EDUC** (years of education), **H_CARE** (healthcare spending), **SOCIAL** (social spending). Dynamic and spatial autocorrelations are also significant, which corroborates the choice of the modelling tool.

Table 2. Estimation results for the models (4–5) with DALYs as endogenous variable

Exogenous variables	β_j	$s(\beta_j)$	*t* statistic	p-value*
GDP	-0.010	0.000	-1.169	0.243
ΔGDP	-0.697	0.161	-4.328	**0**
ALCOH	-0.052	0.020	-2.622	**0.009**
AIR_POL	-0.005	0.004	-1.183	0.237
EDUC	0.017	0.009	2.027	**0.043**
H_CARE	0.051	0.025	1.999	**0.046**
N_BEDS	0.010	0.040	0.246	0.806
SOCIAL	-0.040	0.012	-3.347	**0.001**
Parameter				
ρ	0.972	0.035	28.047	**0**
λ	0.241	0.109	2.222	**0.026**

* The estimates in bold are statistically significant (α = 0.1)

Source: own calculation.

As some variables were discovered to be insignificant, the model was subsequently re-estimated with the significant variables only. The estimates obtained during this step are shown in Table 3.

Table 3. The estimation results of the respective models (4-5) for DALYs

Exogenous variables	β_j	$s(\beta_j)$	t statistic	p-value*
ΔGDP	-0.790	0.194	-4.072	**0**
ALCOH	-0.058	0.021	-2.761	0.006
EDUC	0.018	0.006	2.782	**0.005**
H_CARE	0.042	0.024	1.781	**0.075**
SOCIAL	-0.032	0.013	-2.552	**0.011**
Parameter				
ρ	0.991	0.027	36.530	**0**
λ	0.239	0.114	2.109	**0.035**

* The estimates in bold are statistically significant (α = 0.1)

Source: own calculation.

As a result of the respecification, all the exogenous variables are statistically significant. The DALY measure for the entire population is correlated with both economic factors (GDP growth rate, social and healthcare spending in relation to GDP) and social factors (years of education, lifestyle, i.e. alcohol consumption). Although the p-value for these factors is below 0.1, which means they are all statistically significant, DALYs are characterized by near-unit-root behaviour and the error terms are also significantly spatially correlated, although the correlation is rather weak (value of lambda). In consequence, the time autocorrelation parameter ρ is close to one, so a slower decrease in the DALYs measure (its higher growth rates) is negatively correlated with GDP dynamics, alcohol consumption, and social spending, and positively related to the level of education and healthcare spending. The manifestations of some relationships can be explained by considering the dependence of the components of the DALY measure – YLL and YLD. The estimated model takes the following form:

$$
\begin{aligned}
DALYs_{it} &= 0.99 DALYs_{it-1} - 0.79\,\Delta GDP_{it} - 0.06\,ALCOH_{it} + \\
&+ 0.02\,EDUC_{it} + 0.04\,H\_CARE_{it} - 0.03\,SOCIAL_{it}
\end{aligned} \tag{6}
$$

$$
\hat{\varepsilon}_{it} = 0.23 \sum_{l=1}^{N} w_{il} \varepsilon_{jt} \tag{7}
$$

The second stage of the study involved the estimation of the model (4-5) and the testing of the significance of parameters for the DALYs_70+ series. Initially, all the regressors were included. The results are presented in Table 4.

Table 4. Estimation results for model (4-5) for the DALYs_70+ measure

Exogenous variables	β_j	$s(\beta_j)$	*t* statistic	p-value*
GDP	-0.003	0.000	-0.013	0.989
ΔGDP	-2.778	1.222	-2.274	**0.023**
ALCOH	-0.091	0.062	-1.452	0.146
AIR_POL	-0.031	0.013	-2.285	**0.022**
EDUC	0.100	0.040	2.483	**0.013**
H_CARE	0.057	0.076	0.754	0.451
N_BEDS	0.036	0.087	0.413	0.680
SOCIAL	-0.022	0.024	-0.907	0.364
Parameter				
ρ	0.957	0.047	20.564	**0**
λ	0.282	0.069	4.107	**0**

* The estimates in bold are statistically significant (α = 0,1)

Source: own calculation.

In this case, only three out of eight exogenous variables are statistically significant at the level of significance equal to 01. These are: **ΔGDP** (GDP growth rate), **AIR_POL** (air pollution), and **EDUC** (average years of education). Dynamic and spatial autocorrelations are also statistically significant. Subsequently, the model is respecified to eliminate the insignificant variables. The results are presented in Table 5.

Table 5. The estimation results of the respecified model for DALYs_70+

Exogenous variables	β_j	$s(\beta_j)$	*t*- statistic	p-value*
ΔGDP	-2.254	1.214	-1.858	**0.063**
AIR_POL	-0.023	0.011	-2.026	**0.043**
EDUC	0.098	0.037	2.661	**0.008**
Parameter				
ρ	0.958	0.027	35.001	**0**
λ	0.279	0.053	5.248	**0**

* The estimates in bold are statistically significant (α = 0,1)

Source: own calculation.

Disability-adjusted life years for people over 70 years of age in the investigated countries are affected by a GDP growth rate (economic factor) as well as social factors such as the years of education and air pollution. These relationships are accompanied by very strong dynamics and weak spatial autocorrelation of DALYs_70+. The estimated model has the following form:

$$DALYs\_70^{+}{}_{it} = 0.96\ DALYs\_70^{+}{}_{it-1} - 2.25\,\Delta GDP_{it} - 0.02\ AIR\_POL_{it} + \\ +0.10\ EDUC_{it} \tag{8}$$

$$\hat{\varepsilon}_{it} = 0.28 \sum_{l=1}^{N} w_{il} \varepsilon_{jt} \tag{9}$$

Compared to the model in which the DALY was an endogenous variable, differences in the set of variables statistically significant for the model with the endogenous variable DALY_70+ indicate that the selection of population (based on age) is an important aspect in the identification of socio-economic factors determining disease burden.

5. CONCLUSION

The paper identified the factors affecting disability-adjusted life years based on the sample that consisted of 17 European countries and span the years 2003–2013. Spatial dynamic panel models were used to analyse the persistence and spatial autocorrelation of the variables under study. The novel estimation method introduced recently by Yang (2018) was applied. The DALY measure was found to be significantly related to several economic, social, and environmental factors such as healthcare spending, alcohol consumption or air pollution, but also a GDP growth rate, and years of education. Apart from the former, rather obvious, factors identified by several other studies, the latter two carry interesting policy implications. Significant correlation with a GDP growth rate implies that the DALY indicator was affected by business cycle fluctuations and, in particular, the recent financial crisis. It means that not only does preventing severe economic crises have an indirect impact on public health through healthcare or social spending, but it also affects public health directly. Finally, the correlation of DALYs with years of education confirms an important role of education in improving the level of health in the society.

Due to fundamental differences between the countries in Group A and Group B (defined in section 2.2), the countries from Region A could not be included in the sample. Moreover, the countries from Group A cannot form a separate sample since their number did not allow for the generation of the amount of data sufficient for the development of spatial models.

The factors affecting DALYs can still be identified irrespective of spatial dependence and using standard panel data models, which, however, is beyond the scope of this paper.

The investigation of the spatial-temporal relationship (accounting for a dynamic effect) between the DALY measure for European countries and socio-economic factors requires several important choices relating to such issues as the period of a study, a population in terms of age, and the geopolitical location of countries.

REFERENCES

ANAND, S. and HANSON, K. (1997), 'Disability-adjusted lost years – a critical review', *Journal of Health* Economics, 16, pp. 685–702.

ANAND, S. and HANSON, K. (1998), 'DALYs: efficiency versus equity', *World Development*, 26 (2), pp. 307–310.

ANSELIN, L. (1988), *Spatial Econometrics: Methods and Models*, The Netherlands: Kluwer Academic Press.

ANSELIN, L. (2001), 'Spatial Econometrics', [in:] BALTAGI, B. H. (eds.), *A companion to theoretical econometrics*, Massachusetts: Blackwell Publishers Ltd., pp. 310–330.

ANSELIN, L., LE GALLO, J. and JAYET, J. (2008), 'Spatial panel econometrics', [in:] MATYAS, L., SEVESTRE, P. (eds.), *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*, Berlin-Heidelberg: Springer-Verlag, pp. 625–660.

BARKER, C. and GREEN, A. (1996), 'Opening the Debate on DALYs'*, Health Policy and Planning*, 11, pp. 179–183.

BERMAN, S. (1995), 'Otitis media in developing countries', *Pediatrics*, 96, pp. 126–131.

BINDER, M., HSIAO, C. and PESARAN, M. H. (2005), 'Estimation and inference in short panel vector autoregressions with unit roots and cointegration', *Econometric Theory*, 21, pp. 795–837.

BUN, M. J. and CARREE, M. A. (2005), 'Bias-corrected estimation in dynamic panel data models', *Journal of Business and Economic Statistics*, 23, pp. 200–210.

DAŃSKA-BORSIAK, B. (2011), *Dynamiczne modele panelowe w badaniach ekonomicznych*, Łódź: Wydawnictwo Uniwersytetu Łódzkiego.

DESJARLAIS, R., EISENBERG, L., GOOD, B., and KLEINMAN, A. (1995), *World mental health: problems and priorities in low income countries*, New York: Oxford University Press.

DEVLEESSCHAUWER, B., HAVELAAR, A. H., MAERTENS DE NOORDHOUT, C., HAAGSMA J. A., PRAET, N., DORNY, P., DUCHATEAU, L., TORGERSON, P. R., VAN OYEN H. and SPEYBROECK, N. (2014), 'DALY calculation in practice: a stepwise approach', *International Journal of Public Health*, 59 (3), pp. 571–574.

ELHORST, J. P. (2005), 'Unconditional maximum likelihood estimation of linear and loglinear dynamic models for spatial panels', *Geographical Analysis*, 37, pp. 85–106.

ELHORST, J. P. (2010a), 'Spatial Panel Data Models', [in:] FISCHER, M. M., GETIS, A., (eds), *Handbook of Applied Spatial Analysis*, Springer, Berlin.

ELHORST, J. P. (2010b), 'Applied spatial econometric: raising the bar', *Spatial Economic Analysis*, 5 (1), pp. 9–28.

ELHORST, J. P. (2010c), 'Dynamic panels with endogenous interaction effects when T is small', *Regional Science and Urban Economics*, 40, pp. 272–282.

EUROSTAT'S REPORT FOR THE EUROPEAN COMMISSION (2017), 'Global Europe 2050'.

GOURIEROUX, C. and PHILLIPS, P. C. B., YU, J. (2010), 'Indirect inference for dynamic panel models', *Journal of Econometrics*, 157, pp. 68–77.

HSIAO, C., PESARAN, M. H. and TAHMISCIOGLU, A. K. (2002), 'Maximum likelihood estimation of fixed effects dynamic panel data models covering short time periods', *Journal of Econometrics*, 109, pp. 107–150.

HUBER, P. J. (1981), *Robust Statistics*. New York: Wiley.

KORNIOTIS, G. M. (2010), 'Estimating panel models with internal and external habit formation', *Journal of Business and Economic Statistics*, 28, pp. 145–158.

KRUINIGER, H. (2013), 'Quasi ML estimation of the panel AR(1) model with arbitrary initial conditions', *Journal of Econometrics*, 173, pp. 175–188.

LAURELL, A. C. and ARELLANO, L. O. (1996), 'Market commodities and poor relief: The World Bank proposal for health', *Journal of Health Economics*, 26 (1), pp. 1–18.

LEE, L. F. and YU, J. (2010a), 'Estimation of spatial autoregressive panel data model a with fixed effects', *Journal of Econometrics*, 154(2), pp. 165–185.

LEE, L. F. and YU, J. (2010b), *Estimation of spatial panels: random components vs. fixed effects*, Manuscript, Ohio State University.

LEE, L. F. and YU, J. (2010c), 'Some recent developments in spatial panel data models', *Regional Science and Urban Economics*, 40, pp. 255–271.

LEE, L. F. and YU, J. (2010d), 'A spatial dynamic panel data model with both time and individual fixed effects', *Econometric Theory*, 26, pp. 564–597.

LOZANO, R., MURRAY, C. J. L., FRENK, J. and BOBADILLA, J. L. (1995), 'Burden of diseases assessment and health system reform: results of a study in Mexico', *Journal of International Development*, 7 (3), pp. 555–564.

MARTENS, W. J., NIESSEN, L. W., ROTMANS, J., JETTEN, T. H. and McMICHAEL A.J. (1995), 'Potential impact of global climate change on malaria risk', *Environmental Health Perspectives*, 103 (5), pp. 458–64.

MURRAY, C. J. L. (1994), 'Quantifying the burden of disease: the technical basis for disability-adjusted life years', *Bulletin of the World Health Organization*, 72 (3), pp. 429–445.

MURRAY, C. J. L. (1996), 'Rethinking DALYs', [in:] MURRAY, C. J. L., LOPEZ, A. D., *The Global Burden of Disease and Injury Series*, Harvard School of Public Health, World Health Organization, World Bank, Boston, 1, pp. 1–98.

MURRAY, C. J. L. and LOPEZ, A. D. (1994), *Global comparative assessments in the health sector: disease burden, expenditures and intervention packages*., Geneva, World Health Organization.

MURRAY, C. J. L. and LOPEZ, A. D. (1996a), 'A comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020', *The Global Burden of Disease and Injury* Series. *The Global Burden of Disease*, 1, Harvard School of Public Health, World Bank, World Health Organization.

MURRAY, C. J. L. and LOPEZ, A. D. (1996a), 'A compendium of incidence, prevalence and mortality estimates for over 200 conditions', *The Global Burden of Disease and Injury Series*. *The Global Burden of Disease*, 2, Harvard School of Public Health, World Bank, World Health Organization.

MURRAY, C. J. L., SALOMON, J. A., MATHERS, C. D. and D. LOPEZ A. D. (2002), *Summary measures of population health – concepts, ethics, measurement and applications*. Geneva, World Health Organization.

ROBINE, J. M. (2006), 'Summarizing Health Status' [in:] PENCHEON, D., GUEST, C., MELZER, D. and GRAY, J. A. M., *Oxford Handbook of Public Health Practice*, Oxford University Press.

SU, L. and YANG, Z. (2015), 'QML estimation of dynamic panel data models with spatial errors', *Journal of Econometrics*, 185, pp. 230–258.

TRZPIOT, G. and ORWAT-ACEDAŃSKA, A. (2016), 'Spatial quantile regression in analysis of healthy life years in the European Union countries', *Comparative Economic Research*, 19 (5), pp. 179–199.

VAN DER VAART, A. W. (1998), *Asymptotic Statistics*. Cambridge University Press.

WRÓBLEWSKA, W. (2008), 'Sumaryczne miary stanu zdrowia populacji', *Studia Demograficzne*, pp. 153–154.

YANG, Z. (2018), 'Unified M-Estimation of Fixed-Effects Spatial Dynamic Models with Short Panels', *Journal of Econometrics*, 205, pp. 423–447.

YANG, Z., LI, C. and TSE, Y. K. (2006), 'Functional form and spatial dependence in dynamic panels', *Economics Letters*, 91, pp. 138–145.

YU, J., DE JONG, R. and LEE, L. F. (2008), 'Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large', *Journal of Econometrics*, 146, pp. 118–134.

YU, J. and LEE, L. F. (2010), 'Estimation of unit root spatial dynamic panel data models', *Econometric Theory*, 26, pp. 1332–1362.