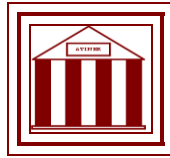


**Athens Institute for Education and Research  
ATINER**



**ATINER's Conference Paper Series  
PUH2015-1586**

**Convergence of Health Status in the  
European Union: A Spatial Econometric  
Approach**

**Renata Stańczyk  
Assistant in the Department of Spatial Econometrics  
University of Lodz  
Poland**

An Introduction to  
ATINER's Conference Paper Series

ATINER started to publish this conference papers series in 2012. It includes only the papers submitted for publication after they were presented at one of the conferences organized by our Institute every year. This paper has been peer reviewed by at least two academic members of ATINER.

Dr. Gregory T. Papanikos  
President  
Athens Institute for Education and Research

This paper should be cited as follows:

**Stańczyk, R. (2015). "Convergence of Health Status in the European Union:  
A Spatial Econometric Approach", Athens: ATINER'S Conference Paper  
Series, No: PUH2015-1586.**

Athens Institute for Education and Research  
8 Valaoritou Street, Kolonaki, 10671 Athens, Greece  
Tel: + 30 210 3634210 Fax: + 30 210 3634209 Email: [info@atiner.gr](mailto:info@atiner.gr) URL:  
[www.atiner.gr](http://www.atiner.gr)  
URL Conference Papers Series: [www.atiner.gr/papers.htm](http://www.atiner.gr/papers.htm)  
Printed in Athens, Greece by the Athens Institute for Education and Research. All rights reserved. Reproduction is allowed for non-commercial purposes if the source is fully acknowledged.  
ISSN: 2241-2891  
14/09/2015

## **Convergence of Health Status in the European Union: A Spatial Econometric Approach<sup>\*</sup>**

**Renata Stańczyk**  
**Assistant in the Department of Spatial Econometrics**  
**University of Lodz**  
**Poland**

### **Abstract**

In this paper we analyse the dynamics of health inequalities across the EU regions of NUTS II level in the time period 2002-2012. Our study is based on specific research methods. First, we apply a structural equation model (MIMIC model) to estimate the synthetic measure of health pressure index.

Second, we apply existing tools developed in the economic growth literature to study health convergence measured by traditional indicators and the obtained MIMIC measure. To verify the hypothesis of beta-convergence we use spatial econometric models which additionally allow us to take into consideration the geographic dependence among the surveyed regions.

We derive two main results. The first is about supporting the hypothesis on the impact of socio-economic factors on the public health status. The second is about reduction in health inequalities in the years 2002-2012.

**Keywords:** EU regions, MIMIC model, Public health.

---

<sup>\*</sup> The project was funded by the National Science Centre on the basis of the decision number DEC-2013/11/N/HS4/03404.

## Introduction

Empirical studies on regional health conducted so far do not explain in a comprehensive (model) manner the relationships that occur between particular determinants and indicators of the health status. These analyses undertaken so far deal mainly with the impact of health on economic growth and vice versa (Preston 1975, Anand and Ravallion 1993, Pritchett and Summers 1996, Schultz 2008) and on the spatial differentiation of health status. Research concerning strictly the assessment of health status is based primarily on the concept of the health production function proposed by Auster et al. (1969), who estimated the relationship between mortality and social and economic factors. This study was followed by Grossman (1972). According to Or (2000) health production function, in the broadest sense, describes the relationship or flows of health inputs and flows of health outputs over a specified period. Production process depends on the health care system and the size of the resources involved in it, as well as social, economic and physical conditions (Or 2000: 55). Folland et al. (2001) argue that health status is an increasing function of the use of health care and the factors that affect this status are biological characteristics, environment and lifestyle.

Population health as an economic category is complex, thing which makes it difficult to measure. The commonly used health status indicators, such as the average life expectancy rate and the infant mortality rate, do not measure the phenomenon as a whole. The diversity of the health status in the EU regions may be explained by the regional diversity of the health-determining factors. Among them one can distinguish elements of the health protection infrastructure and a number of economic and social factors. In order to assess the health status on the macroeconomic level, one may construct a synthetic measure of health status with the application of the MIMIC (Multiple Indicators Multiple Causes) model. The research methodology is based on the application of econometric structural equation models (SEM), which will allow one to estimate models with latent variables.

On the other hand there is the existence of significant health inequalities between particular EU regions. Efforts are being made to achieve an economic and social cohesion, of which the reduction of health inequalities between the EU regions is an important element. The application of the econometric convergence models allows us to assess the level of the public health convergence/divergence. By analysing regional data we can see an unquestionable relationship between the development of a particular region and its impact on the development of neighbouring regions. Here we can talk about a spatial dependence in the case of health status. The convergence process takes place in time and space, which proves the legitimacy of taking the spatial interactions into account. The enrichment of comparative analyses with the spatial interaction aspect will allow us to apply innovative tools in the fields of spatial econometrics and statistics. Standard econometric techniques often fail in the case of spatial dependence (spatial autocorrelation), which is commonplace in regional data sets (Anselin 2001). The modern spatial statistics methods and spatial econometrics methods, presented within the health status convergence, have a very high value both in cognitive and scientific terms and for practical applications. Using these methods in the study can provide important insights into spatial influences on health inequalities. It extends our analysis to neighbourhood effects, which is important for the economic and social environment when planning regional-level interventions to reduce health inequalities.

## Theoretical Assumptions

A review of the research conducted so far shows that the application of the MIMIC model to the assessment of health in regional terms is based on strong theoretical assumptions. This model is underpinned by the concept of the MIMIC-HSI model (*Multiple Indicator Multiple Causes – Health Status Index*) developed in the 70s of the 20<sup>th</sup> century. The main method to estimate the MIMIC model falls under the category of the structural equation models (Jöreskog and Goldberger 1975). The MIMIC-HSI models known in the literature as the Health Status Index Models are multi-equation econometric models with a latent variable (HSI). These models describe health, which is an unobservable variable, indirectly by means of health determinants and health status indicators.

In the health economic literature we can find a numerous number of applications of MIMIC methodology. The Health Status Index Model with latent variables was first applied to the research of the health condition of people living in the 50 States of North America (Robinson and Ferrara 1977). In the 80s and 90s of the 20<sup>th</sup> century the MIMIC-HSI models were extended based on new equations<sup>1</sup>:

- B Wolf, J Van der Gaag (1981): health status index for children with consideration of the socio-economic groups and their families.
- W Van den Ven, EM Hooijmans (1982): a model explaining regional differences in the frequency rate of using health care services and the quantity of the supply of such services.
- G Duru, JHP Paelinck (1991): multi-equation model describing the cost of the health care system, depending on demographic and socio-economic variables.

It should be mentioned that the above described health status models were microeconomic models, i.e. they all described the health status of particular individuals. The model that referred to the health status of the entire population was first applied by the Frenchman A Tibouti (1986) based on a sample of 94 countries from 1973. The exogenous variables of the model included:

- the share of people under 15 years of age in the entire population,
- the number of medical doctors per capita,
- the number of calories consumed daily by an inhabitant of the *i-th* country,
- and
- the share of city dwellers in the entire population.

---

<sup>1</sup> A synthetic review of the applications of the Health Status Index Model is included in the monograph by Suchecka (1998). Health Care Econometrics [*Ekonometria Ochrony Zdrowia*]. Absolutwent. Warsaw, pp. 99-115.

The following variables were applied as the endogenous variables by the author:

- average life expectancy,
- infant mortality rate, and
- unobservable health status index (HSI).

Among recent studies we can indicate many others works that employed this methodology. Leu et al. (1992) used the MIMIC model for measuring health or health impairment due to a specific disease. A MIMIC disability index was estimated for a sample of 145 adults with chronic bronchitis, expressing their self-reported disability caused by the disease on a one-dimensional scale. For instance, Giuffrida et al. (2005) analysed the complex relationships among health, income, health insurance and health care utilization to gain a better understanding of the various factors determining differences in the health of the population, especially between the poor and non-poor. They use a structural equation model in which health, wealth and access to health care are specified as latent variables.

The next part of our work is connected with the reduction of regional health inequalities with applying a beta-convergence methodology. To show that e.g. life expectancy can be modelled using the theory of economic growth, a close relationship of health with income and growth must exist (Mayer-Foulkes 2001). The main study in this field is Prestons' (1975) paper, in which he has shown that life expectancy is positively correlated with income. For instance, Barro (1991) has found the life expectancy indicator to be an important variable of the economic growth model. Arora (2001) has found cointegration between economic growth and health in a 100-125 year time series for seven advanced countries. Mayer-Foulkes (2001) has explored the topic of club-convergence. He analysed convergence clubs in cross-country life expectancy dynamics. Life expectancy was modelled in terms of physical and human capital and technology, the basic economic variables described by economic growth theories. As Jaworska (2014) indicates, on an international scale one can find research of the convergence of the public health status on the local level (Gächter and Theurl 2011). One of the recent works, which is a contribution to the literature on convergence in health status, is a paper by d'Albis et al. (2012). Authors have applied econometric tools commonly used in the economic growth literature to assess the existence convergence across high-income countries. They used both sigma- and beta-convergence methods (Jaworska 2014).

On the other hand Rey and Montouri (1999) emphasize the fact that the conventional model of convergence, when applied in a regional context is misspecified if it does not take into consideration any geographic factors. Thus, many researches use a spatial econometric approach when working with beta-convergence models (Baumont et al. 2003, Fischer and Stirböck 2006, Egger and Pfaffermayr 2006).

Our study makes two principal contributions: first, we model the level of public health (HP index) as a latent variable between observable causes and observable effects; second, we estimate and compare regional health inequalities (measured by HP index) with the application of spatial beta-convergence models.

### Econometric Framework

The MIMIC (*Multiple Indicator Multiple Causes*) model was applied to estimate the public health status in all EU regions. This research methodology is based on the application of structural equation models (SEM) that enable one to estimate models with latent variables. The general SEM (*Structural Equation Model*) model may be expressed as follows (Bentler 1990, Bollen 1989):

Structural model:

$$\eta = \mathbf{B}\eta + \mathbf{\Gamma}\xi + \zeta, \quad (1)$$

$\eta$ : vector of endogenous latent variables,

$\xi$ : vector of exogenous latent variables,

$\zeta$ : vector of remainders in the equations

$\mathbf{B}$ : matrix of structural coefficients  $\eta$ ,

$\mathbf{\Gamma}$ : matrix of structural coefficients  $\xi$ .

Measurement model:

$$\begin{aligned} y &= \mathbf{A}_y\eta + \varepsilon, \\ x &= \mathbf{A}_x\xi + \delta, \end{aligned} \quad (2)$$

$y$ : vector of observable endogenous variables,

$x$ : vector of observable exogenous variables,

$\varepsilon$ : vector of random elements w  $y$ ,

$\delta$ : vector of random elements w  $x$

$\mathbf{A}$ : factor loadings vector.

The MIMIC model known as the multiple indicator and multiple causes model (Jöreskog and Goldberger 1975) is a specific type of the general SEM model. In its original form it describes a structural relation between a single endogenous latent (unobservable) variable and several observable variables. The MIMIC model equations are as follows:

Structural model:

$$\eta_1 = \gamma x + \zeta_1, \quad (3)$$

$\eta_1$ : latent variable,

$\mathbf{x}_{qx1}$ : vector of observable exogenous variables

$\zeta_1$ : random element,

$\gamma_{qx1}$ : factor loadings vector

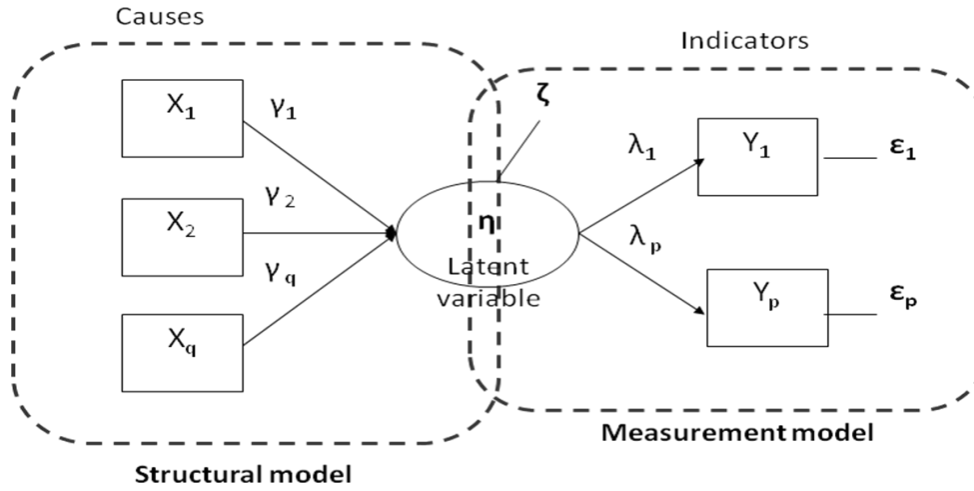
Measurement model:

$$y = \lambda\eta + \varepsilon, \quad (4)$$

$y_{px1}$ : vector of observable endogenous variables,  
 $\varepsilon_{px1}$ : vector of random elements,  
 $\lambda_{px1}$ : factor loadings vector,  
 $\phi$ : covariances between pointer variables (causes).

The MIMIC model consists of a model describing relations between a latent variable (called the structural model) and of a measurement model of observable endogenous variables. The measurement model represents the results of factor analysis, allowing one to calculate loadings of particular factors affecting the latent variable. The structural model, on the other hand, presents a path analysis, which enables one to determine the cause and effect of relations (covariances) between pointer variables (Figure 1).

**Figure 1.** Path Diagram of Simple MIMIC Model



Source: Author's estimations based on Konarski (2009).

Health status is the latent variable in the model applied in the research, whereas the socio-economic determinants of health status and the health status indicators are the describing variables.

A further part of the research is devoted to the analysis of the dynamics of health inequalities using the absolute convergence methods. In general, the unconditional  $\beta$ -convergence model can be proven with the following equation (Kusideł 2013: 47-49):

$$\ln \left( \frac{y_{it0+T}}{y_{it0}} \right) = a + b \ln(y_{it0}) + u_{it0,t0+T} \quad (5)$$

$y_{it0}$ : value of GDP per capita in the initial period,  
 $y_{it0+T}$ : value of GDP per capita in the final period,  
 $T$ : a time interval between the initial and final observation of the dependent variable.



An unconditional beta convergence process occurs when the  $b$  parameter is of negative value and it is statistically significant. A positive and statistically significant value of the parameter indicates the existence of a divergence process. The  $b$  parameter may be calculated using the following transformation:

$$b = -(1 - e^{-\beta T}). \quad (6)$$

A negative value of the parameter means that the economies (regions) with an initial lower level of the indicator achieve a higher rate of growth of this indicator, compared to the economies (regions) with an initial higher level of GDP per capita (Jaworska 2014).

In order to assess the speed at which the economies (regions) tend to reach the state of equilibrium (*steady-state*) the following formula is applied:

$$\beta = -\frac{\ln(1+b)}{T}. \quad (7)$$

With a given coefficient of convergence, we can easily calculate the half-life:

$$hl = -\frac{\ln 2}{\beta}. \quad (8)$$

The half-life coefficient reports to the time required to reduce by half the existing differences between the value of the factor and its value in the steady state.

The convergence process takes place in time and space, which needs to take the spatial interactions into account. In the empirical part of the research we also extend traditional econometric methods and tools to the field of spatial econometrics and statistics. In this context we consider two types of models with spatial interactions (Fischer and Stirböck 2006):

1. The case of Substantive Spatial Dependence (spatial lag model):

$$g = \alpha S + \rho Wg + \varepsilon, \quad (9)$$

$W$ :  $(n,n)$  spatial weight matrix (euclidean distance-based),

$g = \left( \frac{s_{ito+T}}{s_{ito}} \right)$ :  $(n,1)$ -vector of growth rate of life expectancy over the given time period,

$S$ : vector of observations on life expectancy variable in logarithms in initial year,

$\rho$ : spatial autoregressive parameter,

$\varepsilon$ : error term.

2. The case of Spatial Error Dependence (spatial error model):

$$\mathbf{g} = \alpha \mathbf{S} + \boldsymbol{\varepsilon}, \quad (10)$$

$$\boldsymbol{\varepsilon} = \lambda \mathbf{W} \boldsymbol{\varepsilon} + \boldsymbol{\mu}.$$

where,  $\mathbf{g}$ ,  $\mathbf{W}$ ,  $\mathbf{S}$  are defined as before,  $\boldsymbol{\mu}: (n,1)$  is a vector of errors,  $\lambda$  is an autoregressive parameter in the error dependence model.

### Statistical Data

The spatial scope of the research includes the regions of 28 EU member states. After elimination of outliers, i.e. regions in which the variable values significantly differed from the rest, a total of 260 EU regions were qualified for the research<sup>1</sup>. The statistical data for 2002 and 2012 were taken from the Eurostat and OECD databases<sup>2</sup>.

The potential endogenous variables included the average life expectancy at birth and the infant mortality rate.

The remaining exogenous variables include among others the percentage of people with higher education and population density. Since the eighties of the twentieth century when *The Black Report* was published in Great Britain, the socio-economic status (including the level of education) has been commonly believed to be an essential determinant of health behaviour and awareness as well as of the health status (and the related average life expectancy and mortality rate). In the literature one can find a large number of studies (e.g. Korzeniowska 2009) showing a clear dependence between a low level of education and more frequent cases of anti-health behaviour (i.e. smoking, alcohol consumption, unhealthy diet or lack of physical activity).

A population density serves as a stimulant, which means that higher values of this variable should correspond to higher values of health status<sup>3</sup>. The relationship is not

---

<sup>1</sup> The observations which were eliminated refer mainly to the so-called *special territories*, which for historical, geographical or political reasons have different relations with the government of the states within they are located (and hence with the EU), as compared to the rest of the territories. Furthermore, taking into consideration a wide diversification of the NUTS 2 level units in the EU (e.g. large Scandinavian regions versus metropolitan areas), the regions that differ significantly from the other areas in terms of demographic or economic structure were also eliminated. The results were generalized for all units with the significance level at 5%.

<sup>2</sup> <http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/>.

<sup>3</sup> In the literature there is another approach which treats population density as a destimulant. The epidemiological studies confirm the fact that overpopulation may contribute to a spread of diseases (Cassel 1971). A clear example may be diseases caused by stress - coronary heart disease, digestive system dysfunctions, neuroses and other mental disorders. A spread of new infectious diseases is less directly, but also clearly connected with overpopulation. Although a progress in medicine has eliminated many serious infections (of bacterial or parasitic origin), new ones still keep appearing. These are mainly viral infections and infections caused by antibiotic resistant strains of bacteria. The emergence of the latter is the result of a typical feedback loop: common, often excessive use of antibiotics favours a selection of resistant bacterial strains, which in turn leads to an increase of doses or administration of new medicaments, and in consequence to further increase of resistance. Other diseases connected

linear in fact. Up to a particular level, higher population density is related to better access to health care and other health-enhancing facilities.

The structure of population age also has an impact on public health. To test the hypothesis about the age we use the percentage of the people at the age of 65+ in the overall population. Following Fries (1980) we expect that a higher share of elderly people should increase their health status level. Generally the share of elderly has a negative impact on public health but according to Fries' *compression of morbidity* the age at first appearance of aging manifestations and chronic disease symptoms can increase more rapidly than life expectancy (Fries 2005). Thus, the age of population should be positively correlated with health. Table 1 shows the list of all the potential variables included in the model and Table 2 presents the statistical measurement of all the potential variables.

**Table 1. Potential Variables of the Model**

Variable	Description	Direction of relationship*
<b><i>Latent variable</i></b>		
HSI	Health Status Index	
<b><i>Exogenous observable variables</i></b>		
GDP	GDP per capita in real prices (€)	+
INC	Disposable income of households per capita (€)	+
UNM	Unemployment rate (%)	-
SEN	People at the age 65+ in overall population (%)	-/+
DEN	Population density person/km <sup>2</sup>	+
TERT	People with higher education in the economic active population (%)	+
BED	Available beds in hospitals per 100 thousand inhabitants	+
<b><i>Endogenous observable variables</i></b>		
LIFE_EXP	Average life expectancy at birth	+
MORT	Infant mortality rate	-
DEATH	Crude death rate <sup>1</sup>	-

Note: \* the "+" sign indicates that the given variable is a stimulant, whereas the "-" sign describes a destimulant.

Source: Author's estimations.

---

with overpopulation result rather from human activity. They include food poisoning, air and water pollution, civilization hazards and lifestyle.

<sup>1</sup> We can replace the crude death rates by the age-adjusted mortality rate as populations of the regions in EU differ in terms of their populations age composition. Using standardized measure was impossible because of unavailable data.

**Table 2.** *Summary Statistics of Potential Variables*

Variable	Year	Mean	Std.	Min	Max	CV (%)
IM	2002	5.3	2.9	2.0	20.6	54.7
	2012	3.7	1.6	0	11.6	43.2
LE	2002	77.7	2.6	70.1	81.9	3.3
	2012	80.3	2.5	72.9	84.7	3.1
DR	2002	1,025.4	163.8	549.2	1,825.7	15.9
	2012	979.6	185.5	457.2	1,917.8	18.9
DEN	2002	344.7	845.3	3.3	9,087	245.2
	2012	365.8	943.6	3.3	10,294.8	257.9
GDP	2002	19,956	8038	4,200	66,100	40.3
	2012	20,182	8343	3,865	72,919	41.3
TER	2002	22.6	8.1	8.1	48.1	35.8
	2012	30.5	9.1	13.2	68.9	29.8
UNM	2002	8.4	5.4	2.12	27.3	64.3
	2012	9.9	5.9	2.5	34.4	59.6
SEN	2002	16.3	2.8	8.6	25.6	17.2
	2012	18.3	3.4	8.7	27.3	18.6
INC	2002	12,030	3,886	2,365	20,015	32.3
	2012	12,463	4,181	2,400	22,346	33.5

Source: Author's estimations based on Eurostat<sup>1</sup> and OECD data<sup>2</sup>.

One of the most popular population health indicators is the average life expectancy at birth. It indicates the number of years a newborn infant would live, assuming that the probability of death determined at the time of its birth is the same throughout its life. According to statistical data, in the European Union the inhabitants of regions situated in the Mediterranean Sea basin, southern part of German (Stuttgart, Tübingen), southern Sweden (Stockholm, Östra Mellansverige, Småland med öarna) and the southern counties of England live the longest. The lowest indicator was reported in the regions of Eastern Europe – Lithuania, Latvia, regions of Romania, Bulgaria and Poland. The average life expectancy of regions stayed within the range of 70.1 - 80.9 years, in 2002 and 72.9 - 84 years, in 2012.

The infant mortality rate is the next indicator used to compare the health status of the population at the national and international level. It shows the number of deaths of infants that were under one year old in a given year per 1,000 live births in the same year. In the highly developed countries this rate remains at a relatively low level of around 2.5. The highest infant mortality rate of about 11, in 2012 was characteristic of the regions of Romania and Bulgaria.

Considering the public health, the population density is a factor of epidemiological character. Abnormally high density of the population is conducive to a spread of diseases. Generally, in Europe there is no evidence of overpopulation, however areas of high population density can be identified. These are the regions of Central and Western Europe with the largest urban agglomerations, like London, Paris or the German Ruhr region. Population density is highly varied spatially. The minimum value equalled 3.3, in 2012 and the maximum value 10,295 in the same year.

<sup>1</sup> <http://ec.europa.eu/eurostat/data/database>.

<sup>2</sup> <https://data.oecd.org/health.htm>.

The GDP per capita shows the overall economic situation in the regions. In terms of the GDP, the regions of Sweden, Netherlands, Belgium and Southern Germany constitute a group of the most developed regions of the European Union.

The next important indicator of well-being and living standards is the disposable income of households. According to OECD household disposable income can be defined as the sum of the household final consumption expenditure and savings, minus the change in net equity of households in pension funds<sup>1</sup>. It is clear that there is a positive correlation between the GDP per capita and the disposable household income per capita. From an econometric point of view a strong correlation between exogenous variables is not appropriate for the model. The MIMIC model differs from the traditional econometric models by using a confirmatory factor analysis. CFA allows us to test the hypothesis that a relationship between observed variables and their underlying latent constructs exist. The relationship is based on knowledge of the theory, postulated a priori and then tests the hypothesis statistically. According to the economic theory the GDP per capita is an economy-wide measure when the household income is a direct measure of the living standards. The correlation is logical because more developed regions characterize by higher living standards. Nevertheless, we decided to include both these indicators at the model.

### Results of the Empirical Analysis

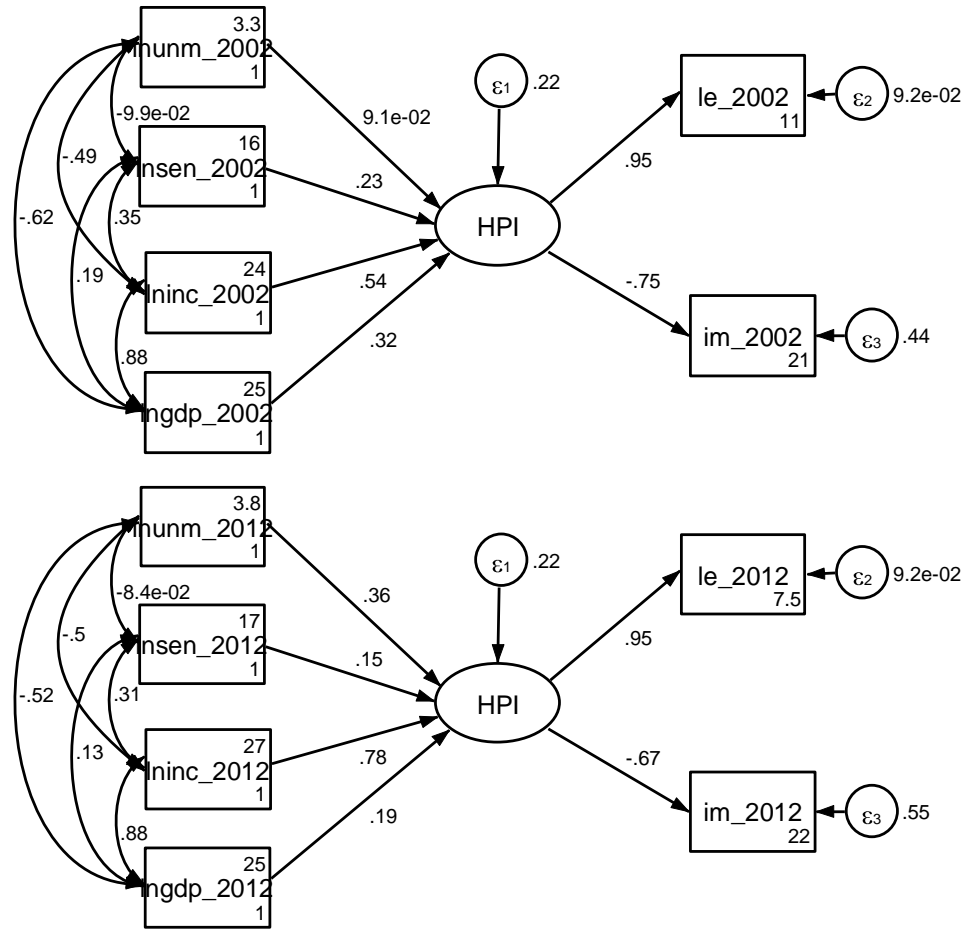
The MIMIC model was estimated with the application of the Maximum Likelihood method<sup>2</sup>. The variables selected for the proposed model also failed to fulfill the mentioned criteria. Hence, it is necessary to estimate the model also by the application of a method resistant to distribution. The results are reported on the Figure 2. These graphs show standardized coefficients of the parameters as they reflect the relative effects of the causes on the latent variable and the direction of relations between the analysed variables.

---

<sup>1</sup> <https://data.oecd.org/hha/household-disposable-income.htm>

<sup>2</sup> The Maximum Likelihood method is the most frequently applied structural equation model estimation method. The purpose of this estimation procedure is to determine the values of parameters and covariances for the covariance matrix used in the model, so that the differences between the covariance matrix used in the model and the one used in the sample would be as small as possible. However, due to quite restrictive assumptions of the ML method (multivariate normal distribution, large sample size), other methods are acceptable, e.g. those "resistant" to distribution method – ADF – *Approximately Distribution Free*.

**Figure 2.** *Standardized Coefficients of the Estimated Model in 2002 and 2012*



Source: Author's estimations.

From among the obtained estimates only four exogenous variable parameters proved to be statistically significant: GDP, INC, SEN and UNM. Other estimates show that causes of health status are also statistically significant and substantially correct. The estimated relationship between all final indicators in 2002 and 2012 is presented in Figure 2. As the above figure suggests the standardized factor loadings in the model enable us to compare the intensity of correlations between all variables. Standardized factor loadings in the measurement model should fall between 0 and 1 with higher values suggesting better indications of the observed variables for the latent variable (Pui-Wa and Qiong 2007). For instance, 1-standardized score increase in logarithm of GDP causes 0.32-standardized score increase in health (HPI). Based on the obtained results, it can be stated that the GDP per capita, the household disposable income, the percentage of people at the age 65+ and unemployment rate have a positive effect on the health (HPI), whereas the last one indicator should affect the health status negatively. Thus, in the next step of calculating the HP index we do not take into consideration the unemployment rate. Other achieved correlations are not very strong, particularly in the case of 2012, but they are statistically significant and they are in line with our expectations and economic intuition. The biggest impact on health these two years has a disposable income of households, which reflects living standards of particular regions.

We can find a significant correlation between health and its indicators: life expectancy and infant mortality. Life expectancy was chosen to be normalized. Typically the variable with the highest factor loading is chosen for this purpose (Bollen 1989). We choose to normalize the LE to a value of 1, resulting in a standardized coefficient of 0.95 (Figure 2). The indicator IM turns out to be significantly negatively related to the latent variable HPI, which is as expected.

The estimated models had to be verified in terms of adjustment and the significance level of the parameters. From among the measures of adjustment of the SEM model the most relevant ones are: chi-square statistics, mean square error of approximation, Comparative Fit Index, Tucker-Lewis Index, also known as Nonnormed Fit Index, Standardized Root Mean Square Residual and Coefficient of Determination.

The value of chi-square statistics ( $\chi^2$ ) is the most popular indicator enabling one to fit a model to the data. In accordance with the null hypothesis of the  $H_0: \Sigma = \Sigma(\Theta)$  test, the covariance matrix of the  $\Sigma$  sample is equal to the covariance matrix implied by the  $\Sigma(\Theta)$  model. It means that standardized residuals of empirical and theoretical matrices amount to 0, which indicates that the restrictions imposed by the researcher are accurate. Nevertheless, the application of the above test is justified in case of numerous samples. The susceptibility of the  $\chi^2$  test to the sample size resulted in the application of different, alternative indicators of fit.

Also the root mean square error of approximation (RMSE) is commonly applied in the evaluation of the SEM model. When calculating the RMSE, there is no comparison of the estimated model with the basic model. A lower value of the RMSE calculated based on the model, means a better fit to the data. It is assumed that RMSE values lower than 0.08 ensure a proper fit to the data.

The CFI and TLI indexes are measures of a relative model fit. They indicate an adequacy of the tested model with reference to the base model. Based on the above indexes one can determine the level of model fit to a continuum from 0 to 1, where 0 is the worst fit and 1 is the most optimal fit level (Bentler 1990).

The standardized root mean square residual (SRMR) shows differences between the covariance matrix used in the sample and the one estimated on the basis of the model. Values closer to 0 indicate a better fit (Jöreskog and Sörbom 1981: 41, 1989: 44]. The coefficient of determination plays a role similar to that of the  $R^2$  in the traditional regression model, and it means the fit of the entire model. The obtained measures of the fit are presented in Table 3.

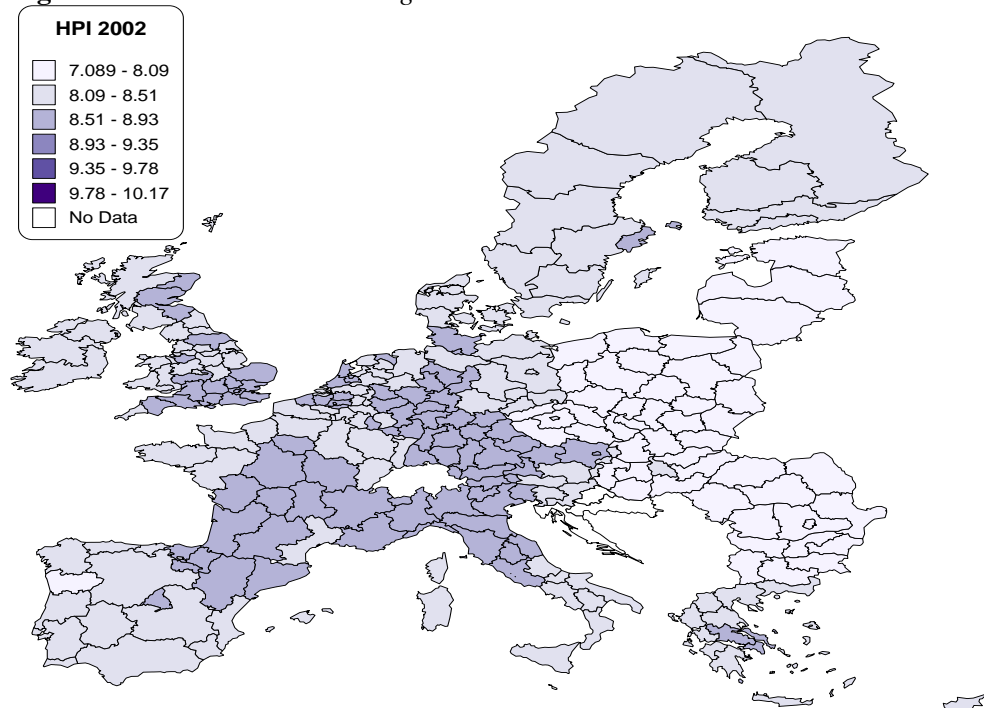
**Table 3. Model Fit Measures**

Fit statistic	2002	2012
<b>Likelihood ratio</b>		
chi-square (MS)	17.23	7.22
$p > \text{chi-square}$	0.00	0.12
chi-square (BS)	522.13	456.36
$p > \text{chi-square}$	0.00	0.000
<b>RMSEA</b>		
RMSEA	0.11	0.06
$p$	0.02	0.00
<b>Information criteria</b>		
AIC	1888.70	1685.89
BIC	1941.40	1738.60
<b>Baseline comparison</b>		
CFI	0.97	0.99
TLI	0.94	0.98
<b>Size of residuals</b>		
SRMR	0.12	0.11
CD	0.78	0.78

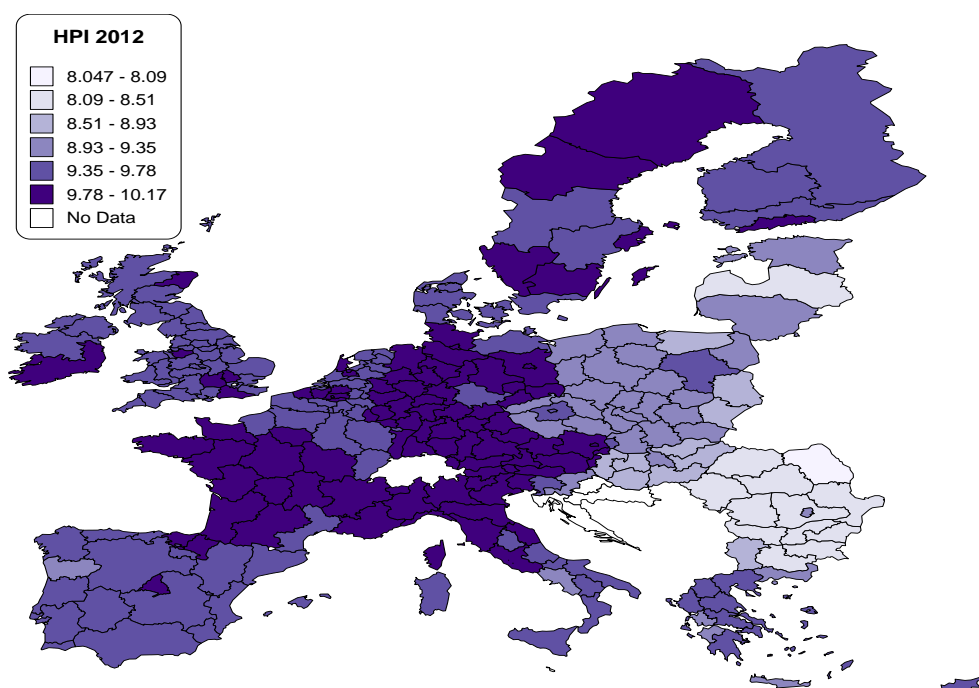
Source: Author's estimations.

In the further part of the study, by applying the values to the structural model equation (3) a health impact index was achieved. The index was calculated by applying the coefficients of the statistically significant with the expected sign causal variables to the corresponding observed variables. The values of the index for all analysed regions are presented in Figure 3. It is assumed that higher values of the index indicate higher levels of the regions in terms of public health.

**Figure 3. HP Index in the EU Regions in 2002 and 2012**







Source: Author's estimations.

The ranking of the regions is not surprising: the highly developed regions characterize the highest values of HP index. The spatial differentiation of the index is similar in the first and final year of the analysis, whereas there are higher values in 2012. The increase of HP index is connected with the growing economic development, life expectancy and the decrease of infant mortality over an analysed period.

The next step was to evaluate the dynamics of health inequalities measured with the estimated health pressure index and classic indicators, like average life expectancy. For this purpose, absolute beta-convergence models (developed based on the theory of economic growth) were applied.

Firstly we have estimated a stationary model by Ordinary Least Squares for the entire sample (see first column of Table 4). The estimated coefficient  $b$  (-0,1) indicates that the life expectancy variable in an initial year ( $\ln Y_{2002}$ ) is significant with an appropriate sign on the coefficient estimate. It shows that beta-convergence has taken place in the period of 2002-2012, which means regions with a lower initial life expectancy have obtained the largest increases in life expectancies. We have obtained similar results for HP index OLS estimation (Table 5).

**Table 4.** *Results of Beta-Convergence Model for Life Expectancy*

	OLS		ML (lag/error)	
	<i>coefficient</i>	<i>t-value</i>	<i>coefficient</i>	<i>z-value</i>
<b>a</b> (constant)	0.48***	8.54	0.39***/0.65***	5.36/7.99
<b>b</b> (lnY2002)	-0.10***	-7.97	-0.087***/-0.14***	-5.29/-7.59
$\lambda$ (spatial error)			0.63***	5.66
$\rho$ (spatial lag)			0.39***	3.03
converg. speed (annual)	1.05%		0.91%/1.6%	
half-life	66 (years)		76/46 (years)	
Diagnostic measures	$R^2=0.19$ Log likelihood=932.36 AIC criterion=1860 Moran $I=0.13$ *** LM (error)=42.89*** Robust LM (error)=45.12*** LM(lag)=13.15*** Robust LM (lag)=15.37***		$R^{*2}=0.22/0.27$ Log likelihood=936.7/943.9 AIC criterion=-1867/-1883	

Source: Author's estimations.

In the case of life expectancy estimation of the rate of convergence is above 1 percent per year and is below the standard convergence speed of 2 percent for regional economies (Fischer and Stirböck 2006). The half-distance to the steady-state is equal to 66 years. The rate of convergence of the HS index equals 1.16% with the half-life being about 60 years.

**Table 5.** *Results of Beta-Convergence Model for HP Index*

	OLS		ML (lag/error)	
	<i>coefficient</i>	<i>t-value</i>	<i>coefficient</i>	<i>z-value</i>
<b>a</b> (constant)	0.26***	6.03	0.11***/0.40***	2.72/6.49
<b>b</b> (lnY2002)	-0.06***	-2.84	-0.04***/-0.12***	-2.42/-4.28
$\lambda$ (spatial error)			0.88***	17.54
$\rho$ (spatial lag)			0.86***	16.11
converg. speed (annual)	0.62%		0.62%/1.27%	
half-life	112 (years)		112/54 (years)	
Diagnostic measures	$R^2=0.03$ Log likelihood=729.76 AIC criterion=-1455.53 Moran $I=0.39$ *** LM (error)=394.18*** Robust LM (error)=8.61*** LM(lag)=437.09*** Robust LM (lag)=51.52***		$R^{*2}=0.41/0.43$ Log likelihood=787.50/793.05 AIC criterion=-1569/-1582	

Source: Author's estimations.

The spatial character of our data has required a consideration of spatial interactions in the models. We have found the evidence of spatial dependence in the analysed phenomenon. The diagnostic measure of the Moran  $I$  statistic is

highly significant, confirming a problem with spatial autocorrelation. The presence of spatial autocorrelation can invalidate the inferential basis by OLS (Jaworska 2014). It can violate one of the basic assumptions of the OLS estimation - the assumption of uncorrelated errors (Fischer and Stirböck 2006). Thus, we need to estimate a convergence model with spatial interactions. The results of the Maximum Likelihood estimation of the spatial lagged model (9) and spatial error model (10) are displayed in the second column of Table 4 and Table 5.

The ML estimation has given quite similar results with *b*-parameters equal to -0.087 for life expectancy and to -0.04 for HP index (spatial lag model) and equal to -0.14 and -0.12 respectively for these indicators (spatial error model). The *b*-parameters are also significant and have a negative sign which is to be expected. Relative to OLS-estimates, ML-estimates has achieved a higher log likelihood indicating the better quality of the models with spatial dependency (Jaworska 2014).

It is difficult to distinguish between lag and error models because both LM tests of the lag and error are significant. Robust measures of both error and lag models are still significant but the Robust LM (error) test has the highest value, which speaks in favour of the spatial error model. In addition higher value of pseudo- $R^2$  and higher log likelihood show that the overall fit of the spatial error model is better in both cases.

## Conclusions

The first part of obtained results supports the hypothesis about the impact of socio-economic factors on the public health status. In 2002 and 2012 the public health status in the EU regions was determined mainly by economic factors: disposable income of households and economic development level measured by GDP per capita, and the age of population. We found that the MIMIC model is a good alternative way of many solutions for the presentation of a synthetic measure which is the HP index. But the main advantage of this tool is the possibility of analysing relationships between health status causes and its indicators simultaneously. It needs to be mentioned that the estimated measure is based on several selected generally available indicators. A model presentation of the regions with detailed variables was limited due to a lack of such data for all the EU regions.

In conclusions it should be stated that in the years 2002-2012 the process of convergence in health condition measured by the average life expectancy and the HP index occurred between the EU regions. It indicates a decrease in health inequalities during the analysed period, which is a very positive phenomenon from an economic point of view.

Considering the conducted analyses that confirm a high degree of variability of the analysed areas, it is difficult to determine a uniform, best practice for all regions of the European Union. A significant role is played by the coherence policy, the purpose of which is to support the activities focused

on equalizing socio-economic conditions in all EU regions. This leads to the improvement of public health in the most backward regions of the European Union.

## References

- Anselin L (2001) *Spatial Econometrics. A companion to theoretical econometrics*. Blackwell Publishing, pp. 310-330.
- Anand S, Ravallion M (1993) Human development in poor countries: on the role of private incomes and public services. *Journal of Economic Perspectives* 7(Winter): 133-150.
- Arora S (2001) Health, human productivity and long-term growth. *Journal of Economic History* 61(3): 699-749.
- Auster R, Leveson I, Sarachek D (1969) The production of health, an explanatory study. *Journal of Human Resources* 4(4): 411-436.
- Barro RJ (1991) Economic Growth in a Cross Section of Countries. *Quarterly Journal of Economics* 106(2): 407-443.
- Baumont C, Ertur C, Le Gallo J (2003) Spatial Convergence Clubs and European Regional Growth. In Fingleton B (Edn.) *European Regional Growth. Advances in Spatial Science*. Berlin: Springer-Verlag, pp. 136-151.
- Bentler PM (1990) Comparative fit indices in structural models. *Psychological Bulletin* 107(2): 238-246.
- Bollen KA (1989) *Structural Equations with Latent Variables*. New York: Wiley.
- Cassel J (1971) Health consequences of population density and crowding. In *Raport Population Growth: Consequences and Policy Implications*. Baltimore MD: John Hopkins University Press, pp. 462-478.
- d'Albis H, Esso LJ, Pifarre i Arolas H (2012) *Mortality Convergence Across High-Income Countries: An Econometric Approach*. University of Sorbonne Working Papers.
- Duru G, Paelinck Jean HP (1991) *Econometrics of Health Care*. Kluwer Publishers.
- Egger P, Pfaffermayr M (2006) Spatial convergence. *Regional Science* 85(2): 199-215.
- Fischer MM, Stirböck C (2006) Pan-European regional income growth and club-convergence. Insights from a spatial econometric perspective. *Annals of Regional Science* 40(4): 693-721.
- Folland S, Goodman AC, Stano M (2001) *The Economics of Health and Health Care*. New Jersey: Prentice Hall.
- Fries JF (1980) Aging, natural death and compression of morbidity. *The New England Journal of Medicine* 303(3): 130-136.
- Fries JF (2005) The compression of morbidity. *The Milbank Quarterly* 83(4): 801-823.
- Grossman M (1972) On the concept of health capital and the demand for health. *Journal of Political Economy* 80(2): 223-255.
- Giuffrida A, Iunes R, Savedoff W (2005) *Health and Poverty in Brazil: Estimation by Structural Equation Model with Latent Variables*. Inter-American Development Bank. No. 1/2005 IDB Publications (Working Papers).
- Gächter M, Theurl E (2011) Health status convergence at the local level: empirical evidence from Austria. *International Journal for Equity in Health* 10(34). Retrieved from <http://goo.gl/vUM49c>.

- Jaworska R (2014) Health inequalities across the European union regions: a beta-convergence approach. *Comparative Economic Research* 17(4): 71-86.
- Jöreskog KG, Goldberger AS (1975) Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association* 70(351): 631-639.
- Jöreskog KG, Sörbom D (1981) *Analysis of Linear Structural Relationship by Maximum Likelihood and Least Square Methods*. Research Raport 81-8. Uppsala. Sweden.
- Jöreskog KG, Sörbom D (1989) *Lisrel 7: A Guide to Program and Applications*. SPSS Inc: Chicago, Illinois.
- Konarski R (2009) *Modele Równań Strukturalnych [Structural Equation Modeling]*. Warsaw: Wydawnictwo Naukowe PWN.
- Korzeniowska E (2009) Wykształcenie w badaniach socjologii zdrowia-przemiany zjawiska i dylematy interpretacyjne [Education in sociology of health studies - transition phenomenon and interpretive dilemmas]. In Bąk A, Kubisz-Muła Ł (Edn.) *Metody, techniki i praktyka badań społecznych [Methods, techniques and practice of social research]*. Bielsko-Biała: Wyd. Naukowe ATH, pp. 129-147.
- Kusideł E (2013) *Konwergencja gospodarcza w Polsce i jej znaczenie w osiąganiu celów polityki spójności [Economic convergence in Poland and its importance in achieving the objectives of cohesion policy]*. Wydawnictwo UŁ.
- Leu R, Gerfin M, Spycher S (1992) The validity of the MIMIC (Multiple Indicators/Multiple Causes) health index-some empirical evidence. *Development in Health Economics and Public Policy* (1): 109-42.
- Mayer-Foulkes D (2001) *Convergence Clubs in Cross-Country Life Expectancy Dynamics*. Discussion Paper 2001/134. WIDER Development Conference, 25-26 May 2001, Helsinki.
- Or Z (2000) Determinants of health outcomes in industrialised countries: a pooled, cross-country, time-series analysis. *OECD Economic Studies* 30(2000/I): 53-77.
- Preston S (1975) The changing relation between mortality and level of economic development. *Population Studies* 29(2): 231-248.
- Pritchett L, Summers L (1996) Wealthier is healthier. *Journal of Human Resources* 31(4): 842-68.
- Pui-Wa L, Qiong W (2007) Introduction to structural equation modelling. Issues and practical considerations. *Instructional Topics in Educational Management* 26(3): 33-43.
- Rey S, Montouri B (1999) US regional income convergence: a spatial econometric perspective. *Regional Studies* 33(2): 143-156.
- Robinson PM, Ferrara MC (1977) The estimation of a model for unobservable variable with endogenous causes. In Aigner DJ, Goldberger AS (Edn.) *Latent Variables in Socio-Econometrics Models*. Amsterdam: North-Holland Publication Co.
- Schultz TP (2008) *Handbook of Development Economics*. Elsevier 4: 3406.
- Suchecka J (1998) *Ekonometria Ochrony Zdrowia [Health Care Econometrics]*. Warsaw: Absolwent.
- Tibouti A (1986) *État de santé, parités de pouvoir d'achat et croissance économique [Health, purchasing power and economic areas and believe growth]*. Ph.D thesis. Aix-Marseille II University, France.
- Van de Ven W, Hooijmans EM (1982). *The MIMIC-Health status index - What it is and what it does*. Leyden: Centre for Research Economics.

Wolf B, Van der Gaag J (1981) New Health Status Index for Children. *Health, Economics and Health Economics*. Amsterdam: North-Holland.