EKONOMIA I PRAWO. ECONOMICS AND LAW
Volume 24, Issue 3, 2025
p-ISSN 1898-2255, e-ISSN 2392-1625
www.apcz.umk.pl/EiP

Original article
received 28.04.2025; accepted 09.09.2025; published 30.09.2025
doi: 10.12775/ EiP.2025.18

Analysis of the Determinants of Healthcare Expenditure Based on Panel Data from 2000–2020

PIOTR MALINOWSKI

University of Warsaw, Faculty of Economics Sciences, ul. Długa 44/50, 00-241, Warsaw, Poland
piotrmalinowski13@gmail.com
ORCID: https://orcid.org/0009-0002-9835-6621

Abstract

Motivation: The rising costs of healthcare, fueled by aging populations, increasing life expectancy, and expanding access to medical technologies, have intensified the need to understand the underlying drivers of healthcare expenditure. While previous studies often emphasize the income elasticity of healthcare spending, many fail to incorporate broader institutional and socio-demographic factors or rely on limited cross-sectional data from a small subset of countries.

Aim: This study aims to provide a comprehensive and updated analysis of the determinants of healthcare expenditures using panel data from 155 countries over the period 2000–2020. By including a wide range of economic, demographic, and institutional variables, and employing a robust two-way fixed effects model, the research seeks to uncover key factors shaping global health spending.

Results: The final significant variables influencing healthcare expenditure include GDP per capita, the share of public healthcare expenditure in GDP, the proportion of the population aged 65 and older, urbanization rates, the number of physicians per 1,000 inhabitants and out-of-pocket healthcare expenses. Additionally, a quadratic term for the logarithm of GDP per capita was introduced to account for the non-linear relationship between income and healthcare spending. The findings suggest that both economic and demographic factors are crucial in determining healthcare expenditures, offering insights for policymakers aiming to improve the efficiency of healthcare systems.

Keywords: Healthcare expenditures, Healthcare costs, Public spending, Institutional factors, Socio-demographic factors, Panel data, Fixed effects model
JEL: I13, I14, I15, I18

1. Introduction

In light of aging societies, the development of medical technologies, and the growing public expectations for the highest possible standard of healthcare provision, the topic of healthcare expenditure and its outcomes has gained significant importance and is increasingly featured in public discourse. Existing research on the determinants of healthcare expenditure places particular emphasis on the relationship between GDP per capita and the level of healthcare spending (Younsi et al., 2016, pp. 580–601; Xu et al., 2011, pp. 1–28; Leu, 1986, pp. 41–63). To estimate this relationship, researchers often use the income elasticity of GDP per capita with respect to aggregate healthcare expenditure. This focus on income elasticity analysis frequently leads to other potentially important variables being either overlooked or insufficiently examined. Yet institutional conditions in individual countries, the structure of public sector healthcare funding, or even individual consumer preferences regarding private healthcare spending may also play a crucial role.

Many existing studies are based on cross-sectional data or analyze small samples of countries, often limited to OECD members or developing economies (Gerdtham et al., 1992, pp. 63–84; Gbesemete & Gerdtham, 1992, pp. 303–308). Only a few studies to date have analyzed large samples using panel data and included a broader set of variables–something that appears increasingly insufficient in the context of the current public debate (Younsi et al., 2016, pp. 580–601; Xu et al., 2011, pp. 1–28).

The aim of this study is to conduct an in-depth analysis of the determinants of healthcare expenditure using panel data from 155 countries for the years 2000–2020, taking into account the differences arising from varying levels of development and institutional conditions. This paper constitutes a direct continuation and extension of the author's previous work (*Malinowski, 2024, pp.1-11*), which analyzed the determinants of healthcare spending using cross-sectional data from 2018. While that study confirmed the significance of key economic and demographic variables in explaining cross-country variation, the current analysis deepens the investigation by introducing a dynamic, panel-based approach that captures both temporal and regional heterogeneity. For this purpose, the most suitable econometric model will be estimated, based on a two-way fixed effects model with robust standard error matrices. The study will also justify the drawbacks of applying alternative approaches. In line with the methodological choices made in Malinowski (2024), this study also incorporates the **squared logarithm of GDP per capita** to model the **non-linear relationship** between national income and healthcare spending - a specification that previously yielded strong explanatory power and continues to prove statistically valid in the extended panel framework.

2. Literature on the Determinants of Healthcare Expenditure

To obtain the correct form of the estimated model, the selection of independent variables remains crucial. Due to significant discrepancies in the literature regarding the proposed explanatory variables, a detailed review of the most highly regarded academic articles has been conducted. The literature review is divided into two stages. In the first stage, cross-sectional studies are analyzed, as they often cover larger samples of countries for a single point in time or focus on specific groups of countries with potentially similar individual effects. In the second stage, panel data studies are examined, which, although often considering fewer variables, include multiple time periods, allowing for the observation of temporal effects. This division enables the clear identification of independent variables specific to each study and facilitates the selection of those to be used in this paper.

Gbesemete and Gerdtham analyzed healthcare expenditure in Africa using a cross-sectional study of 30 countries. They found that 78.3% of the variance in healthcare expenditure could be explained by the combined variability of GDP per capita, the percentage of births attended by medical personnel, and international aid received (Gbesemete & Gerdtham, 1992, pp. 303–308). This study is particularly relevant when selecting variables for low-income countries, most of which are African nations.

On the opposite end of the spectrum is the study by Gerdtham et al., who conducted a cross-sectional analysis of 19 OECD countries (Gbesemete & Gerdtham, 1992, pp. 303–308). In addition to the commonly analyzed relationship between GDP per capita and healthcare expenditure, this study focuses on the share of public financing in healthcare spending and its implications. This issue is also addressed by other authors. Leu argues that an increase in the public share of healthcare financing raises total expenditure, linking it to the decrease in the cost of medical services for individual consumers (Leu, 1986, pp. 41–63). Culyer draws similar conclusions, adding that an increase in the number of entities financing healthcare services (insurance companies) also raises expenditure (Culyer, 1989, pp. 21–32), attributing this to the difficulty in centralized budget control. In later works, Gerdtham et al. conclude that providing services mainly through the public sector reduces total healthcare expenditure (Gerdtham et al., 1998, pp. 113–134). To incorporate these insights, Gerdtham et al. include several additional variables in their model: the share of total public expenditure allocated to healthcare, the percentage spent on hospital care, a dummy variable indicating the presence of copayments in ambulatory care, and another dummy variable for state-funded hospitals (Gerdtham et al., 1992, pp. 63–84).

Gerdtham et al. also introduce a simplified measure of supply-induced demand, as analyzed by Leu and Thornton & Rice (Leu, 1986, pp. 41–63; Gerdtham et al., 1992, pp. 63–84; Thornton & Rice, 2008, pp. 2873–2889). This lies at the intersection of microeconomic perspectives–viewing healthcare spending through household budgets and income-maximizing medical service providers–and the macroeconomic approach taken in this paper, which considers healthcare expenditure from a national perspective. As a proxy for supply-induced demand, they use the number of physicians per 1,000 inhabitants, unlike Leu, who used the percentage of hospital beds provided by the public sector relative to total beds in both public and private sectors (Leu, 1986, pp. 41–63; Gerdtham et al., 1992, pp. 63–84) Statistically significant variables included GDP per capita, the share of public financing, urbanization, and the dummy variable for copayments in ambulatory care. In contrast, the number of physicians per 1,000 inhabitants and the percentage of the population over 65 were statistically insignificant.

The authors understood urbanization, similarly to Leu, as a proxy for healthcare facility accessibility and the associated costs for consumers, which tend to be lower in urban centers (Leu, 1986, pp. 41–63). Their obtained negative correlation, consistent with Thornton & Rice, supports this interpretation (Leu, 1986, pp. 41–63; Thornton & Rice, 2008, pp. 2873–2889). However, it contradicts studies that interpret urbanization as a proxy for environmental pollution. It's also important to consider that the effect of urbanization may differ across country groups. In high-income countries, high urbanization likely facilitates better access to healthcare and lower costs, as mentioned above. In contrast, in low-income countries, following Gugler and Flanagan's argumentation, healthcare expenditure may increase (Gugler & Flanagan, 1978, pp. 26–46), due to the formation of cities with poor sanitation, high pollution from early industrialization, and low housing standards, which can promote the spread of epidemics.

Thornton and Rice conducted an intranational analysis of 50 U.S. states (Thornton & Rice, 2008, pp. 2873–2889). This unique study focuses on internal variation and the alignment of supply and demand in the medical market. Its exceptional perspective highlights additional aspects worth considering in our model. Uniquely, the study identifies education as a key determinant of healthcare expenditure. This finding is critical from a cost-benefit analysis standpoint. According to the study, increasing education levels reduces healthcare spending. It's worth noting that more years of formal education also enhance economic efficiency in other areas, including productivity/

However, the impact of education on healthcare expenditure is somewhat ambiguous. On one hand, longer education improves health habits and thus public health, indirectly reducing spending. On the other hand, from an individual perspective, increased awareness of health issues may raise spending. GDP per capita was also significant, but its elasticity was only 0.37, explained by the intranational scope of the analysis. Unlike Gbesemete and Gerdtham, the age structure of the population, expressed as the percentage over 65, was statistically significant (Gbesemete & Gerdtham, 1992, pp. 303–308). These are the three key takeaways from the study. Given its uniqueness, education should be considered a potentially important factor in healthcare spending. However, in this paper, we will omit education itself and instead focus on population health indicators–such as alcohol consumption, smoking, and the percentage of overweight individuals–which are likely outcomes of formal education. Therefore, the focus will be on analyzing these variables.

Regarding panel data, the two most important studies are by Xu et al. and Younsi et al., both of which attempt a comprehensive analysis of healthcare expenditure determinants over longer time spans (Younsi et al., 2016, pp. 580–601; Xu et al., 2011, pp. 1–28). These studies largely replicate the estimations of previously discussed variables, which will not be repeated here. Instead, we will focus on the innovations introduced.

Younsi et al. introduce an additional variable: out-of-pocket expenditure, referring to healthcare costs borne directly by households, such as physician copayments (earlier captured by Gerdtham et al.'s dummy variable), payments in primary care, and drug purchases (Younsi et al., 2016, pp. 580–601; Gerdtham et al., 1992, pp. 63–84). This innovative approach allows for a more comprehensive view of consumer expenditure, complementing government-funded healthcare spending. Younsi et al. include it as a percentage of total healthcare spending (Younsi et al., 2016, pp. 580–601). A key insight from this study–absent in earlier works–is that healthcare expenditure grows more slowly over time than GDP per capita. Xu et al. reach similar conclusions (Xu et al., 2011, pp. 1–28).

In addition to traditional econometric approaches, recent studies have applied methods such as fuzzy-set Qualitative Comparative Analysis (fsQCA) to examine complex interactions between determinants of healthcare spending. For instance, Nie et al. identified multiple configurations of institutional and demographic factors that jointly explain variations in health expenditures among OECD countries, emphasizing the importance of multi-causal analysis. (Nie et al, 2025, pp. 1–18)

3. Data and Methodology

In selecting the variables for the model, the author relied on the existing literature discussed above, giving priority to those factors that consistently demonstrated the highest statistical significance in previous empirical studies or appeared to be the most relevant from a theoretical perspective. This approach ensured that the final set of explanatory variables reflected both well-established determinants of healthcare expenditure and variables with strong potential explanatory power. This leads to a baseline model consisting of 12 explanatory variables and one dependent variable, defined as follows (Equation 1). This model structure was originally developed by the author in a previous study (Malinowski, 2024, pp. 1-11), which employed cross-sectional data from 153 countries for the year 2018. Through a series of modifications to the model's functional form, that study achieved a cross-sectional specification explaining 96.5% of the variation in healthcare expenditure. It is therefore justified to begin with the functional form proposed by these authors, while applying the necessary modifications required for a panel model framework, such as the inclusion of adding country-specific effect (Equation 1) and year-specific effect (Equation 2).

$$\begin{aligned}\mathbf{CHE_{it}} = \boldsymbol{\beta_0} &+ \boldsymbol{\beta_{1i}}\mathbf{lnGDP_{it}} + \boldsymbol{\beta_{2i}}\mathbf{GHE_{it}} + \boldsymbol{\beta_{3i}}\mathbf{AGE65_{it}} + \boldsymbol{\beta_{4i}}\mathbf{URB_{it}} + \boldsymbol{\beta_{5i}}\mathbf{lnPHY_{it}} \\ &+ \boldsymbol{\beta_{6i}}\mathbf{lnLEB_{it}} + \boldsymbol{\beta_{7i}}\mathbf{lnIMR_{it}} + \boldsymbol{\beta_{8i}}\mathbf{OOP} + \boldsymbol{\beta_{9i}}\mathbf{BMI_{it}} + \boldsymbol{\beta_{10i}}\mathbf{lnALC_{it}} \\ &+ \boldsymbol{\beta_{11i}}\mathbf{CIG_{it}} + \boldsymbol{\beta_{12i}}\mathbf{ODA_{it}} + \mathbf{v_i} + \boldsymbol{\varepsilon_{it}}\end{aligned} \tag{1}$$

where:
i – 1, ... , N – country;
t – 1, ... , N – time period;
v_i – country-specific effect;
ε_{it} – random error term;
CHE – Current health expenditure per capita, expressed in current international dollars, adjusted by purchasing power parity (PPP);
GDP – Gross Domestic Product per capita in current international dollars (PPP-adjusted);
GHE – Domestic general government health expenditure as a percentage of GDP;
AGE65 – Population aged 65 and over as a percentage of the total population. Based on the de facto population definition, including all residents regardless of legal status or citizenship;
URB – Urban population as a percentage of the total population, according to definitions from national statistical offices;

PHY – Number of physicians per 1,000 people, including general practitioners and specialists;
LEB – Life expectancy at birth, defined as the number of years a newborn is expected to live if prevailing mortality patterns remain constant;
IMR – Infant mortality rate, defined as the number of infants dying before age 1 per 1,000 live births in a given year;
OOP – Out-of-pocket expenditure as a percentage of total healthcare expenditure. Includes expenses such as medication purchases, doctor visits, and additional tests, borne directly by households;
BMI – Percentage of adults aged 18 and older with a body mass index (BMI) of 25 kg/m^2 or higher;
ALC – Total alcohol consumption per capita (age 15+), measured in liters of pure alcohol per calendar year, including both recorded and unrecorded alcohol, adjusted for tourist consumption;
CIG – Percentage of the population aged 15 and older currently using any tobacco product (smoked and/or smokeless), either daily or occasionally;
ODA – Official development assistance received by the country, expressed in constant 2022 US dollars. Percentage-based or PPP-adjusted international dollar data were not available, although they would have been more appropriate for this analysis;

These definitions are consistent with those employed in earlier studies, including Malinowski (2024), enabling comparability and continuity in analysis.

All data were retrieved from the World Bank database, except for BMI, sourced from the World Health Organization (WHO), and ODA, sourced from the Organization for Economic Co-operation and Development (OECD). The variable definitions are the official ones provided by the World Bank, WHO (BMI), and OECD (ODA). Throughout this paper, healthcare expenditure refers to aggregated per capita healthcare spending expressed in current international dollars, PPP-adjusted (Malinowski, 2024, pp. 1–11).

International dollars are a World Bank-defined monetary unit representing the amount of goods and services that a U.S. dollar would buy in the United States. The term 'current' refers to the purchasing power of the unit at the most recent estimation, which, for the dataset used, is the year 2022 (Malinowski, 2024, pp. 1–11).

The choice of variables and model functional form was based on the related literature cited above. It is worth noting that data for alcohol consumption (ALC) and tobacco use (CIG) were reported in five-year intervals. To address the loss of a significant number of observations due to the gaps between reporting periods, linear interpolation and extrapolation were applied to fill the missing values. A similar approach was applied to the PHY variable (physicians per 1,000 inhabitants), where some isolated data points were missing despite expected annual reporting. These adjustments should not affect the precision of coefficient estimates for these variables but may improve the estimation precision for the remaining variables.

It should also be emphasized that the term *country* refers not to the political sovereignty of a geographic region, but rather to the independent reporting of the variables of interest. Therefore, some dependent territories are listed as separate entities, and their data are not included in the 'dominant' country's dataset. The complete map of countries included in the analysis is presented in Chart 1.

In this study, a 5% significance level is used as the threshold for statistical significance. Additionally, variables potentially significant at the 10% level are marked in the output tables. However, this is merely an auxiliary guide for further model interpretation and is not used as a definitive criterion for statistical significance.

3.1. Descriptive statistics

Table 1. presents the basic descriptive statistics for the variables analyzed. Based on the table, significant differences in healthcare expenditure levels across countries can be observed. The average healthcare expenditure across the entire sample amounted to $1,162.28. The highest healthcare expenditure per capita was recorded in the United States in 2020, reaching $11,702.41, while the lowest was observed in the Democratic Republic of the Congo in 2000, amounting to $6.72.

When analyzing the share of healthcare expenditure in GDP over time, a very slight upward trend can be noted Chart 2. In the analyzed dataset, healthcare expenditure grew faster than GDP per capita; the slope coefficient of the trend line for the logarithmised values was 0.0495 for lnCHE and 0.043 for lnGDP, respectively. Thus, the findings here contrast with those of Younsi et al., who argued that healthcare expenditure grows more slowly than GDP per capita (Younsi et al., 2016, pp. 580–601).

Public healthcare expenditure represents a subgroup of total healthcare spending. The average share of public healthcare expenditure in GDP across the analyzed countries was 3.38%. It ranged from 0.06% to 22.25% of GDP. Chart 3. shows how this share evolved over time. From a temporal perspective, we can observe a steady, slight increase in the share of public healthcare expenditure throughout the years.

A key variable for the following analysis, as well as for estimating income elasticity, is GDP per capita, which averaged $16,531.37 across the countries analyzed. The highest GDP per capita was observed in Qatar in 2012 at $163,219.49, while the lowest was recorded in the Democratic Republic of the Congo in 2001 at $420.27.

The number of observations remained stable across most variables. However, significant gaps were noted for the ALC and CIG variables, resulting from their reporting at five-year intervals, as discussed earlier.

3.2. Choice of Functional Form of the Model

A key step in the analysis is the selection of the most suitable panel data model that best estimates the parameters of interest. The model selection process is based on estimating successive model variants. A logarithmic transformation was applied to all variables not expressed as percentages, marked by the prefix 'ln' in the model specification. Although many authors apply logarithmic transformations to all variables, diagnostic testing in this study indicated that such an approach would not improve the estimation quality or functional form (Younsi et al., 2016, pp. 580–601; Xu et al., 2011, pp. 1–28). It was therefore more appropriate to retain percentage variables in their original form.

The correlation between independent variables used in the model is relatively high. The estimated values are shown in Chart 4. However, considering the nature of the variables, a certain degree of correlation between them is unavoidable. An attempt was made to remove the most highly correlated variables–lnIMR, AGE65, and lnPHY–but doing so did not affect the results of functional form or heteroskedasticity tests. As such, the decision was made to retain these variables.

The estimation began with a Pooled OLS model. The parameter estimates are presented in a comparative table (Table 2.). The Breusch-Pagan LM test was used to examine the significance of individual effects across countries and whether the model reduces to a simple pooled regression. This hypothesis was rejected based on a p-value of 0.000, indicating non-zero individual effects and invalidating the use of a simple Pooled OLS model.

Panel cointegration techniques have also been employed in regional studies to model long-run equilibrium relationships between expenditure and macroeconomic indicators, highlighting the suitability of panel-based specifications (Samadi & Homaie Rad, 2013, pp. 63-68).

The focus then shifted to the Fixed Effects (FE) and Random Effects (RE) estimators. Theoretical considerations point to the Fixed Effects model being more appropriate, for two main reasons. First, the dataset includes the entire population rather than a random sample, favoring FE. Second, it is likely that country-specific effects are correlated with the explanatory variables. To finalize model selection, a Hausman test was performed, yielding a p-value of 0.000 (Hausman, 1978, pp. 1251–1271). This led to the rejection of the null hypothesis of no correlation between individual effects and explanatory variables, thus confirming the use of the Fixed Effects model. Estimates for both FE and RE models are shown in Table 2.

However, the FE model does not account for year-specific effects. To address this, a Two-Way Fixed Effects model was estimated, adding a time effect term resulting in the following equation:

$$\mathbf{CHE_{it} = \beta_0 + \beta_{1i}lnGDP_{it} + \beta_{2i}GHE_{it} + \beta_{3i}AGE65_{it} + \beta_{4i}URB_{it} + \beta_{5i}lnPHY_{it} + \beta_{6i}lnLEB_{it} + \beta_{7i}lnIMR_{it} + \beta_{8i}OOP + \beta_{9i}BMI_{it} + \beta_{10i}lnALC_{it} + \beta_{11i}CIG_{it} + \beta_{12i}ODA_{it} + v_i + \gamma_t + \varepsilon_{it}}, \tag{2}$$

where:

γ_t – year-specific effect.

Most of the time fixed effects in the Two-Way Fixed Effects model were statistically significant. Omitting them could bias the estimates of other variables. Younsi et al. and Xu et al. also included a time variable to obtain valid estimators, although they did not explicitly refer to it as a Two-Way Fixed Effects model (Younsi et al., 2016, pp. 580–601; Xu et al., 2011, pp. 1–28). Furuoka et al. explicitly used this model and found it to be the best fitting one (Furuoka et al., 2017, pp. 12–25).

To verify the functional form, the RESET test was performed in both fitted and regressor versions. The fitted version returned an F-statistic of 24.0 and p-value of 0.000, leading to the rejection of the null hypothesis of correct functional form. The regressor version gave an F-statistic of 2.58 with a p-value of 0.0762, suggesting no grounds for rejecting the null. These conflicting results imply that more weight should be given to the fitted version of the RESET test, while the regressor version is considered supplementary.

Based on the statistical significance of interactions, squared and cubic terms in the regressor version, the model was extended by adding the square of lnGDP, denoted as intlnGDPxlnGDP, as previously done by Malinowski (2024). Reapplying the RESET fitted test to this new specification yielded an F-statistic of 2.90 and p-value of 0.0553–indicating no reason to reject the null hypothesis. Thus, this extended model now has the correct functional form:

$$\mathbf{CHE_{it} = \beta_0 + \beta_{1i}lnGDP_{it} + \beta_{2i}GHE_{it} + \beta_{3i}AGE65_{it} + \beta_{4i}URB_{it} + \beta_{5i}lnPHY_{it} + \beta_{6i}lnLEB_{it} + \beta_{7i}lnIMR_{it} + \beta_{8i}OOP + \beta_{9i}BMI_{it} + \beta_{10i}lnALC_{it} + \beta_{11i}CIG_{it} + \beta_{12i}ODA_{it} + \beta_{13i}intlnGDPXlnGDP_{it} + v_i + \gamma_t + \varepsilon_{it}} \tag{3}$$

To test for heteroskedasticity, the modified Wald test was used, yielding a p-value of 0.000, indicating heteroskedasticity. The Wooldridge test for first-order autocorrelation returned an F-statistic of 100.596 and p-value of

0.000, confirming the presence of autocorrelation. Given the short time dimension and large number of entities in the panel dataset, cross-sectional dependence was not tested, as it would not affect subsequent steps.

To address heteroskedasticity and autocorrelation, the model was re-estimated using robust standard error matrices. This specification was considered the best model in this study and is referred to henceforth as the 'best model' or the 'Two-Way Fixed Effects model with correct functional form and robust standard errors.' The parameter estimates for the correct functional form using robust errors are shown in Table 3., alongside estimates from other model variants (Table 2.). With this revised specification and robust standard errors, the following variables were statistically significant: lnGDP, GHE, AGE65, URB, lnPHY, OOP, and intlnGDPxlnGDP.

To assess the impact of interpolation and extrapolation of lnPHY, lnALC, and CIG, the best model was re-estimated using only the raw data (without interpolation), resulting in just 468 observations from 142 countries. In contrast, the interpolated dataset contained 3,182 observations across 155 countries. Year-specific effects remained significant regardless of data completion. Most coefficient estimates were of similar magnitude, with differences deemed qualitatively minor. Therefore, interpolation did not appear to distort the results, while allowing for a greater number of observations, enhancing the credibility of the findings. Data imputation was also employed by other authors (Xu et al., 2011, pp. 1–28).

4. Model Estimation and Results

The parameter estimates are presented in Table 3., and their analysis provides key findings for this study. The variables that proved to be statistically significant include: GDP per capita (lnGDP), public health expenditure as a share of GDP (GHE), population age structure (AGE65), urbanization level (URB), number of physicians per 1,000 inhabitants (lnPHY), out-of-pocket expenditure share (OOP), and the square of the logarithm of GDP per capita (intlnGDPXlnGDP).

The obtained income elasticity is 1.711. However, it is not appropriate to directly compare this elasticity with values reported in other studies, as it cannot be interpreted independently of the squared logarithmic term introduced here, which is not present in other works. The square of the logarithm of income (intlnGDPXlnGDP) is negatively correlated with healthcare expenditure and has a complex interpretation.

An attempt was made to use the logarithm of GDP squared instead, which would have been easier to interpret. However, this approach was unsuccessful due to a negative RESET test result for that model. Therefore, the square of the logarithm of GDP per capita was used in this study, as it is necessary to achieve a correct functional form. This allows for the estimation of non-constant income elasticity between GDP per capita and healthcare expenditure (Wooldridge, 2013, p. 198). In other words, the level of income elasticity in the model depends on the specific value of GDP per capita–or more precisely, on the logarithm of that value. This relationship is captured in Equation (4). The income elasticity would be defined solely by the parameter if were equal to zero.

$$\%\Delta \mathbf{CHE} \approx [\boldsymbol{\beta}_1 + 2\boldsymbol{\beta}_{13}\mathbf{lnGDP}]\%\Delta \mathbf{GDP} \tag{4}$$

Public health expenditure as a share of GDP is positively correlated with healthcare expenditure. This aligns with the arguments of Leu and Culyer (Leu, 1986, pp. 41–63); (Culyer, 1989, pp. 21–32), but contradicts the findings of Gerdtham et al., who observed a negative correlation (Gerdtham et al., 1992, pp. 63–84). This finding is also supported by recent evidence from the European Union, where institutional quality and fiscal capacity have been shown to significantly influence public healthcare spending (Piscopo et al., 2024, pp. 1-19). The results support the concept that subsidizing medical services by the government reduces their effective price for consumers, increasing demand. As government funding increases, the market equilibrium shifts toward higher consumption of healthcare services, ultimately raising total expenditure.

However, the positive correlation of out-of-pocket expenditure with total healthcare spending appears contradictory to the above. Though close to zero, this may indicate the existence of unknown factors influencing how the distribution of healthcare costs between the state and households affects total expenditure. This relationship has been further explored in recent empirical work on low- and middle-income countries, such as India, where household-level data confirms the burden of direct payments and their uneven distribution across regions and income groups (Sofi & Yasmin, 2024). It may be worthwhile to more deeply analyze insurance system structures in different countries to determine whether they impact the estimated values.

Urbanization turned out to be negatively correlated with healthcare spending (Thornton & Rice, 2008, pp. 2873–2889). Thus, it can be treated as a proxy for the cost of access to healthcare facilities. Somewhat contradicting this is the positive correlation found for the number of physicians per 1,000 inhabitants. Initially, these variables were expected to complement one another and show similar directional relationships with income.

A justification for the positive correlation of the physician variable may lie in the supply-induced demand hypothesis. The underlying intuition is that physicians, as producers of medical services, may tend to maintain their income levels regardless of the physician-to-population ratio (Eastaugh, 1992, pp. 410–424); (Frech, 1996, pp. 84–101).

The share of the population over 65 years was also statistically significant and positively correlated with healthcare spending, in line with initial expectations. Surprisingly, all lifestyle-related variables (obesity, smoking, alcohol consumption) turned out to be statistically insignificant. This contradicts both the well-known negative health impacts of those behaviors and some prior research findings (Thornton & Rice, 2008, pp. 2873–2889).

The analyzed model appears to be the best specification and should serve as a reference point for inferring the determinants that influence healthcare expenditure. However, it is essential to view any study through the lens of its limitations. Data availability varies greatly across years, necessitating interpolations and approximations. Moreover, many studies use different estimations for comparable qualitative variables. Debates continue over using age brackets such as under 15 or over 65 as proxies for population age structure, and over the cut-off thresholds used to define obesity. These small differences–though individually minor–are numerous and can significantly impact final results, particularly in terms of statistical significance.

Therefore, all estimates should be treated as guidelines regarding the importance of specific determinants in shaping healthcare expenditure, and should be taken into account when designing health policies in individual countries.

5. Conclusion

The aim of this study was to identify the most important determinants of healthcare expenditure. Using panel data from 155 countries over the period 2000–2020, a series of econometric models was estimated. This enabled the selection of the Fixed Effects model with both country and year effects, along with robust standard error matrices, as the best tool for modeling healthcare spending. Introducing the square of the logarithm of GDP per capita proved crucial for achieving the correct functional form of the model. The final model demonstrated a high level of fit across the countries analyzed.

This work constitutes a further step in the investigation of healthcare expenditure determinants and contributes new insights - particularly concerning the role of income. By applying a Fixed Effects model with year-specific effects and robust error matrices, and by introducing the square of the logarithm of GDP per capita, the study enables the estimation of non-constant income elasticity with respect to healthcare spending.

The findings show that income elasticity of healthcare expenditure varies greatly depending on the type of data and variables used.

The findings of this study are largely consistent with the literature reviewed. GDP per capita and public health expenditure as a share of GDP were confirmed as key determinants, in line with Leu, Culyer, and Younsi et al., while the quadratic GDP term highlighted a non-linear relationship noted only in some prior works (Leu, 1986, pp. 41–63); (Culyer, 1989, pp. 21–32);(Younsi et al., 2016, pp. 580–601). The positive effect of population ageing supports earlier results (Thornton & Rice, 2008, pp. 2873–2889). However, lifestyle-related variables proved insignificant, and the negative link between urbanization and health spending contrasts with studies associating urbanization with higher costs, suggesting instead that greater urban access may reduce expenditure.

The obtained results provide important institutional guidelines for actions aimed at improving the efficiency of healthcare expenditure and enhancing access to medical services. They also form a solid foundation for further, more detailed analyses and for discussion on healthcare financing strategies and the diversity of health system models. Future research should explore the effectiveness of different health system models, particularly by comparing public and private systems, in order to identify the most efficient approaches to healthcare financing and organization. It is especially important to gain a better understanding of how specific health policies impact access to and quality of healthcare services, and what mechanisms can most effectively allocate resources within the healthcare sector.

References

Culyer, A. (1989). Cost containment in Europe. In Health Care Financing Review: Annual supplement (pp. 21–32). U.S. Department of Health and Human Services.

Eastaugh, S.R. (1992). Health economics, efficiency, quality, and equity. Auburn House.

Frech, H.E. (1996). Competition and monopoly in medical care. The AEI Press.

Furuoka, F., Lim, B., Kok, E., Hoque, M.Z., & Munir, Q. (2017). What are the determinants of health care expenditure? Empirical results from Asian countries. Sunway Academic Journal, 8, 12–25. https://doi.org/10.2139/ssrn.2967109

Gbesemete, K.P., & Gerdtham, U.G. (1992). Determinants of health care expenditure in Africa: A cross-sectional study. World Development, 20(2), 303-308. https://doi.org/10.1016/0305-750X(92)90108-8.

Gerdtham, U.G., Søgaard, J., Andersson, F., & Jönsson, B. (1992). An econometric analysis of health care expenditure: A cross-section study of the OECD countries. Journal of Health Economics, 11(1), 63–84. https://doi.org/10.1016/0167-6296(92)90025-V.

Gerdtham, U.-G., Jönsson, B., MacFarlan, M., & Oxley, H. (1998). The determinants of health expenditure in the OECD countries: A pooled data analysis. In Developments in Health Economics and Public Policy, 6 (pp. 113–134). Springer.

Gugler, J., & Flanagan, W.G. (1978). Urbanization and social change in West Africa. Cambridge University Press.

Hausman, J.A. (1978). Specification tests in econometrics. Econometrica, 46(6), 1251–1271. https://doi.org/10.2307/1913827.

Malinowski, P. (2024). Determinants of healthcare expenditure: A cross-sectional analysis at the country level. Journal of Education, Health and Sport, 70, 55541, 1–11. https://doi.org/10.12775/JEHS.2024.70.55541.

Nie, S., Liu, D., & Chen, S. (2025). Complex interactions in healthcare expenditure through the years: A panel data analysis using fsQCA in OECD countries with policy implications. PLOS ONE, 20(5), e0324497, 1-18. https://doi.org/10.1371/journal.pone.0324497.

Leu, R.E. (1986). Public and private health services: complementarities and conflicts. In A.J. Culyer & B. Jönsson (Eds.), Basil Blackwell (pp. 41–63). Oxford.

Piscopo, J., Groot, W., & Pavlova, M. (2024). Determinants of public health expenditure in the EU. PLOS ONE, 19(3), e0299359, 1–19. https://doi.org/10.1371/journal.pone.0299359.

Samadi, A. & Homaie Rad, E. (2013). Determinants of Healthcare Expenditure in Economic Cooperation Organization (ECO) Countries: Evidence from Panel Cointegration Tests. *International Journal of Health Policy and Management, 1*(1), 63-68. https://doi.org/10.15171/ijhpm.2013.10

Sofi, S.A., & Yasmin, E. (2024). Out-of-Pocket Health Expenditure and Associated Factors: Insights From National Health Accounts (NHA) Using Panel Data Analysis. *A journal of medical care organization, provision and financing, 61,* 469580241309903. https://doi.org/10.1177/00469580241309903.

Thornton, J.A., & Rice, J.L. (2008). Determinants of healthcare spending: A state level analysis. Applied Economics, 40(22), 2873–2889. https://doi.org/10.1080/00036840600993973.

Wooldridge, J. M. (2013). Introductory econometrics: A modern approach (5th ed.). South-Western Pub.

Xu, K., Saksena, P., & Holly, A. (2011). The determinants of health expenditure: A country-level panel data analysis(No. 26). World Health Organization. Retrieved from https://apps.who.int/iris/handle/10665/44700.

Younsi, M., Chakroun, M., & Nafla, A. (2016). Robust analysis of the determinants of healthcare expenditure growth. The International Journal of Health Planning and Management, 31(4), 580–601. https://doi.org/10.1002/hpm.2358.

Acknowledgements

Author contributions: author/authors has/have given an approval to the final version of the article. Author's total contribution to the manuscript: P. Malinowski (100%).

Funding: This research received no external funding.

Supplementary information: author/authors acknowledge following people and institution for help with the preparation of the article: University of Warsaw, Faculty of Economics Sciences.

Note: The present publication builds upon and significantly expands the previous research conducted by Malinowski (2024), offering a more comprehensive and detailed analysis of the determinants of healthcare expenditure.

Appendix

Table 1. Descriptive statistics for the entire dataset

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Number of observations	Mean	Standard deviation	Min	Max	IQR
CHE	3,219.00	1,162.27	1,521.23	6.72	11,702.41	1,360.29
GDP	3,231.00	16,531.37	19,183.74	420.27	163,219.49	20,704.04
GHE	3,221.00	3.38	2.39	0.06	22.25	3.25
AGE65	3,255.00	8.10	5.79	0.86	29.58	9.74
URB	3,255.00	55.86	22.99	8.25	100.00	37.36
PHY	2,043.00	2.02	1.48	0.01	7.55	2.70
LEB	3,255.00	69.90	9.00	41.96	84.56	12.89
IMR	3,255.00	27.13	25.26	1.70	138.60	35.90
OOP	3,219.00	33.52	19.51	0.08	86.07	29.58
ALC	771.00	5.78	4.25	0.00	19.05	7.48
CIG	1,085.00	23.70	11.32	3.50	68.50	17.20
ODA	3,255.00	512.24	1,039.76	-866.95	27,361.74	601.07
BMI	3,255.00	44.32	18.45	4.33	89.90	28.37

Source: Own elaboration based on data from the World Bank, the World Health Organization, and the Organization for Economic Co-operation and Development.

Table 2. Comparative summary of estimator values using the following models: Pooled OLS, Fixed Effects, Two-way Fixed Effects, and Two-way Fixed Effects with correct functional form and robust standard errors (Two-way Fixed Effects Robust)

	(1)	(2)	(3)	(4)
VARIABLES	PooledOLS	Fixed Effects	Two-way Fixed Effects	Two-way Fixed Effects Robust
lnGDP	0.941***	0.951***	0.896***	1.711***
	(0.015)	(0.017)	(0.018)	(0.287)

	(1)	(2)	(3)	(4)
VARIABLES	PooledOLS	Fixed Effects	Two-way Fixed Effects	Two-way Fixed Effects Robust
GHE	0.147***	0.144***	0.142***	0.143***
	(0.004)	(0.004)	(0.004)	(0.014)
AGE65	0.007***	0.011***	0.006*	0.021**
	(0.003)	(0.003)	(0.003)	(0.009)
URB	-0.004***	-0.007***	-0.009***	-0.010***
	(0.001)	(0.001)	(0.001)	(0.003)
lnPHY	0.025***	0.032***	0.040***	0.031**
	(0.007)	(0.007)	(0.007)	(0.015)
lnLEB	0.495***	0.657***	0.443***	0.142
	(0.091)	(0.096)	(0.102)	(0.274)
lnIMR	-0.009	-0.029	-0.008	-0.024
	(0.020)	(0.022)	(0.021)	(0.049)
OOP	0.004***	0.003***	0.003***	0.004**
	(0.000)	(0.000)	(0.000)	(0.002)
BMI	-0.002**	-0.002	0.001	0.003
	(0.001)	(0.001)	(0.001)	(0.005)
lnALC	0.034***	0.041***	0.048***	0.034
	(0.009)	(0.011)	(0.011)	(0.039)
CIG	-0.001*	0.001	0.003***	0.004
	(0.001)	(0.001)	(0.001)	(0.003)
ODA	0.000	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
intlnGDPXlnGDP				-0.047***
				(0.015)
Constant	-4.779***	-5.390***	-4.185***	-6.534***
	(0.447)	(0.480)	(0.508)	(1.865)
Number of observations	3,182	3,182	3,182	3,182
Number of countries	155	155	155	155
Country FE	NO	YES	YES	YES
Year FE	NO	NO	YES	YES

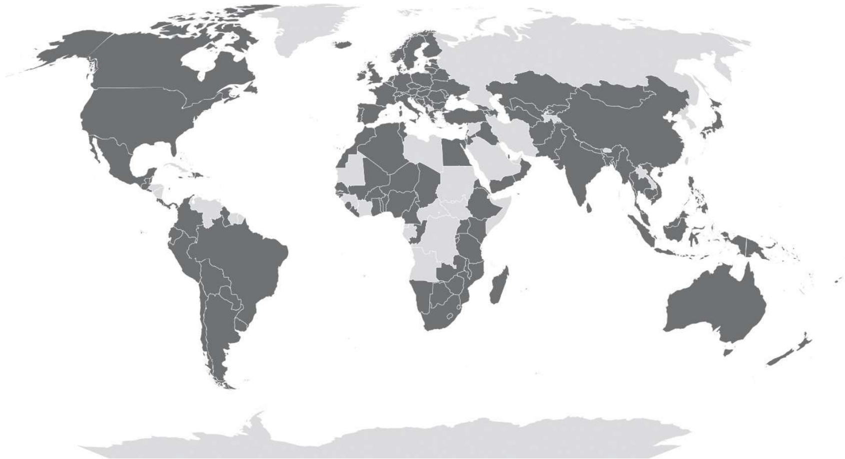
Additional information: For clarity, individual year effects have been omitted from the table.
Legend: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors are reported in parentheses according to the model used.
Source: Own elaboration based on data from the World Bank, the World Health Organization, and the organization for Economic Co-operation and Development.

Table 3. Estimation of the Two-way Fixed Effects model with correct functional form and robust standard errors

	(1)	(2)	(3)	(4)	(5)
VARIABLES	Coefficients	Robust standard error	t	p > | t |	95% Confidence Interval
lnGDP	1.711***	(0.287)	5.969	0.000	1.145 - 2.278
GHE	0.143***	(0.014)	10.273	0.000	0.116 - 0.171
AGE65	0.021**	(0.009)	2.361	0.019	0.003 - 0.038
URB	-0.010***	(0.003)	-3.116	0.002	-0.017 - -0.004
lnPHY	0.031**	(0.015)	2.009	0.046	0.001 - 0.061
lnLEB	0.142	(0.274)	0.518	0.605	-0.400 - 0.684
lnIMR	-0.024	(0.049)	-0.485	0.629	-0.120 - 0.072
OOP	0.004**	(0.002)	2.174	0.031	0.000 - 0.007
BMI	0.003	(0.005)	0.564	0.573	-0.007 - 0.012
lnALC	0.034	(0.039)	0.885	0.377	-0.042 - 0.111
CIG	0.004	(0.003)	1.388	0.167	-0.002 - 0.010
ODA	-0.000	(0.000)	-0.683	0.496	-0.000 - 0.000
intlnGDPXlnGDP	-0.047***	(0.015)	-3.103	0.002	-0.076 - -0.017
Constant	-6.534***	(1.865)	-3.504	0.001	-10.218 - -2.850
Number of observations	3,182				
Number of countries	155				
Country FE	YES				
Year FE	YES				

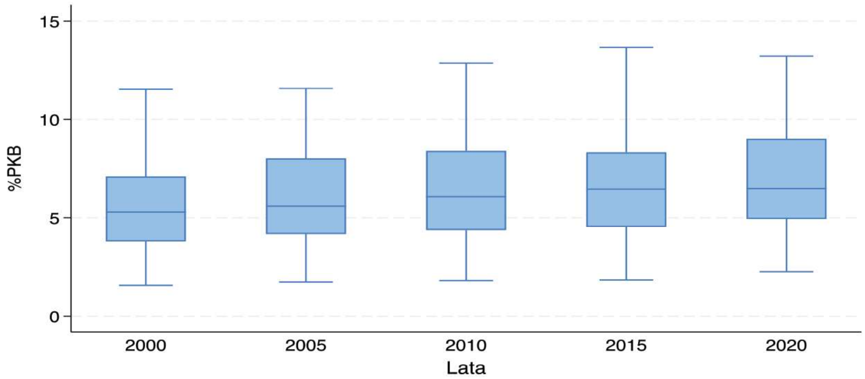
Additional information: For clarity, individual year effects have been omitted from the table.
Legend: *** p<0.01, ** p<0.05, * p<0.1.
Source: Own elaboration based on data from the World Bank, the World Health Organization, and the Organization for Economic Co-operation and Development.

Chart 1. Map of countries included in analysis



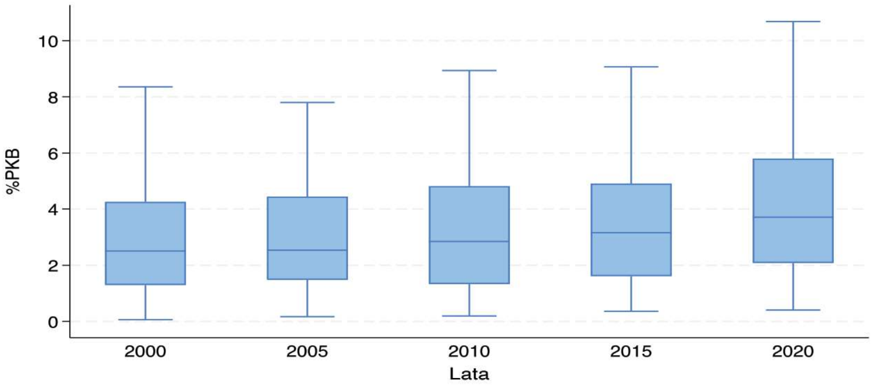
Additional information: The countries included in the analysis are marked in dark gray.
Source: Own elaboration based on World Bank data.

Chart 2. Share of healthcare expenditure in GDP for the aggregated group of all analyzed countries over the years



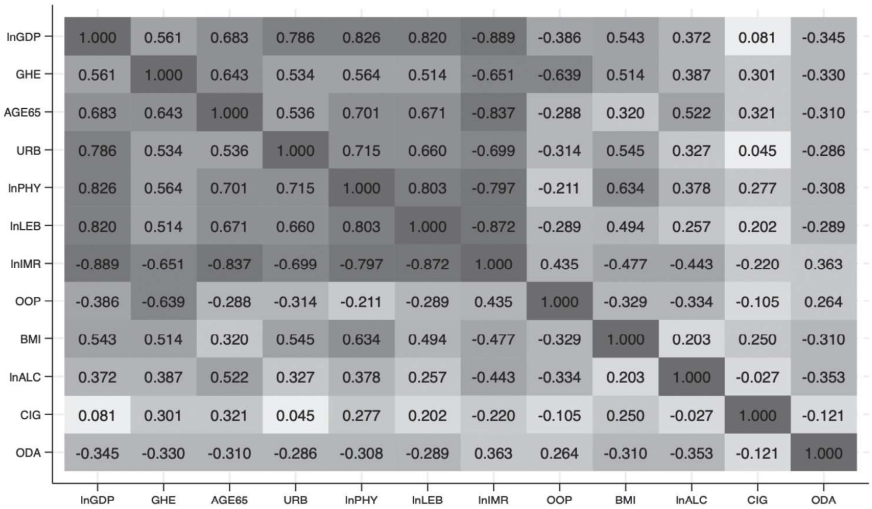
Additional information: Outliers are not included in the chart.
Source: Own elaboration based on data from the World Bank.

Chart 3. Share of public healthcare expenditure in GDP for the aggregated group of all analyzed countries over the years



Additional information: Outliers are not included in the chart.
Source: Own elaboration based on data from the World Bank.

Chart 4. Correlation matrix of independent variables



Source: Own elaboration based on data from the World Bank, the World Health Organization, and the Organization for Economic Co-operation and Development.