ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

THE ROLES OF PREVENTIVE AND CURATIVE HEALTH CARE IN ECONOMIC DEVELOPMENT

M.A. THESIS

Anna PIROGOVA

Department of Economics

Economics M.A. Program

JULY 2025

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

THE ROLES OF PREVENTIVE AND CURATIVE HEALTH CARE IN ECONOMIC DEVELOPMENT

M.A. THESIS

Anna PIROGOVA
(412231001)

Department of Economics

Economics M.A. Program

Thesis Advisor: Assoc. Prof. Dr. Osman DOĞAN

JULY 2025

ISTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

ÖNLEYİCİ VE TEDAVİ EDİCİ SAĞLIK HİZMETLERİNİN EKONOMİK KALKINMADAKİ ROLLERİ

YÜKSEK LİSANS TEZİ

Anna PIROGOVA
(412231001)

İktisat Anabilim Dalı
İktisat Tezli Yüksek Lisans Programı

Tez Danışmanı: Doç. Dr. Osman DOĞAN

TEMMUZ 2025

Anna Pirogova, an M.A. student of İTU Graduate School student ID 412231001, successfully defended the thesis entitled "THE ROLES OF PREVENTIVE AND CURATIVE HEALTH CARE IN ECONOMIC DEVELOPMENT", which she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor :	**Assoc. Prof. Dr. Osman DOĞAN** İstanbul Technical University	....
Jury Members :	**Asst. Prof. Dr. Shourjo CHAKRAVORTY** Istanbul Technical University	....
	Asst. Prof. Dr. Mahmut Sami GÜNGÖR Marmara University	....

Date of Submission : 03 July 2025
Date of Defense : 14 July 2025

To my family,

FOREWORD

Writing this thesis as part of the M.Sc. in Economics at Istanbul Technical University has been both a challenging and intellectually rewarding experience. During this time, I have learnt new analytical tools and have deepened my understanding of economic theory.

I was fortunate to have such an exceptional advisor as Prof. Dr. Osman Doğan. I would like to thank him for his superior guidance and also for his inspiring lectures. Seeing his passion for science has been a source of encouragement for me.

I also want to thank Prof. Dr. Bülent Güloglu for the “Time Series” and “Panel Data” courses, which provided me with a strong theoretical foundation for this research.

I am also grateful to my family for their support.

Special thanks go to the academic community, open-source contributors, and developers whose tools and knowledge made this research possible.

July 2025 Anna PIROGOVA

TABLE OF CONTENTS

Page

ABBREVIATIONS

AIC	**:** Akaike Information Criterion
BIC	**:** Bayesian Information Criterion
DML	**:** Double Machine Learning
FE	**:** Fixed Effects
GDP	**:** Gross Domestic Product
GMM	**:** Generalized Method of Moments
Lasso	**:** Least Absolute Shrinkage and Selection Operator
LSMA	**:** Least Squares Model Averaging
OECD	**:** Organisation for Economic Co-operation and Development
OLS	**:** Ordinary Least Squares
RE	**:** Random Effects
RMSE	**:** Root Mean Square Error
XGBoost	**:** Extreme Gradient Boosting

SYMBOLS

$\log(\boldsymbol{GDP_{it}})$	: Natural logarithm of real GDP per capita for country *i* at time *t*
$\boldsymbol{\alpha_i}$	: Country-specific fixed effects
$\boldsymbol{\delta_t}$	: Time-specific fixed effects
$\boldsymbol{HealthExp_{it}}$	: Curative or preventive health expenditure as a share of GDP
$\boldsymbol{HealthExp_{it}^2}$	: Squared term of curative or preventive health expenditure
$\boldsymbol{HealthExp_{i,t-1}}$	: Lagged health expenditure (curative or preventive)
$\boldsymbol{HealthExp_{i,t-1}^2}$	: Squared term of lagged health expenditure
$\boldsymbol{X_{it}}$	: Vector of control variables
$\boldsymbol{X_{i,t-1}}$	: Vector of lagged control variables
$\boldsymbol{\varepsilon_{it}}$	: Idiosyncratic error term
$\boldsymbol{g} \in \{\boldsymbol{low}, \boldsymbol{high}\}$	: Income group indicator
$\boldsymbol{\alpha_i^g}, \boldsymbol{\delta_t^g}$	: Country and year fixed effects estimated separately by income group
$\boldsymbol{Curative_{it}}$	: Curative health expenditure (% of GDP)
$\boldsymbol{Educ_{it}}$	: Education expenditure (% of GDP)
$\boldsymbol{Curative_{it}^2}$	: Squared term of curative health expenditure
$\boldsymbol{Curative_{it}} \times \boldsymbol{Educ_{it}}$	: Interaction term of curative and education expenditure
$\boldsymbol{Curative_{it}^2} \times \boldsymbol{Educ_{it}}$	: Interaction of squared curative and education expenditure
$\boldsymbol{Curative_{i,t-1}}, \boldsymbol{Educ_{i,t-1}}$	: Lagged values of curative and education spending

$\boldsymbol{Curative_{i,t-1} \times Educ_{i,t-1}}$	: Lagged interaction term
$\boldsymbol{Curative_{i,t-1}^{2} \times Educ_{i,t-1}}$	: Lagged interaction of squared curative and education
$\boldsymbol{Y_{it}}$	: Log of real GDP per capita
$\boldsymbol{D_{it}}$	: Health expenditure (curative or preventive)
$\boldsymbol{D_{it}^{2}}$	: Squared health expenditure
$\boldsymbol{D_{i,t-1}}$	: Lagged health expenditure
$\boldsymbol{D_{i,t-1}^{2}}$	: Squared lagged health expenditure
$\boldsymbol{g}(\cdot)$	: High-dimensional, nonparametric function of covariates
$\boldsymbol{Z_{it}}$	: Set of all regressors, including transformations
$\boldsymbol{l(Z_{it}, D_{it})}$	: Conditional expectation of Y_{it} given Z_{it}, D_{it} (outcome model)
$\boldsymbol{m(Z_{it})}$	: Conditional expectation of D_{it} given Z_{it} (treatment model)
$\boldsymbol{\bar{X}_i, \bar{X}_t}$	: Country and year means of control variables
$\boldsymbol{\bar{X}_i^{*}, \bar{X}_t^{*}}$	: Means of lagged covariates excluding first period
$\boldsymbol{X_i^{0}}$	: Initial (1998) values of control variables
$\boldsymbol{Y_i^{0}}$	: Initial (1998) value of outcome variable
$\boldsymbol{\hat{Y}_i}$	: Predicted value of outcome
$\boldsymbol{\beta}$	: Treatment effect of linear health expenditure term
$\boldsymbol{\gamma}$	: Treatment effect of squared health expenditure term
$\boldsymbol{\lambda, \lambda_1, \lambda_2}$	: Regularization parameters controlling penalty strength
$\boldsymbol{\hat{\psi}_j}$	: Penalty weight for coefficient b_j in weighted Lasso
$\boldsymbol{b_j}$	: Coefficient for the j^{th} predictor
$\boldsymbol{b \in R^{p}}$	: Coefficient vector in p-dimensional space

$\sum_{i=1}^{n}(Y_i - b^\top X_i)^2$	: Residual sum of squares (goodness of fit)
$\sum_{j=1}^{p} \mid b_j \mid$	: L1 norm (Lasso penalty)
$\sum_{j=1}^{p} b_j^2$	: L2 norm (Ridge penalty)
$\hat{g}_{RF}(Z) = \frac{1}{B}\sum_{b=1}^{B} \hat{g}_b\,(Z)$	: Random Forest prediction (mean of B trees)
$\hat{g}(Z) = \sum_{j=1}^{J} \lambda\,\hat{g}_j(Z)$	: Boosted model prediction
$\tilde{g}(Z) = \sum_{k=1}^{K} \hat{\alpha}_k\,\hat{g}_k(Z)$	: Model averaging from K learners
$\hat{\alpha}_k$	: Estimated weight of the k^{th} model in ensemble
$\tilde{Y}_{it}, \tilde{D}_{it}$	: Residualized outcome and treatment variables used in orthogonalization

LIST OF TABLES

Page

LIST OF FIGURES

Page

THE ROLES OF PREVENTIVE AND CURATIVE HEALTH IN ECONOMIC DEVELOPMENT

SUMMARY

In this thesis, we examine the impact of curative and preventive health expenditures on economic performance across Organisation for Economic Co-operation and Development (OECD) countries from 1998 to 2021. The existing literature generally agrees that health spending has a positive effect on Gross Domestic Product (GDP) growth, in line with human capital theory. However, fewer studies make a distinction between types of health investment, such as curative (treating illnesses) and preventive (avoiding illnesses). We believe that optimizing these expenditures is crucial in situations involving policy trade-offs. The aim of this thesis is to estimate the causal effect of curative and preventive spending, and to explore interactions with education investment and differentiation by income level. Additionally, we aim to apply an innovative method, such as Double Machine Learning (DML), in an area where it has not been frequently used before.

We obtained a balanced panel dataset of 29 OECD countries over the period 1998-2021 from OECD Health Statistics and the World Bank. The methodology employs two main approaches: traditional Fixed Effects (FE) models with year and country dummy variables, and a DML framework with high-dimensional predictors, utilizing five machine learning methods (Lasso, Ridge, Elastic Net, Random Forest, and Extreme Gradient Boosting (XGBoost)). Our primary variables of interest are curative and preventive health expenditures as a percentage of GDP. Controls include demographic variables (e.g., life expectancy, elderly population) and economic variables (e.g., savings, trade openness, short-term interest rate).

FE models reveal that curative expenditure may show a nonlinear, U-shaped relationship with the income level. The turning point is 2.87% of GDP after passing that point, curative health spending positively contributes to economic development. This finding is consistent across contemporaneous and lagged FE models. Regarding preventive expenditure, we were unable to identify a significant impact on the logarithm of GDP per capita. Additionally, regressions based on income level showed that the positive effects of curative spending become significant only for higher-income OECD countries, but not for lower-income ones. We also found that higher education spending moderates the effect of curative health expenditure in promoting economic development.

The DML estimates provide interesting results as well. For curative spending, DML models agree on a convex relationship. Contrary to FE models, some DML models identified an inverted U-shaped relationship between preventive health expenditure and the dependent variable; however, for lagged preventive spending models disagree.

We can derive important policy implications from these results. Firstly, curative health investment can increase the income level, but only if the spending exceeds the minimum threshold, and investment in education is also necessary. Secondly, we

suggest that countries should first invest in education and then expand curative healthcare spending. Thirdly, countries need to improve the tracking and classification of preventive expenditures to assess their effect on economic performance with certainty. Our DML results suggest that returns to prevention may be underestimated due to current data limitations. For high-income countries, policymakers should consider a higher threshold for the curative spending for it to contribute to economic development. To sum up, this thesis demonastrates that we need to explore novel approaches to drive discoveries in health economics.

ÖNLEYİCİ VE TEDAVİ EDİCİ SAĞLIK HİZMETLERİNİN EKONOMİK KALKINMADAKİ ROLLERİ

ÖZET

Bu tezde, 1998-2021 döneminde Ekonomik Kalkınma ve İşbirliği Teşkilatı (OECD) ülkelerinde tedavi edici ve önleyici sağlık harcamalarının ekonomik performans üzerindeki etkilerini inceliyoruz. Mevcut literatür, sağlık harcamalarının insan sermayesi teorisiyle uyumlu olarak Gayri Safi Yurt İçi Hasıla (GSYİH) büyümesi üzerinde olumlu bir etkisi olduğu konusunda genel bir uzlaşıya sahiptir. Ancak, daha az sayıda çalışma tedavi edici (hastalığı tedavi eden) ve önleyici (hastalığı önleyen) gibi sağlık yatırımı türleri arasında ayrım yapmaktadır. Bu harcamaların optimizasyonunun, politika öncelikleri arasında denge kurmak gereken durumlarda kritik olduğunu düşünüyoruz. Bu tezin amacı, tedavi edici ve önleyici harcamaların nedensel etkilerini tahmin etmek ve bu etkilerin eğitim yatırımı ve gelir düzeyine göre farklılaşıp farklılaşmadığını araştırmaktır. Ayrıca, bu alanda daha önce sık kullanılmamış olan Çift Makine Öğrenmesi (DML) yöntemini uygulamayı da amaçlıyoruz.

1998–2021 dönemi için 29 OECD ülkesinden oluşan dengeli panel veri seti, OECD Sağlık İstatistikleri ve Dünya Bankası kaynaklarından elde edilmiştir. Metodoloji iki temel yaklaşıma dayanmaktadır: yıl ve ülke sabit etkilerini içeren geleneksel Sabit Etkiler (FE) modelleri ve beş makine öğrenme yöntemi (Lasso, Ridge, Elastic Net, Random Forest ve Extreme Gradient Boosting (XGBoost)) kullanan, yüksek boyutlu değişkenler içeren DML çerçevesi. Ana ilgi değişkenlerimiz, GSYİH'ye oranla tedavi edici ve önleyici sağlık harcamalarıdır. Kontrol değişkenleri arasında demografik değişkenler (örneğin yaşam beklentisi, yaşlı nüfus oranı) ve ekonomik değişkenler (örneğin tasarruf oranı, ticaret açıklığı, kısa vadeli faiz oranı) yer almaktadır.

FE modelleri, tedavi edici harcamaların gelir düzeyine göre U şeklinde doğrusal olmayan bir ilişki gösterebileceğini ortaya koymaktadır. Dönüm noktası GSYİH'nin %2,87'sidir; bu eşiğin aşılmasından sonra tedavi edici sağlık harcamaları ekonomik gelişmeye olumlu katkı sağlamaktadır. Bu bulgu hem eşzamanlı hem de gecikmeli FE modellerinde tutarlıdır. Önleyici harcamalara gelince, kişi başına GSYİH'nın logaritması üzerinde anlamlı bir etki tespit edemedik. Gelir düzeyine göre yapılan regresyonlar, tedavi edici harcamaların pozitif etkilerinin yalnızca yüksek gelirli OECD ülkelerinde anlamlı hale geldiğini, düşük gelirli ülkelerde ise etkili olmadığını göstermektedir. Ayrıca, daha yüksek eğitim harcamalarının, tedavi edici sağlık harcamalarının ekonomik gelişmeyi desteklemesini güçlendirdiğini de bulduk.

DML tahminleri de dikkat çekici sonuçlar vermektedir. Tedavi edici harcamalar için DML modelleri dışbükey (U şeklinde) bir ilişki üzerinde hemfikirdir. FE modellerinden farklı olarak, bazı DML modelleri önleyici sağlık harcamaları ile bağımlı değişken arasında ters U şeklinde bir ilişki belirlemiştir; ancak gecikmeli önleyici harcama modelleri arasında fikir birliği yoktur.

Bu sonuçlardan önemli politika çıkarımları elde edilebilir. İlk olarak, tedavi edici sağlık yatırımları gelir düzeyini artırabilir, ancak yalnızca harcamalar asgari eşik düzeyini aştığında ve eğitim yatırımı da varsa. İkinci olarak, ülkelerin önce eğitime yatırım yapmaları, ardından tedavi edici sağlık harcamalarını artırmaları önerilmektedir. Üçüncü olarak, ülkelerin önleyici harcamaların ekonomik büyüme üzerindeki etkilerini güvenilir şekilde değerlendirebilmeleri için bu harcamaların izlenmesini ve sınıflandırılmasını geliştirmeleri gerekmektedir. DML bulgularımız, mevcut veri sınırlamaları nedeniyle önleyici harcamaların getirilerinin olduğundan düşük görünebileceğini göstermektedir. Yüksek gelirli ülkelerde politika yapıcıların, tedavi edici harcamaların ekonomik gelişmeye katkı sağlayabilmesi için daha yüksek bir eşik düzeyi dikkate almaları gerekebilir. Özetle, bu tez sağlık ekonomisinde yeni buluşlara ulaşmak için yenilikçi yöntemlere başvurulması gerektiğini göstermektedir.

1. INTRODUCTION

Health plays an important role in the economy, both as a fundamental component of individual well-being and as a factor that influences long-term economic growth and economic performance. Improvements in population health can raise productivity, extend working lives, and reduce the economic costs of illness. These ideas are not new: early work by Mushkin (1962) and Becker (1964) introduced the notion of health as part of human capital, while Grossman's (1972) model formalized how individuals invest in their health stock over time.

What is less often emphasized in policy and empirical work is that not all health spending serves the same function. In this thesis, we distinguish between curative spending, which targets existing diseases, and preventive spending, which aims to reduce the likelihood of future health problems. These two types of expenditure operate through different mechanisms and may therefore have different effects on economic outcomes.

Despite growing interest in the economics of health, much of the existing literature does not differentiate health expenditure by type. As a result, we know relatively little about the specific contributions of curative and preventive spending to growth. When we consider trends such as rising healthcare costs, aging societies, and tightening public budget constraints, optimizing healthcare expenses is one way to contribute to sustainable growth (Prince et al., 2015; Wang, 2018).

Some recent studies suggest that both types of health expenditure may follow nonlinear patterns in their effects, in other words showing gains at low levels but diminishing returns beyond a certain point (Wang, 2016; Wang & Wang, 2021). Others have found that these effects may depend on a country's income level or complementary factors, such as education, but this has not been systematically tested. Moreover, endogeneity remains a serious concern in most models, especially when attempting to assess causality.

In this study, we address these gaps by investigating the separate and potentially nonlinear effects of curative and preventive health spending on economic development, using panel data from 29 OECD countries between 1998 and 2021. The data were compiled from OECD Health Statistics and World Bank sources, resulting in a balanced panel of 648 observations. Our dataset includes variables such as real GDP per capita, curative and preventive health expenditures as a percentage of GDP, and various demographic and socioeconomic controls.

To estimate the relationship between health spending and economic performance, we use fixed effects regression models with both country and year effects. This approach accounts for unobserved heterogeneity and global shocks. We also introduce one-period lagged specifications in order to account for possible endogeneity. Additionally, we provide two extensions of the model, which decompose the effects by income levels and include an interaction term with education spending.

In addition to traditional econometric models, we employ an innovative Double Machine Learning (DML) framework for causal inference. DML is based on sample splitting and orthogonalization, aiming to reduce overfitting and correct for biases in high-dimensional settings. In our study, we use five machine learning methods to derive DML estimates. We include both linear and squared terms of health expenditure and consider lagged specifications.

Our research questions are:

- What are the separate effects of curative and preventive health spending on the real GDP per capita?
- Do these effects exhibit nonlinear or threshold-based behavior?
- How do these effects vary across countries with different income levels?
- Does the impact of curative health spending depend on a country's level of education investment?

From these questions, we propose the following hypotheses:

- **H1**: Curative health spending is expected to have a stronger and more direct impact on the real GDP per capita than preventive spending.
- **H2**: The relationship between health spending and real GDP per capita is nonlinear, involving threshold effects.

- **H3**: The positive effect of health spending on real GDP per capita is stronger in higher-income countries.
- **H4**: Greater education investment strengthens the positive effects of curative health spending.

Our findings indicate that curative expenditure exhibits a statistically significant U-shaped relationship with real GDP per capita, with a turning point of approximately 2.87% of GDP. In contrast, for preventive health spending, we do not observe a consistent significant effect on economic performance, likely due to low variation and issues with data quality. Extensions show that curative spending is more effective in higher-income countries, and a high level of education spending complements curative spending.

This thesis consists of four main chapters. Chapter 1, Literature Review, outlines the theoretical foundations and empirical evidence on the relationship between healthcare spending and economic growth, with particular emphasis on the distinction between curative and preventive expenditure. Chapter 2, Data and Methodology, introduces the dataset and provides descriptive statistics. In that chapter, we discuss the details of the methodological approach for FE and DML estimators. Chapter 3, Results and Discussion, presents the empirical findings across several model specifications. We examine the effects of curative and preventive spending, test for nonlinearities, and explore heterogeneity by income and education levels. After that, we compare the results from the FE and DML models. Finally, in Chapter 4, Conclusions, we summarize the main findings, discuss limitations of this study, and propose directions for future research.

2. LITERATURE REVIEW

Health is widely recognized not only as an important element of individual well-being but also as a determinant and driver for economic development. In modern literature, health expenditures are considered to be a form of “health human capital investment”. Investing in “health human capital” is associated with increasing the number and quality of the workforce, reducing time lost due to sickness, and therefore promoting higher economic performance in the long run.

However, as we mentioned earlier, we cannot view healthcare as a homogeneous concept. While empirical research has established a positive relationship between total healthcare expenditure and economic growth, the separate roles of preventive and curative healthcare spending in economic performance have received relatively less attention. This distinction is particularly relevant for high-income economies, which have achieved longer life expectancy, but whose ageing population is struggling with chronic diseases that incur high societal costs (Prince et al., 2015). We may note that this tendency has intensified the debates on optimal allocation of scarce healthcare resources.

In this context, OECD countries can be an important group for this study. On the one hand, that group of countries has relatively well-developed healthcare systems, with high curative healthcare spending, ranging from 3.5% to 6.5% of GDP (OECD data, author’s calculations). On the other hand, preventive healthcare allocation remains at a low level, within a 2-4% band of total healthcare spending for most OECD countries (Gmeinder et al., 2017), despite emerging evidence of the cost-effectiveness of preventive strategies.

Recent research explores the potential of preventive healthcare as a driver for economic development. There are several which argue that spending on prevention may have non-linear effects on the economic growth. Until the optimal level there are increasing returns, and after passing a certain point, the returns decline. Moreover, as societies age and healthcare costs escalate, the strategic role of prevention in reducing the future burden on curative services becomes more prominent.

This literature review aims to provide a structured overview of the existing theoretical and empirical research on the relationship between healthcare expenditure — particularly its preventive and curative components — and economic development.

Special attention is given to studies focused on OECD countries, as well as research utilizing panel data methodologies similar to those applied in this thesis.

The literature review will have the following outline. The next section focuses on the theoretical background, describing the role of health for economic performance and introducing the concepts of "health human capital" and "health as an investment." The subsequent sections present empirical findings from the literature on curative and preventive healthcare spending. The final section compares our thesis to the previous literature.

2.1 Theoretical Background: Health, Human Capital, and Economic Performance

We can define economic growth as a process in which there is an increase in the economy's productive capacity. Conventionally, it is measured by the increase in real GDP. In neoclassical economic theory, economic growth is based on three factors: the capital stock, the labor stock, and productivity. Productivity depends on technological progress and is given as exogenous. However, in the latter studies, researchers introduced a function to explain technological progress induced by investment in human capital, which was initially formulated by Gary Becker in 1964.

Becker's formulation of a theory of human capital formation states that investing in human capital increases an individual's productivity. Becker indicated health as one of the components of the human capital stock; however, the latter focus was mainly on education. In this sense, it is also worth mentioning Mushkin's work (1962), in which he first explicitly introduced health as a form of human capital investment. In this paper, health is viewed not only as a consumption good but also as a productive factor, because healthier individuals work more efficiently and for longer periods, thereby affecting national output. This idea built the foundation for what later became known as the health-led growth hypothesis — the view that improvements in population health can stimulate economic growth through several mechanisms: higher labor productivity, increased labor force participation, and lower absenteeism due to illness.

The next important formulation was done by Michael Grossman (1972), who constructed a model for the demand for health capital itself. According to Grossman's model, the consumer initially owns a stock of health, and that stock depreciates over time. However, using investments such as medical care, diet, and exercise, the individual can maintain and increase the stock of health. The health stock produces the output of “healthy time,” which is available for both labor market participation and leisure. To this day, Grossman’s model stands as the main model for the demand for health.

Connecting the theory to measurable outcomes, real GDP per capita is often used in empirical studies as a proxy for economic development (von Haldenwang & Ivanyna, 2010; Sharma, 2018). According to the World Bank Glossary, GDP per capita serves as an indicator of economic output per individual and indirectly reflects income levels across the population. Sustained economic growth leads to higher average incomes and is closely associated with poverty reduction (World Bank, 2025). In a sense, real GDP per capita also reflects aspects of levels of economic development, even though we cannot say that this measure is a complete measure of economic welfare (Eurostat, 2025).

In this context, while health expenditure in general is accepted as beneficial for economic development and economic growth, the debate on how to allocate resources between preventive and curative healthcare is frequently discussed in the literature. As we have discussed in the introduction, these two expenditure categories have distinct functions and different effects on the economy and cannot be replaced interchangeably. Let us define those notions.

According to the System of Health Accounts 2011, the aim of curative care is to relieve symptoms of illness or injury, to reduce the severity of it, or to protect against complications. Curative care consists of inpatient, outpatient, and day curative care. Under preventive care, we understand any measure that focuses on avoiding or reducing the number or the severity of injuries, diseases, and their complications. Preventive care, in turn, comprises information, education, and counseling programs, immunization programs, early disease detection programs, healthy condition monitoring programs, epidemiological surveillance, and risk and disease control programs, as well as disaster and emergency response programs (OECD et al., 2011).

There is an ongoing debate in the literature regarding preventive vs. curative healthcare, namely, which one is “better” and whether prevention is better than cure. Some researchers claim that preventive measures result in savings, while others point to limitations and notice controversy regarding the potential cost-effectiveness of preventive medicine compared to more “traditional” curative healthcare.

In that regard, let us first present some arguments, which are from a somewhat sceptical side. When considering the question of “Can preventive medicine save money?” supporters of the sceptical side tend to reformulate the question and look from a new angle: “How can we measure the effectiveness of preventive medicine?”. While the efficiency of curative medicine is somewhat easier to assess (we can observe the result of the treatment directly), regarding the preventive interventions, the situation is far more complex, because often we need to observe “absences” that were prevented. For example, if preventive measures are successful, say, a specific disease of a certain population, which could have happened 10-20 years later, was prevented, it is very hard to definitely attribute it to the preventive measures that were taken. Over a long span, many factors may have affected the (non)appearance of the disease, and hence, the effect is hard to observe.

Moreover, proponents of that side are also concerned about whether generalizations can be made in the same manner as those in curative medicine. For instance, Russell (1986) expressed the idea that it is challenging to make generalizations about preventive interventions, moreover due to the context-sensitive nature of public health, firstly we should estimate economic benefits and cost-effectiveness potential of preventive measure on the individual level and in various settings and longer timeframes, before we can make general conclusions. Approximately 20 years later, Russell, in the presentation for the Economics of Prevention Workshop, concludes: decision makers have to make the decision based on the individual merit of the intervention rather than generalizations. Additionally, some reviews suggest that the effect of preventive measures is not uniform: some studies have shown that around 20% of preventive interventions reduce medical spending. Meaning that in some cases, preventive measures are indeed cost-saving; however, in some other cases, it is just adding to healthcare costs (Cohen et al., 2008).

The third argument relates to the general potential inefficiency of healthcare, underuse of services, including some specific prevention measures, and therefore, room for growth for the healthcare delivery (Cohen et al., 2008).

Keeping in mind certain limitations and potential issues of preventive medicine which are rightfully pointed out by critics, the author of this thesis suggests to shift the focus from comparing and contrasting the efficiency and/or cost-effectiveness of both medicines to rather integrative approach and focus on the question: “given limited resources what can be the best optimal allocation between curative and preventive medicine”? In other words, if we accept that both types of medical expenditures have the right to exist and benefit the health and hence the human capital, therefore economic prosperity, can we find such an allocation that will produce the desired effect?

The works of researcher Fuhmei Wang (2016, 2018, 2021) made significant contributions to this area. Wang’s research suggests that both types of spending have non-linear effects on economic performance: increasing returns at low levels of investment, but diminishing returns as spending rises beyond certain thresholds. Moreover, Wang’s models indicate that there is an optimal allocation of healthcare spending between prevention and cure that maximizes economic returns. Importantly, these models suggest that underinvestment in prevention — a common feature of many OECD healthcare systems — may result in missed opportunities to improve productivity and reduce long-term healthcare costs. At the same time, if there is an over-investment in prevention, going beyond the optimal level, gains are marginal. Even inefficiencies may arise, especially if preventive efforts detect minor health issues that do not significantly affect productivity.

2.2 Empirical Evidence on Healthcare Expenditures and Economic Performance

Beylik et al. (2022) considered OECD countries' panel data and examined the association between health expenditure indicators and economic growth. The study found that increasing the health expenditure relative to GDP by 1% results in an increase in GDP of 0.09% and of 0.06% in the amount of per capita income.

Piabuo and Tieguhong (2017) focused on several African countries and observed a 0.3-0.4 unit increase in GDP per capita for those countries that achieved the Abuja target (15% of government expenditure allocation to healthcare). Researchers also noted that there exists a long-run relationship between healthcare expenses and economic outcomes.

Baldacci et al. (2004) utilized a panel of 120 developing countries and found that healthcare expenditure had a significant positive effect on growth during a specific timeframe.

Atilgan et al. (2024) also provided empirical evidence from OECD countries that supports the health-led growth hypothesis. The authors also highlighted that countries possessing an efficient health financing system may realize larger growth from health expenditure than others.

Hence, the literature shows that different samples of both developed and developing countries experienced a significant positive effect of healthcare expenditure on economic growth. This finding is consistent with the theory of health as human capital introduced earlier. However, higher levels of spending could yield only marginal gains or even inefficiencies, especially if preventive efforts detect minor health issues that do not significantly affect productivity.

In 2024 study F. Wang explores how prevention influences human capital loss from catastrophic disease, for example cancer, end-stage kidney disease and major psychiatric diseases for the case of Taiwan. The author emphasizes that these kinds of diseases not only impose a high financial burden on the healthcare system but also result in the loss of human capital. Fuhmei Wang concludes that investing in preventive services helps prevent those diseases at early stages, and therefore saves labor productivity and the lifetime employment duration (Wang et al., 2024).

In another study, the research is focused on the role of prevention for aging societies, especially for OECD countries. The authors show that prevention decreases the prevalence rate of illnesses leading to sustainable growth in productivity (Wang & Wang, 2021).

In the research by Maciosek et al. (2010), the authors concluded that increasing the use of preventive services to a 90 percent level would result in savings of approximately 0.2 percent of U.S. personal health care spending.

Wang and Wang (2021) note that there exists an optimal share of health expenditure on prevention to GDP, which is approximately 1.175%, maintaining a rate of illness prevalence at 6.13%. At the same time, excessive expenditure on preventive healthcare will result in detecting new, mild-severity chronic diseases, resulting in longer illness durations.

Hu and Wang (2024) have obtained fascinating results, finding that when household consumption is below a certain threshold, the effect of health expenditure is economic performance is negative; however, after passing the threshold, the effect turns significantly positive.

Let us compare our study with the literature. First of all, like Wang (2018), Atilgan et al. (2024), and Beylik et al. (2022), our focus is on OECD countries and we use the official OECD Health Statistics. Variables and macroeconomic controls are also similar to those established in previous literature. However, our study has some differences in terms of data structure; we aimed for a balanced panel from 1998 to 2021. Therefore, we selected only countries with a minimum of missing data and used imputation techniques to fill in a few missing observations. This contrasts with earlier studies, such as Wang (2018), which use unbalanced panels and shorter time periods.

Our study also has differences in methodology: this thesis uses FE panel regressions and a novel DML approach rather than Generalized Method of Moments (GMM) estimators. Additionally, we included square terms, similar to those in Wang (2018), and our research also includes a more extended analysis. Firstly, we tested for heterogeneous effects by income. Secondly, we include interaction terms between curative health expenditure and public education spending, while in previous studies, education is typically treated as an additional control variable.

3. DATA AND METHODOLOGY

3.1 Data Description and Summary Statistics

Data from 29 OECD countries were used; the panel datasets of dependent and explanatory variables were observed over the period from 1998 to 2021. These countries include: Australia, Austria, Belgium, Canada, Czechia, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Italy, Japan, Luxembourg, Mexico, Netherlands, Norway, Poland, Portugal, Slovakia, South Korea, Spain, Sweden, Switzerland, United Kingdom and the United States. The data is collected from OECD statistics and The Global Economy data provider, which compiles data from multiple sources, including The World Bank, The United Nations, and UNESCO.

Table 3.1 summarizes the descriptive statistics and shows the means, standard deviations, the minimum, and the maximum. According to the summary statistics, the study is based on a large range of values; therefore, the estimated samples are more comprehensive. Table 3.2 provides country-level summary statistics.

There is significant cross-country heterogeneity in all dimensions, which highlights the diversity of economic conditions and health system priorities among OECD members. The average real GDP per capita in 2010 USD is approximately $38,789.60, with a standard deviation of $21,778.28. In our sample, Luxembourg has the highest per capita real GDP (US$95,358.79), while Mexico has the lowest per capita GDP (US$9,247.33).

Health outcomes also vary widely. Life expectancy ranges from a low of 73.83 years (Mexico) to a high of 82.76 years (Japan), with an average of 79.49 years. The proportion of the population aged 65 and above — a key demographic determinant of healthcare demand — ranges from under 5% to nearly 30%, with a mean of 16.20%. We observe that curative healthcare expenditure averages 4.98% of GDP, with a wide range from 1.72% to 11.90%. The highest and lowest percentage of curative health expenditures are observed in the USA (10.2%) and Mexico, respectively (2.96%).

In contrast, preventive spending is much lower overall, with a sample average of just 0.26%, ranging from effectively 0% to a maximum of 1.27%. For preventive health expenditures, we observe that Canada's share is the highest (0.59%), while Greece has the lowest (0.12%).

Table 3.1 : Summary Statistics.

Variable	Mean	Std. Dev.	Minimum	Maximum
Real GDP per capita 2010 US dollars	38789.60	21778.28	5627.55	127117.10
Trade as percent of GDP (%)	90.94	56.64	18.13	393.14
Population size in millions	41.44	62.02	0.27	332.05
Population ages 65 and above percent of total (%)	16.20	3.78	4.78	29.79
Life expectancy in years	79.49	2.80	70.13	84.56
Public spending on education percent of GDP (%)	5.18	1.19	2.89	8.61
Savings as percent of GDP (%)	6.32	6.18	-15.71	27.92
Short term interest rate (%)	2.83	3.51	-0.80	27.10
Total health care as percent of GDP (%)	8.81	2.22	3.60	18.60
Curative health care as percent of GDP (%)	4.98	1.46	1.72	11.90
Preventive health care as percent of GDP (%)	0.26	0.15	0.00	1.27

Table 3.2 : Country-Level Summary Statistics.

Country	Real GDP per capita 2010 US dollars	Trade as percent of GDP (%)	Population size in millions	Population ages 65 and above percent of total (%)	Life expectancy in years	Public spending on education as a percentage of GDP (%)	Savings as percent of GDP (%)	Short-term interest rate (%)	Total health care as a percent of GDP (%)	Curative health care as percent of GDP (%)	Preventive health care as percent of GDP (%)
Australia	44839.19	41.91	21.97	13.89	81.37	5.13	5.54	3.88	8.77	5.89	0.19
Austria	41541.85	96.91	8.41	17.27	80.19	5.47	7.87	1.64	10.03	5.96	0.25
Belgium	39303.89	150.27	10.86	17.63	79.87	6.17	7.83	1.64	9.78	5.22	0.18
Canada	40119.55	67.89	33.94	14.56	80.8	4.93	5.64	2.33	10.05	4.87	0.6
Czechia	16455.45	125.83	10.41	16.13	77.05	4.2	3.59	2.68	6.99	3.85	0.2
Denmark	53681.48	95.12	5.55	16.93	79.02	7.73	9.41	1.89	9.63	5.61	0.26
Finland	41090.37	74.77	5.35	17.97	79.87	6.3	7.56	1.64	8.62	4.89	0.34
France	36742.06	56.96	64.51	17.82	81.12	5.53	5.98	1.64	10.8	5.72	0.25
Germany	38871.18	76.05	82.11	19.63	79.71	4.6	8.18	1.64	10.81	5.54	0.37
Greece	22142.35	60.2	10.89	19.37	79.99	3.8	-5.94	2.52	8.44	5.26	0.11
Hungary	12048.05	147.6	9.99	16.9	73.9	4.78	4.49	6.44	7.18	3.7	0.29
Iceland	56812.82	82.25	0.32	12.63	81.7	7.26	-0.5	7.62	8.7	5.1	0.23
Italy	32499.59	52.92	58.84	20.63	81.61	4.28	2.88	1.7	8.45	4.44	0.3
Japan	37671.62	28.8	127.27	23.5	82.76	3.38	4.72	0.24	9.1	5.49	0.26
Luxembourg	96735.18	306.89	0.52	14.09	80.35	4.22	13.09	1.64	5.98	3.43	0.13
Mexico	9990.65	60.97	111.74	6.28	73.83	4.65	3.68	8.71	5.28	2.82	0.18
Netherlands	46036.47	135.32	16.61	16	80.22	5.13	10.73	1.64	9.58	4.63	0.43
Norway	75433.68	69.85	4.89	15.68	80.89	7.1	19.78	3.19	9.06	4.54	0.22
Poland	10946.62	82.08	38.12	14.34	75.81	4.97	5.76	5.97	6.13	3.72	0.14

Table 3.2 (continued) : Country-Level Summary Statistics.

Country	Real GDP per capita 2010 US dollars	Trade as percent of GDP (%)	Population size in millions	Population ages 65 and above percent of total (%)	Life expectancy in years	Public spending on education as a percentage of GDP (%)	Savings as percent of GDP (%)	Short-term interest rate (%)	Total health care as a percent of GDP (%)	Curative health care as percent of GDP (%)	Preventive health care as percent of GDP (%)
Portugal	20080.03	72.32	10.41	18.87	79.15	5.03	-0.63	1.67	9.35	5.75	0.2
Slovakia	14562.64	155.31	5.4	13.14	75.11	4.04	1.32	3.59	6.69	3.15	0.15
South Korea	22721.08	77.73	49.42	10.97	79.75	3.91	16.59	3.88	5.84	3.55	0.19
Spain	27667.06	58.96	44.82	17.62	81.43	4.37	7.36	1.67	8.42	5	0.23
Sweden	47630.12	83.32	9.45	18.4	81.27	7	11.66	1.59	9.31	5.43	0.3
Switzerland	71454.81	112.04	7.84	16.82	81.97	4.95	10.96	0.46	10.32	5.23	0.24
USA	49677.63	26.38	306.48	13.55	77.82	5.96	2.62	2.14	15.35	10.2	0.53
United Kingdom	40563.73	56.84	62.65	16.8	79.78	5.04	0.59	2.73	9.16	5.52	0.26

Figure 3.1 and Figure 3.2 display the average trend in health spending over time. We can observe a clear upward trajectory in curative expenditures, especially in the periods following global crises such as the 2008–2009 financial collapse and the COVID-19 pandemic (2020–2021). Preventive spending, in contrast, remained relatively flat for most of the time span, with a notable but slightly delayed increase post-2019 — we believe that it is likely reflecting expanded public health campaigns and epidemic preparedness measures in response to the global pandemic.

As we can see from these temporal patterns, most OECD countries are increasingly prioritizing curative healthcare expenditure. While curative health services are necessary and often benefit from strong institutional and interest group support, preventive measures — despite potentially being more cost-effective in the long term — often suffer from underinvestment.

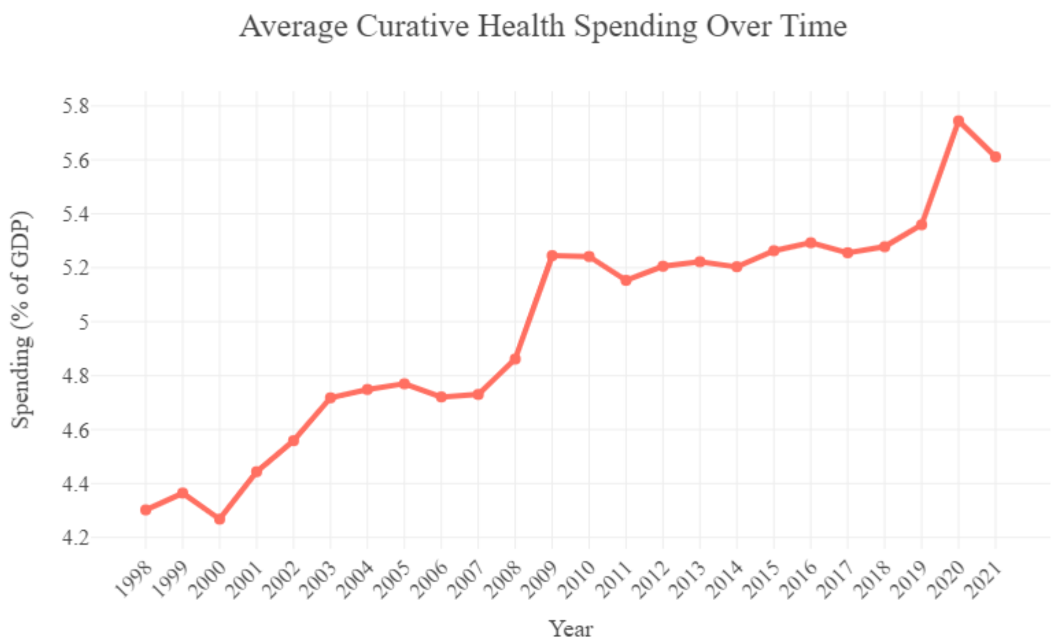

Figure 3.1 : Average Curative Health Expenditure Over Time.

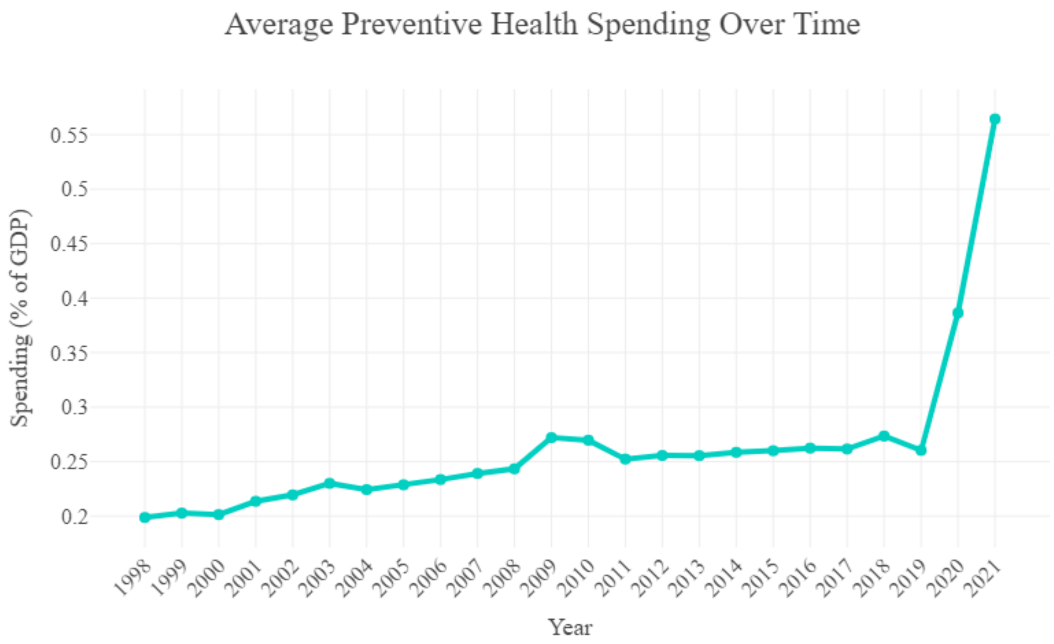

Figure 3.2 : Average Preventive Health Expenditure Over Time.

Figure 3.3 and Figure 3.4 further explore the distribution of curative and preventive spending across countries on a yearly basis. In the boxplots in these figures, we showed not only central tendencies (medians and means) but also the degree of variation and the presence of outliers.

In Figure 3.3, curative spending shows consistent median growth and a moderate increase over time, suggesting that while most countries gradually increased their curative spending, the increase in a few countries was much higher.

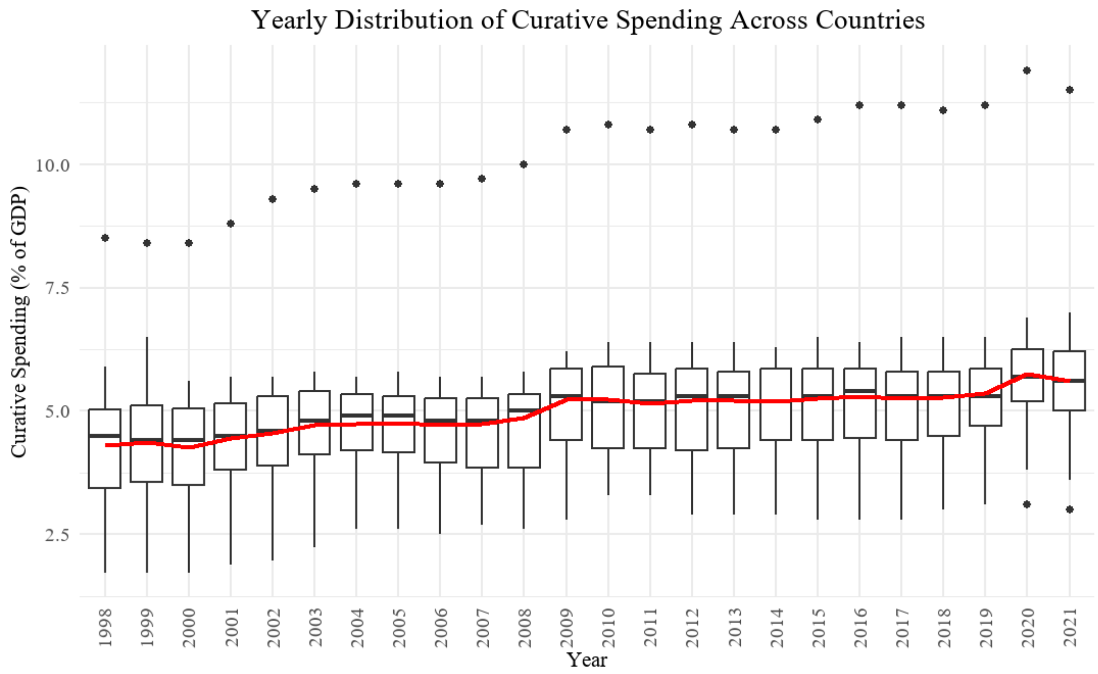

Figure 3.3 : Box-plot Distribution of Curative Expenditure Over Time.

Figure 3.4 shows a more uneven trend for preventive health spending. The distribution remains tightly clustered around low values for much of the time period, with visible increases in variability — and some countries sharply diverging from the rest — particularly after 2020.

These graphical patterns indicate that while preventive spending is slowly gaining attention, it still occupies a relatively marginal role in most public budgets. The sharp post-2020 increase in some countries may indicate a turning point, although this remains to be seen in the context of long-term structural reforms.

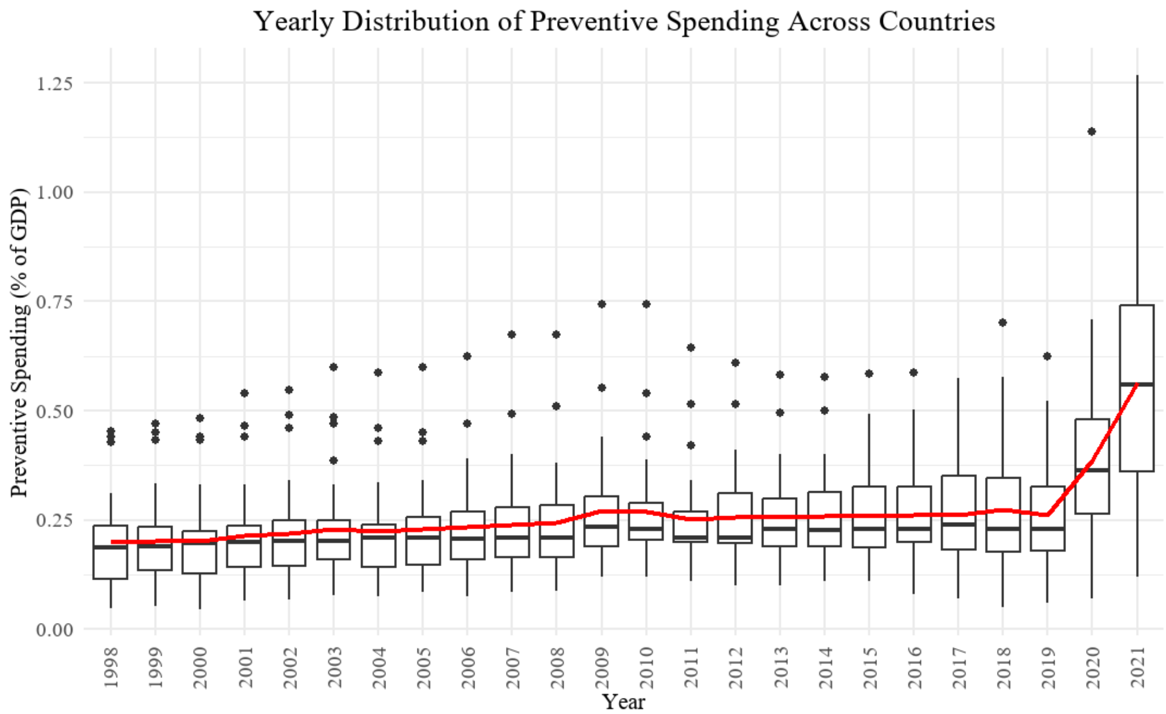

Figure 3.4 : Box-plot Distribution of Preventive Expenditure Over Time.

3.2 Fixed Effects Regression Approach

To analyze the relationship between healthcare expenditure and economic development, we employ a panel data regression framework using FE estimators. This approach allows us to control for time-invariant unobserved heterogeneity across countries, such as institutional quality, geography, and long-standing structural differences that might confound the observed relationship. As noted by Wooldridge (2010) and Baltagi (2005), the fixed effects estimator is particularly useful in eliminating country-specific bias in panel datasets.

We estimate two-way fixed effects models, which include both country fixed effects, α_i and year fixed effects, δ_t. Country effects control for unobservable but constant characteristics of each country, while time effects capture global trends or shocks affecting all countries in a given year (such as pandemics or international economic crises). This specification is commonly used in cross-country studies of public health and economic outcomes, as seen in works by Baldacci et al. (2004) and Moreno-Serra and Smith (2011). GDP per capita, expressed in constant 2010 USD, serves as a commonly accepted proxy for economic development in cross-country panel analyses.

Referring to Barro and Sala-i-Martin (2004), the control variables X_{it} are composed of the trade balance as percent of GDP to reflect openness, population size in millions for presenting human capital quantity, population ages 65 and above percent of total to examine the impact on economic performance, life expectancy in years for researching the impact of the population's health status on economic prosperity, public spending on education percent of GDP for presenting the quality of human capital, savings as percent of GDP to reflect the accumulation of private capital, short-term interest rate to reflect the price of physical capital and the interest rate.

The influence that each variable has on economic performance has been of particular interest in the literature. This research also addresses the aforementioned concerns.

Our baseline regression model is defined as:

$$\log(\mathrm{GDP}_{it}) = \alpha_i + \delta_t + \beta_1 \cdot \mathrm{HealthExp}_{it} + \beta_2 \cdot \mathrm{HealthExp}_{it}^2 + \gamma' \mathrm{X}_{it} + \varepsilon_{it}, \quad (3.1)$$

where:

- $\log(\mathrm{GDP}_{it})$: the natural logarithm of real GDP per capita for the country i at time t,
- α_i: country-specific fixed effects,
- δ_t: captures time-specific effects,
- $\mathrm{HealthExp}_{it}$: denotes either curative or preventive health spending as a share of GDP,
- $\mathrm{HealthExp}_{it}^2$: squared term of either curative or preventive health spending,
- X_{it}: vector of controls,
- ε_{it}: the idiosyncratic error term.

The inclusion of both a linear and a squared term for healthcare expenditure follows previous studies that explore nonlinearities in the growth-healthcare literature (Suhrcke et al., 2006). It allows us to detect potential threshold effects or diminishing returns.

To account for the dynamic nature of healthcare effects and reduce concerns about reverse causality, we also consider lagged fixed effects models, where all explanatory variables (including healthcare and controls) are introduced with a one-period lag:

$$\log(\mathrm{GDP}_{it}) = \alpha_i + \delta_t + \beta_1 \cdot \mathrm{HealthExp}_{i,t-1} + \beta_2 \cdot \mathrm{HealthExp}_{i,t-1}^2 + \gamma' \mathrm{X}_{i,t-1} + \varepsilon_{it}, \tag{3.2}$$

where:

- $\log(\mathrm{GDP}_{it})$: the natural logarithm of real GDP per capita for the country i at time t,
- α_i: country-specific fixed effects,
- δ_t: captures time-specific effects,
- $\mathrm{HealthExp}_{i,t-1}$: lagged curative or preventive health spending,
- $\mathrm{HealthExp}_{i,t-1}^2$: lagged squared term of either curative or preventive health spending,
- $\mathrm{X}_{i,t-1}$: vector of lagged control variables,
- ε_{it}: the idiosyncratic error term.

This dynamic specification is consistent with the empirical strategies employed by Aghion et al. (2010), who argue that improvements in population health may not generate immediate productivity gains, but rather manifest over time. Introducing lagged variables also mitigates simultaneity bias and provides a stronger causal interpretation of the relationship.

To explore heterogeneous effects, we follow a strategy similar to Baldacci et al. (2008) by conducting separate regressions for subsamples. Specifically, we divide countries into two samples based on average income level and estimate fixed effects models for each subgroup.

Our contemporaneous equation is:

$$\log(GDP_{it}) = \alpha_i^g + \delta_t^g + \beta_1^g \cdot Curative_{it} + \beta_2^g \cdot Curative_{it}^2 + \gamma^{\mathrm{g}\prime} X_{it} + \varepsilon_{it}, \tag{3.3}$$

where:

- $g \in$ {low,high}: the income group,
- α_i^g and δ_t^g: country and year fixed effects estimated separately for each group,
- X_{it}: vector of the control variables,
- ε_{it}: the idiosyncratic error term.

In this approach, we assume that the parameters of the model may not be the same across high- and low-income countries. The lagged equation version takes the following form:

$$\log(GDP_{it}) = \alpha_i^g + \delta_t^g + \beta_1^g \cdot Curative_{i,t-1} + \beta_2^g \cdot Curative_{i,t-1}^2 + \gamma^{g\prime} X_{i,t-1} + \varepsilon_{it}, \tag{3.4}$$

where:

- $g \in$ {low,high} indicates the income group,
- α_i^g and δ_t^g: country and year fixed effects estimated separately for each group,
- $X_{i,t-1}$: the vector of lagged control variables,
- ε_{it}: the idiosyncratic error term.

We will also explore the education decomposition by adding continuous interaction terms. Our goal is to establish whether the effect of curative spending depends on the level of education spending. The contemporaneous equation with the interaction terms takes the following form:

$$\log(GDP_{it}) = \alpha_i + \delta_t + \beta_1 \cdot Curative_{it} + \beta_2 \cdot Curative_{it}^2 + \beta_3 \cdot (Curative_{it} \times Educ_{it}) + \beta_4 \cdot (Curative_{it}^2 \times Educ_{it}) + \gamma' X_{it} + \varepsilon_{it}, \tag{3.5}$$

where:

- α_i, δ_t: country and year fixed effects,
- $Curative_{it}$, $Educ_{it}$: current levels of curative health and education spending (% of GDP),

- $Curative_{it}^2$: squared term for curative health expenditure,
- $Curative_{it} \times Educ_{it}, Curative_{it}^2 \times Educ_{it}$: continuous interaction terms,
- X_{it}: vector of control variables,
- ε_{it} : the idiosyncratic error term.

The lagged version with the interaction terms take the following form:

$$\begin{aligned}\log(GDP_{it}) = \alpha_i + \delta_t + \beta_1 \cdot Curative_{i,t-1} + \beta_2 \cdot Curative_{i,t-1}^2 \\ + \beta_3 \cdot \left(Curative_{i,t-1} \times Educ_{i,t-1}\right) + \beta_4 \\ \cdot \left(Curative_{i,t-1}^2 \times Educ_{i,t-1}\right) + \gamma' X_{i,t-1} + \varepsilon_{it},\end{aligned} \quad (3.6)$$

where:

- α_i, δ_t: country and year fixed effects,
- $Curative_{i,t-1,}$, $Educ_{i,t-1}$: lagged levels of curative health and education spending,
- $Curative_{it}^2$: squared term for curative health expenditure,
- $Curative_{i,t-1} \times Educ_{i,t-1}, Curative_{i,t-1}^2 \times Educ_{i,t-1}$: continuous interaction terms,
- $X_{i,t-1}$: vector of lagged control variables,
- ε_{it}: the idiosyncratic error term.

We do not consider these extensions involving differential effects across income groups or interaction effects through education spending for preventive health expenditures because our results do not provide statistical evidence supporting them.

3.3 Double Machine Learning Approach

The DML estimator is a method designed to estimate causal parameters in models where high-dimensional or complex machine learning methods are used to control for confounding. It relies on the principle of Neyman orthogonality to construct estimators that are robust to small errors in the estimation of nuisance parameters (the outcome or treatment models). DML employs sample-splitting and cross-fitting techniques to reduce overfitting and ensure valid inference, even when using flexible ML models.

This makes it particularly suitable for estimating treatment effects in partially linear models or other semiparametric settings, where the main interest lies in a low-dimensional target parameter while high-dimensional data are available (Chernozhukov et al., 2024; Chernozhukov et al., 2018).

We estimate the causal effect of curative health expenditures—both linear and quadratic—on economic performance using a Double Machine Learning (DML) framework. The outcome model is a partially linear regression of the form:

$$Y_{it} = \beta D_{it} + \gamma D_{it}^2 + g\left(X_{it}, X_{it}^2, X_{it}^3, X_{it}X_{jt}, \bar{X}_i, \bar{X}_t\right) + \varepsilon_{it}, \tag{3.7}$$

where:

- Y_{it}: the log of real GDP per capita,
- D_{it}: health expenditure (curative or preventive),
- D_{it}^2: square of health expenditure,
- $g(\cdot)$: high-dimensional, nonparametric function of covariates and their nonlinear transformations and interactions,
- $\bar{X}_i$ and $\bar{X}_t$: country and year means are used to account for unobserved country and time heterogeneity,
- ε_{it}: the idiosyncratic error term.

In these approach we have two tretament variables (or variables of interest) D_{it} and D_{it}^2. Our aim is to estimate β and γ after controlling the effects of covariates on the dependent variable through g. To that end, the DML approach proceeds in two sequential steps, each targeting one of the treatment variables while treating the other as part of the control variables.

Let Z_{it} be the set of all regressors, including all their technical transformations. In the first step, to estimate the linear effect of D_{it} on the dependent variable, we absorb D_{it}^2 into the control function and use DML to estimate the residual-on-residual regression:

1. First stage (nuisance function estimation):

- Estimate

$$\hat{\ell}(Z_{it}, D_{it}^2) = E[Y_{it} \mid D_{it}^2, Z_{it}], \tag{3.8}$$

- Estimate

$$\widehat{m}(Z_{it}, D_{it}^2) = E[D_{it} \mid D_{it}^2, Z_{it}], \tag{3.9}$$

where Z_{it} includes covariates and their transformations, interactions, and fixed effects.

2. Second stage (target parameter regression):

- Form residuals:

$$\tilde{Y}_{it} = Y_{it} - \hat{\ell}(Z_{it}, D_{it}^2), \tag{3.10}$$

$$\tilde{D}_{it} = D_{it} - \widehat{m}(Z_{it}, D_{it}^2) \tag{3.11}$$

- Estimate β via the following Ordinary Least Squares (OLS) regression model:

$$\tilde{Y}_{it} = \beta \tilde{D}_{it} + \varepsilon_{it} \tag{3.12}$$

In the second step, to estimate the nonlinear effect γ, we now absorb D_{it} into the control variables and treat D_{it}^2 as the treatment variable:

1. First stage (nuisance function estimation):

- Estimate

$$\hat{\ell}(Z_{it}, D_{it}) = E[Y_{it} \mid D_{it}, Z_{it}] \tag{3.13}$$

- Estimate

$$\widehat{m}(Z_{it}, D_{it}) = E[D_{it}^2 \mid D_{it}, Z_{it}] \tag{3.14}$$

2. Second stage (target parameter regression):

- Form residuals:

$$\tilde{Y}_{it} = Y_{it} - \hat{\ell}(Z_{it}, D_{it}), \tilde{D}_{it}^2 = D_{it}^2 - \widehat{m}(Z_{it}, D_{it}) \tag{3.15}$$

- Estimate γ via the following OLS regression model:

$$\tilde{Y}_{it} = \gamma \tilde{D}_{it}^2 + \varepsilon_{it} \tag{3.16}$$

In both steps, we use the cross-fitting method described in Chernozhukov et al. (2024) to formulate the residualized variables $\tilde{Y}_{it}$, $\tilde{D}_{it}$ and $\tilde{D}_{it}^2$. Also, in all final OLS regression models, we use the heteroskedasticity-robust standard errors.

To test our model's sensitivity to different specifications, we will use two additional options, in addition to our base option (No Initial X values), for the contemporaneous version. These versions are:

1. The technical regressors include all initial regressors (Includes both X_i^0 and Y_i^0):

$$Y_{it} = \beta D_{i,t} + \gamma D_{i,t}^2 + g\left(X_{i,t}, X_{i,t}^2, X_{i,t}^3, X_{i,t}X_{j,t}, \bar{X}_i, \bar{X}_t, X_i^0, Y_i^0\right) + \varepsilon_{it}, \tag{3.17}$$

where:

- Y_{it}: log of real GDP per capita for the country i at time t,
- $D_{i,t}$: health expenditure (curative or preventive),
- $D_{i,t}^2$: square of health expenditure,
- $g(\cdot)$: high-dimensional, nonparametric function of covariates and their nonlinear transformations and interactions,
- $\bar{X}_i$ and $\bar{X}_t$: country and year means, used to account for unobserved heterogeneity,
- X_i^0: initial values (from 1998) of control variables,
- Y_i^0: initial value (from 1998) of the outcome variable,
- ε_{it}: the idiosyncratic error term.

2. The technical regressors do not include the initial outcome variable (Only X_i^0, excludes Y_i^0):

$$Y_{it} = \beta D_{i,t} + \gamma D_{i,t}^2 + g\left(X_{i,t}, X_{i,t}^2, X_{i,t}^3, X_{i,t}X_{j,t}, \bar{X}_i, \bar{X}_t, X_i^0\right) + \varepsilon_{it}, \tag{3.18}$$

where:

- all terms are defined as above,
- X_i^0 is included but Y_i^0 is excluded.

As in the case of the fixed effects regression approach, we also consider a version with lagged regressors to address potential endogeneity issues. These lagged specifications of our partially linear model take the following form:

$$\begin{aligned} Y_{it} &= \beta D_{i,t-1} + \gamma D_{i,t-1}^2 \\ &+ g\left(X_{i,t-1}, X_{i,t-1}^2, X_{i,t-1}^3, X_{i,t-1}X_{j,t-1}, \bar{X}_i^{\,*}, \bar{X}_t^{\,*}\right) + \varepsilon_{it}, \end{aligned} \tag{3.19}$$

where:

- Y_{it}: log of real GDP per capita at time t,
- $D_{i,t-1}$: lagged health expenditure (curative or preventive),
- $D_{i,t-1}^2$: square of lagged health expenditure,
- $g(\cdot)$: high-dimensional, nonparametric function of lagged covariates, including their squared, cubic, and interaction terms,
- $\bar{X}_i^*, \bar{X}_t^*$ country- and time-specific means of lagged covariates, excluding the first time period.

In our implementation of the DML estimator, we employed five machine learning models to estimate nuisance functions $l(Z_{it}, D_{it})$and $m(Z_{it}, D_{it})$ under a cross-fitting scheme with K=5 folds. These learners are Lasso, Ridge, Elastic Net Regression, Random Forest, and XGboost (boosted trees).

Let us describe each learner in more detail. Our first learner is Lasso, also known as the Least Absolute Shrinkage and Selection Operator method. The key characteristic of this regularized regression is the penalty assigned to the size of the coefficients. The penalty is necessary for the model to focus on the most important predictors and ignore the ones that are less useful.

The Lasso estimator solves the following optimization problem:

$$\min_{b \in R^p} \sum_{i=1}^{n} (Y_i - b'X_i)^2 + \lambda \sum_{j=1}^{p} \mid b_j \mid \hat{\psi}_j \tag{3.20}$$

In the equation above, the first term is the standard squared error, which measures the fit to our data. The second term is the penalty, which “punishes” large values of the coefficients b_j. And lastly, we have λ and $\hat{\psi}_j$ standing for a parameter of the strength of the “punishment” and penalty weights, respectively.

As follows from the Lasso specification, it is able to set some coefficients exactly to zero, which is equivalent to variable selection. At the same time, we should keep in mind the limitations of Lasso. In the presence of a high correlation between explanatory variables, it may drop one of them; however, it can hide a real relationship that we may have. Also, Lasso tends to underestimate the size of the coefficients, a phenomenon known as shrinkage bias.

In our setting, we implemented Lasso using the cv.glmnet frunction from glmnet package with default settings for cross-validation to select the regularization parameter (lambda.min).

Ridge regression is also a regularization technique that is suitable for high-dimensional datasets. In the case of Ridge regression, the optimization problem is:

$$\hat{\beta}(\lambda) = \underset{b \in R^p}{\operatorname{argmin}} \sum_{i=1}^{n} (Y_i - b'X_i)^2 + \lambda \sum_{j=1}^{p} b_j^2 \tag{3.21}$$

In the equation above, we can also observe that the first term is similar to Lasso, but the second term is quite different. λ is still the regularization parameter (the shrinkage parameter) controlling the penalty for large coefficients. Unlike Lasso, Ridge will not set the coefficients exactly to zero, but will allow them to be very small in size. Therefore, Ridge does not perform the variable selection. When the number of predictors p is close to or exceeds the number of observations n, this approach helps to prevent overfitting.

At the same time, when the underlying data-generating process is believed to be sparse, the selection of Ridge regression may be undesirable. Since only a small number of predictors will really affect the outcome, setting many coefficients to non-zero values can lead to unnecessary noise and unstable estimates.

In our application, we fitted Ridge via the cv.glmnet function with alpha = 0, applying an L2 penalty, for choosing the shrinkage parameter. Thus, we again resort to the cross-validation method for choosing the optimal lambda value.

The next learner is Elastic Net, which is a useful hybrid of Lasso and Ridge. The Elastic learner is defined as

$$\hat{\beta}(\lambda_1, \lambda_2) = \underset{b \in R^p}{\operatorname{argmin}} \sum_{i=1}^{n} (Y_i - b'X_i)^2 + \lambda_1 \sum_{j=1}^{p} b_j^2 + \lambda_2 \sum_{j=1}^{p} | b_j | \tag{3.22}$$

The first penaly term in the equation comes from Lasso, and the second one comes from Ridge. So, λ_1 controls the Ridge penalty, and λ_2 is controlling the Lasso penalty. Naturally, Elastic Net possesses strengths from two approaches and is able to perform in both sparse and dense settings. The Elastic Net is beneficial when there is a concern about multicollinearity.

We fitted Elastic Net via the cv.glmnet function from glmnet package with a fixed mixing parameter alpha = 0.5. The regularization strength (lambda.min) was selected through the cross-validation method.

Note that Lasso, Ridge and Elsatic Net are linear learners. Next, we move from linear learners to non-parametric ones (non-linear learners). The first non-linear model that we used is the Random Forest. Random Forest is a large collection of decision trees, which differ in structure and are not perfectly correlated. Each tree is trained on bootstrapped samples. It was first introduced by Breiman in 2001. The final random forest prediction is:

$$\hat{g}_{\mathrm{RF}}(Z) = \frac{1}{B}\sum_{b=1}^{B} \hat{g}_b\,(Z) \tag{3.23}$$

In the equation above, B is the total number of trees, $\hat{g}_b(Z)$ is the individual prediction function from the tree. One of the advantages of the Random Forest is that it is very flexible and capable of modeling complex interactions.

In our study, we trained our model using the randomForest function from the randomForest package with default parameters.

Now let us describe boosted trees, which is an ensemble-type learning method. In this method, the boosting process proceeds in the following way: firstly, a relatively simple model is estimated. After that, we examine where the model performed poorly and fit another model to the residuals. The process is repeated until the final prediction is obtained. Mathematically, we start the process by initializing residuals. $\mathrm{R}_i := \mathrm{Y}_i, \mathrm{i} = 1, \dots \mathrm{n}$, then fit the tree to the data. Then, we update residuals by subtracting the tree's predictions, scaled by a factor λ. The process goes from 1 to J steps, and the final model can be expressed as:

$$\hat{g}(Z) = \sum_{j=1}^{J} \lambda\, \hat{g}_j(Z) \tag{3.24}$$

We used the xgboost (Extreme Gradient Boosting) package and the cross-validation method via the xgb.cv function with 5-folds to determine the optimal number of boosting rounds (nrounds = 1000, eta = 0.1, max_depth = 4). The iteration range corresponding to the best cross-validated performance was used for prediction.

To improve the stability and robustness of our estimates, we also used Least Squares Model Averaging (LSMA) across all learners. We consider a linear combination of the predictions generated by five base learners: Lasso, Ridge, Elastic Net, Random Forest, and XGBoost. Let $\hat{g}_k(Z)$ denote the prediction from the learner k, where Z represents the covariate set. The aggregated (averaged) prediction is defined as:

$$\tilde{g}(Z) = \sum_{k=1}^{K} \tilde{\alpha}_k \, \hat{g}_k(Z) \tag{3.25}$$

where $\tilde{\alpha}_k$ are the weights estimated through least squares regression of the observed outcome on the predictions from each learner. In that way, we minimize the in-sample squared prediction error over a validation fold, treating the predictions from the individual learners as regressors.

In our setting, we first compute the residualized predictions $\tilde{Y}_{it}$ and $\tilde{D}_{it}$ for each learner via cross-fitting. Then we use these residuals to generate aggregated predictions by regressing the observed outcomes Y and D on the collection of learner predictions (without intercept). The coefficients that we get are the LSMA weights assigned to each learner.

Finally, we estimate the treatment effect β by regressing the residuals from the combined Y-model on the residuals from the combined D-model. This provides a more accurate estimate by combining information from all models.

4. RESULTS AND DISCUSSION

4.1 Estimation Results

4.1.1 FE models estimation

We have estimated several models in order to analyze the effect of curative and preventive healthcare expenditures on GDP per capita. In total, there were four sets of models. The first set only considers curative health expenditure, the second set focuses on preventive health expenditure, and the third and fourth sets are the extensions of the model for curative health expenditure, which attempt to decompose the effect across income and education spending groups, respectively. Results for the first and second sets of models are presented in the Appendix (see Appendix A.1 and Appendix A.2).

4.1.1.1 FE model for curative expenditure

To be sure that we are using the correct estimator, we conducted the Hausman test (see Table 4.1). The test showed there is a statistically significant difference between the RE (Random Effects) and FE estimator. The null hypothesis is "Random Effects model is preferred". Since the p-value is less than 0.05, we reject the null hypothesis and conclude that we should indeed use the FE estimator.

Table 4.1 : Hausman Test Results for Curative Models

Test Statistic	Degrees of Freedom	p-value
32.812	9	0.000144

In the first set of models, reported in Table 4.2 we show the effect of curative healthcare expenditure on the logarithm of real GDP per capita. Starting from basic OLS (model 1), we make our model more advanced step by step.

OLS is relatively straightforward and suitable for initial exploration, but it does not account for unobserved heterogeneity across countries. In model 1, we find a statistically significant and concave relationship between curative expenditure and

GDP per capita. Our model suggests diminishing marginal returns to curative spending on income levels. The coefficient on curative expenditure is 0.590 ($p < 0.001$), and the squared term is -0.034 ($p < 0.001$). However, this basic model may confuse between-country and within-country variation; therefore, we do not consider it for causal inference.

To address those potential biases, as the second step, we estimate a fixed effects (FE) model without controls (model 2). This specification now controls for time-invariant country characteristics that might influence the relationship, such as institutional quality or healthcare system structure. This model's results are quite different from OLS: the coefficients on both the curative expenditure and its square become statistically insignificant at the 5% level, and the explanatory power drops sharply ($R^2 = 0.007$, adjusted $R^2 < 0$). This suggests that the OLS results may have been driven by cross-country differences rather than true within-country dynamics.

In model 3, we extend the FE framework by including time and individual fixed effects, as well as the set of control variables. The rationale for the inclusion is that we want control for time-specific shocks and macroeconomic covariates that may simultaneously affect both health expenditure and income levels. In this specification, we observe a statistically significant U-shaped relationship. The coefficient on curative expenditure is -0.195 ($p < 0.1$), and its square is 0.032 ($p < 0.05$). These results suggest that at low levels, increases in curative spending are associated with declines in income per capita, but beyond a certain threshold, the effect will become positive. This nonlinearity is consistent with a scenario in which low levels of curative spending may be inefficient or take resources away from other important areas. In contrast, higher levels may reflect effective public health investments that support labor productivity and growth in GDP per capita.

Despite these improvements, fixed effects models do not entirely eliminate concerns about endogeneity. A key limitation is the potential for simultaneous causality bias, where economic performance may influence health expenditure decisions, leading to reverse causality. To address this, we estimate lagged versions of the models (models 4 through 6), using one-period lagged values of curative expenditure and its square as well as lagged values of controls, when applicable.

In the lagged OLS model (model 4), the results are similar to those in model 1: curative expenditure (lagged) has a positive coefficient of 0.607 ($p < 0.001$), and its square is negative (-0.035, $p < 0.001$), indicating again the concave relationship. In the lagged FE model without controls (model 5), the relationship remains statistically insignificant, indicating that control variables are necessary for capturing within-country dynamics.

Model 6 combines the full set of fixed effects and control variables with lagged regressors. Here, the coefficient on lagged curative expenditure and its square are approximately the same as in the model 3. We have to mention that the coefficient for lagged curative expenditure is significant at a 10% significance level, which is below the common 5% significance level threshold (-0.175 ($p < 0.1$)). Nevertheless, when we turn to the squared coefficient, the model suggests that it is positive and significant at the 5% significance level (- 0.030 ($p < 0.05$)). The magnitudes of the coefficients are very similar to those in the contemporaneous specification, suggesting robustness of the effect to lag structure. Therefore, the model 6 confirms the U-shaped relationship, but now we interpret it as a delayed effect: lower past curative expenditures are associated with lower income levels, but higher expenditures eventually yield positive returns.

Table 4.2 : FE regressions for Curative Expenditure.

	OLS Curative (model 1)	FE Curative (model 2)	FE Control Curative (model 3)	OLS lag Curative (model 4)	FE lag Curative (model 5)	FE lag Control lag Curative (model 6)
Curative	**0.590*** **	-0.061	**-0.195+**	**0.607*** **	-0.056	**-0.175+**
	(0.058)	(0.156)	(0.117)	(0.058)	(0.150)	(0.106)
Curative²	**-0.034*** **	0.006	**0.032** **	**-0.035*** **	0.005	**0.030** **
	(0.004)	(0.011)	(0.011)	(0.004)	(0.011)	(0.011)
Num.Obs.	648	648	648	621	621	621
Intercept	Yes	No	No	Yes	No	No
Controls included	No	No	Yes	No	No	Yes
Lagged controls	No	No	No	No	No	Yes
R2	0.244	0.007	0.295	0.253	0.007	0.331

Table 4.2 (continued) : FE regressions for Curative Expenditure.

	OLS Curative (model 1)	FE Curative (model 2)	FE Control Curative (model 3)	OLS lag Curative (model 4)	FE lag Curative (model 5)	FE lag Control lag Curative (model 6)
R2 Adj.	0.242	-0.078	0.225	0.251	-0.081	0.264
Akaike Information Criterion (AIC)	1054.1	-646.3	-853.7	1001.5	-640.8	-872.7
Bayesian Information Criterion (BIC)	1072.0	-632.9	-809.0	1019.2	-627.5	-828.4
Log.Lik.	-523.067			-496.754		
Root Mean Square Error (RMSE)	0.54	0.15	0.12	0.54	0.14	0.12

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Heteroskedasticity-robust standard errors. Models 4-6 use lagged versions of variables, identical row labels are used for readability; see model specifications for details or refer to the full table in the Appendix.

We also computed the marginal effect and the turning point for the model 6. As illustrated in the Figure 4.1, the marginal effect of curative spending on log GDP per capita becomes positive only after curative expenditure exceeds approximately 2.87% of GDP. Below this threshold, additional curative spending yields a negative or negligible contribution to real GDP per capita. We also report the predicted log of GDP per capita depending on the curative expenditure while holding controls at means (see Figure 4.2).

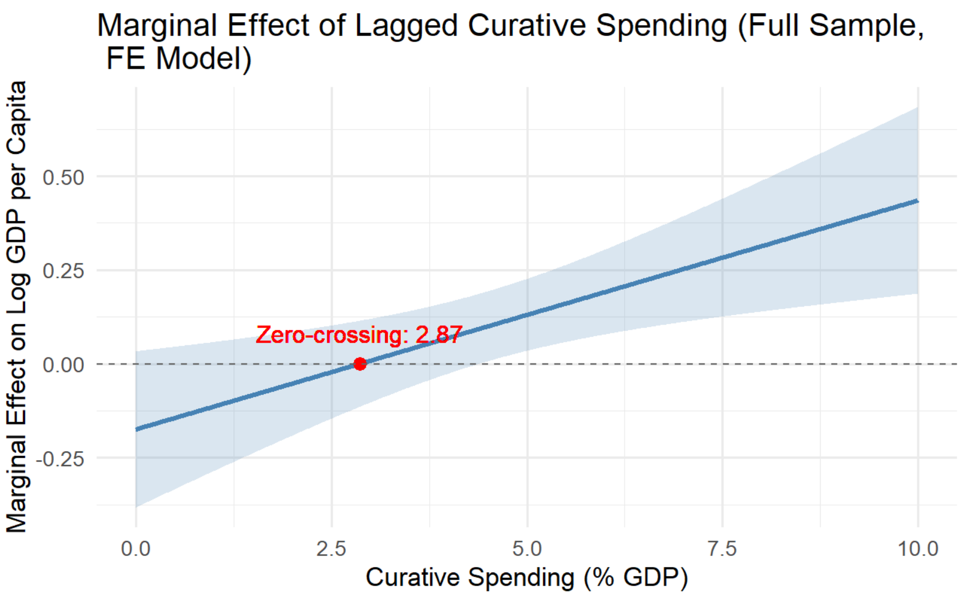

Figure 4.1 : Marginal Effect for Curative Expenditure, 95% CI bands.

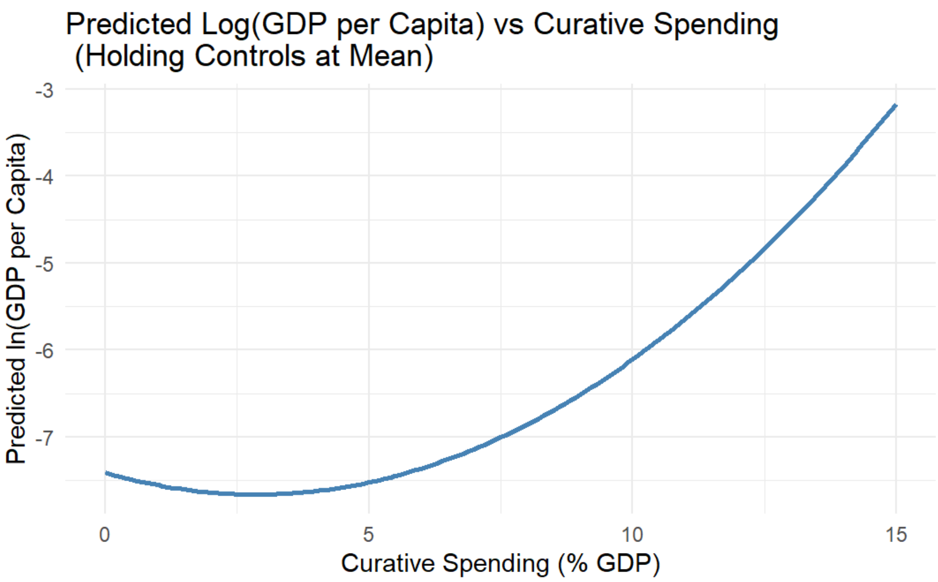

Figure 4.2 : Predicted Log of GDP per Capita and Curative Expenditure.

4.1.1.2 FE model for preventive expenditure

Again, we also conducted the Hausman test (see Table 4.3). As the p-value is less than 0.05, we reject the null hypothesis and conclude that the FE estimator is preferable.

Table 4.3 : Hausman Test Results for Preventive Models

Test Statistic	Degrees of Freedom	p-value
48.135	9	0.000000241

Now we can proceed to our models. In the second set of models, we focus on analyzing the relationship between preventive health expenditure and income per capita. As with curative expenditure, we follow a sequential model strategy to check the robustness of our estimation under different levels of control for heterogeneity and potential endogeneity. The results are presented in Table 4.4.

In the pooled OLS model without controls (model 7), we observe that preventive health expenditure appears to have a strong and statistically significant inverted U-shaped relationship with GDP per capita. The coefficient on preventive expenditure is 1.971 ($p < 0.001$), and the squared term is -1.340 ($p < 0.001$). However, this model again suffers from omitted variable bias and also does not distinguish between variation across countries and within countries over time; hence, its reliability for causal inference is limited.

To control for unobserved country-specific factors, we estimate a fixed effects model without controls (model 8). The model shows that both the linear and quadratic terms become statistically insignificant, and there is also a sharp decline in R^2 to 0.019. We think that the sudden decrease in explanatory power and the absence of significance mean that OLS estimates likely capture cross-country differences rather than true dynamic effects within countries.

We then include control variables and time fixed effects in model 9, but still the coefficients on preventive expenditure (-0.317) and its square (0.378) remain statistically insignificant. In this case, we conclude that the model specification fails to capture any significant effect of preventive health expenditure on income per capita after controlling for country heterogeneity and other determinants of income.

To reduce the risk of simultaneity bias, we include one-period lagged preventive expenditure terms in models 10 through 12. The lagged OLS model (model 10) replicates the inverted U-shaped relationship observed in model 7, but with larger coefficient magnitudes. This suggests that the effect of preventive expenditure on growth may be delayed and have an inverted U-shape. However, the lagged FE model with and without controls (models 11-12) again yields insignificant coefficients.

These findings suggest that, unlike curative expenditure, FE models were unable to detect a robust and consistent relationship between preventive health expenditure and real GDP per capita, even when controlling for unobserved heterogeneity and macroeconomic characteristics.

We think that this result is primarily due to a lack of variation in the preventive expenditure spending and data quality, potential issues, as the structure of the category of preventive expenditure was changing from year to year and country to country.

Table 4.4 : FE Models for Preventive Expenditure.

	OLS Preventive (model 7)	FE Preventive (model 8)	FE Control Preventive (model 9)	OLS lag Preventive (model 10)	FE lag Preventive (model 11)	FE lag Control lag Preventive (model 12)
Preventive	**1.971*** **	-0.576	-0.317	**2.115*** **	-0.741	-0.445
	(0.385)	(0.616)	(0.460)	(0.516)	(0.629)	(0.514)
Preventive²	**-1.340*** **	0.493	0.378	**-1.499***	0.649	0.672
	(0.380)	(0.424)	(0.350)	(0.635)	(0.532)	(0.522)
Num.Obs.	648	648	648	621	621	621
Intercept	Yes	No	No	Yes	No	No
Controls included	No	No	Yes	No	No	Yes
Lagged controls	No	No	No	No	No	Yes
R2	0.056	0.019	0.149	0.056	0.025	0.191
R2 Adj.	0.053	-0.065	0.065	0.053	-0.061	0.109
AIC	1198.6	-653.9	-732.2	1147.1	-652.1	-754.3
BIC	1216.5	-640.5	-687.4	1164.8	-638.8	-710.0
Log.Lik.	-595.319			-569.537		

Table 4.4 (continued) : FE Models for Preventive Expenditure.

	OLS Preventive (model 7)	FE Preventive (model 8)	FE Control Preventive (model 9)	OLS lag Preventive (model 10)	FE lag Preventive (model 11)	FE lag Control lag Preventive (model 12)
RMSE	0.61	0.15	0.14	0.61	0.14	0.13

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Heteroskedasticity-robust standard errors. Models 4-6 use lagged versions of variables, identical row labels are used for readability; see model specifications for details or refer to the full table in the Appendix.

Figure 4.3 illustrates the marginal effect of lagged preventive health spending on log GDP per capita from Model 12. When preventive spending exceeds approximately 0.33% of GDP, the marginal effect becomes positive. This turning point suggests that a minimal baseline of preventive investment is necessary to generate economic returns. Complementing this, in Figure 4.4 we present the predicted log GDP per capita across varying levels of preventive spending, holding all other variables at their means. These results suggest that while low levels of preventive spending may be economically ineffective, modest increases beyond the zero-crossing point yield positive returns, peaking around the optimal level we identified, although the relationship lacks statistical significance.

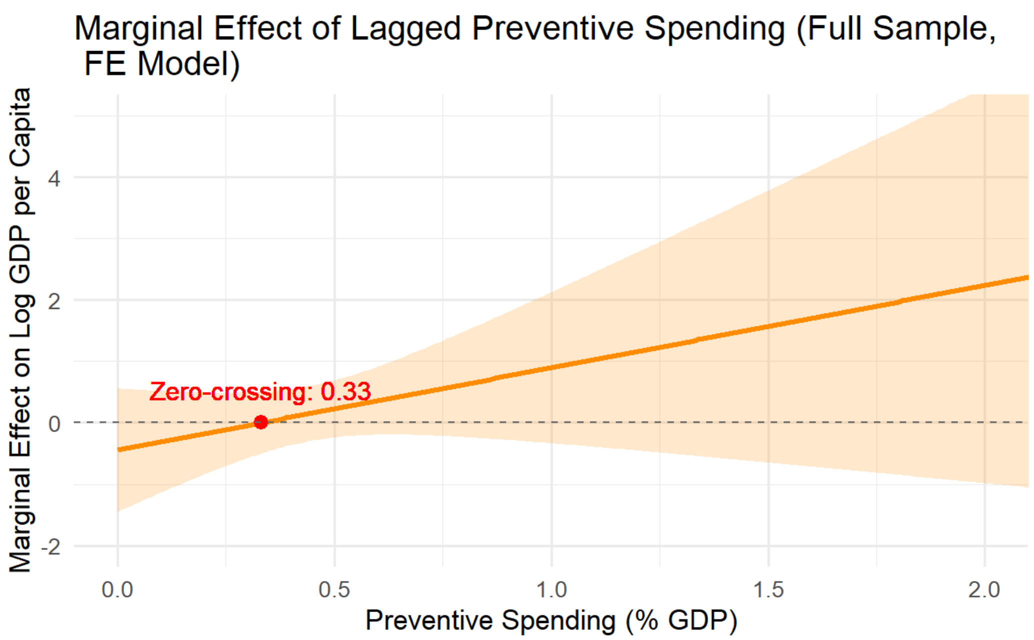

Figure 4.3 : Marginal Effect for Preventive Expenditure, 95% Confidence Bands.

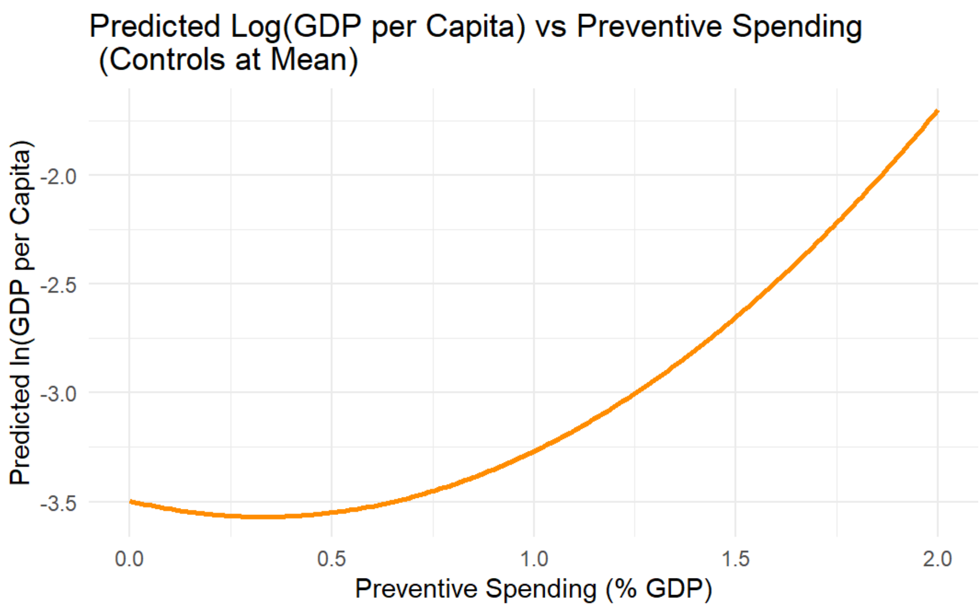

Figure 4.4 : Predicted Log of GDP and Preventive Expenditure (Holding Controls at Means).

4.1.1.3 Income level decomposition: extension of the FE model for curative expenditures

To investigate whether the impact of curative health expenditure on economic performance varies by income level, we divide the sample into two groups based on the mean level of real GDP per capita across all years and countries. By that method, we obtained 240 observations for the lower-income group sample and 408 for the higher-income group sample. It is essential to note that these groups are defined in relation to the OECD sample itself, rather than using an external income classification. We tried to apply Worlds Bank suggested classification for income categories, however most countries from our sample were placed into the high income category, which is expected, as OECD countries are characterized by a higher income level.

After that, we estimated fixed effects models with control variables separately for each group. We report the main results in Table 4.5, full results can be found in Appendix A.4. Low-income group regression shows that there is no significant association between curative expenditure and GDP per capita for contemporaneous and lagged specifications. In contrast, in high-income group regression, we find a statistically significant positive coefficient on curative expenditure (0.211, $p < 0.05$) and a marginally significant negative coefficient on its squared term (-0.022, $p < 0.1$). Therefore, we conclude that there is statistical evidence for a U-shaped relationship where curative spending shows a positive effect on economic performance after a certain threshold.

We can also observe differences across income groups in the coefficients of the control variables. Firstly, life expectancy has a significant negative effect on the log of GDP per capita in the low-income group (-0.101 to -0.107, $p < 0.05$) with no significant effect in the high-income group. Secondly, education spending is positively associated with economic performance in low-income countries (0.076, $p < 0.05$), but has a significant and negative coefficient in high-income countries (-0.042 to -0.067), which, in our opinion, can reflect diminishing returns to education investment at higher levels of development. Lastly, population and the share of the elderly population are significant only in the low-income group. At the same time, short-term interest rates exhibit a strong positive association with growth in the high-income sample.

It will be worth mentioning that R^2 and adjusted R^2 stay quite high for our low-income countries regression; our model is able to explain around 60% of the variation in the economic growth. However, as we move to high-income countries regression, the adjusted R^2 drops to 0.1, meaning that some other factors start to have an impact on economic performance apart from our control variables.

We believe that the model comparisons reveal potentially important heterogeneity, even though we acknowledge that we did not formally test for coefficient equality across groups. We conclude that the extended analysis with larger samples and alternative income classifications can be a direction for future research.

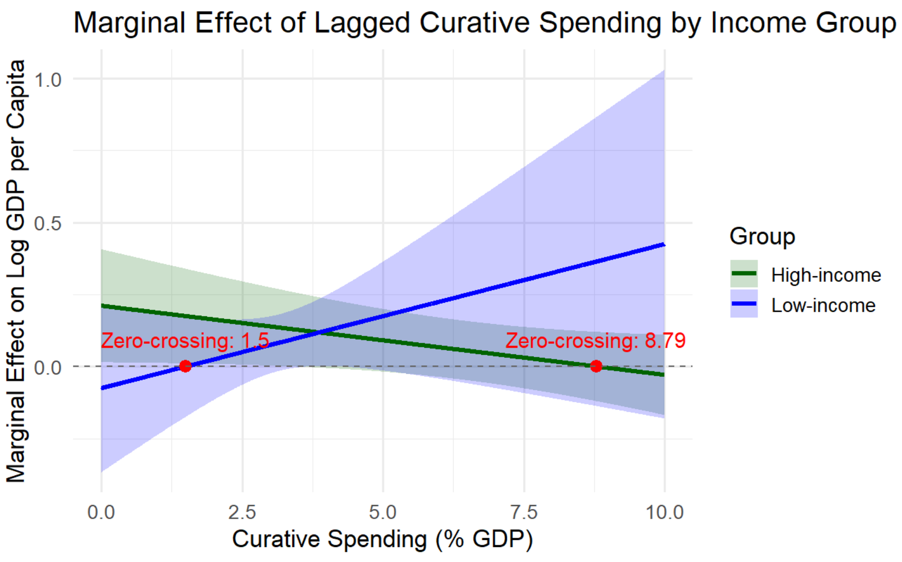

Figure 4.5 : Marginal Effect for Curative Spending by Income Group.

Table 4.5 : Curative Expenditure - Decomposition by Income Groups.

	FE Control Curative Low Income (model 13)	FE Control Lag Curative Low Income (model 14)	FE Control Curative High Income (model 15)	FE Control Lag Curative High Income(model 16)
Curative	-0.055	-0.075	0.129	**0.211***
	(0.132)	(0.149)	(0.088)	(0.100)
Curative²	0.023	0.025	-0.004	**-0.012+**
	(0.020)	(0.022)	(0.005)	(0.007)
Life Expectancy	**-0.101***	**-0.107***	-0.009	-0.010
	(0.040)	(0.045)	(0.045)	(0.043)
Education Spending	**0.076***	0.060	**-0.042***	**-0.067****
	(0.035)	(0.037)	(0.019)	(0.024)
Population	**-0.064*****	**-0.063*****	-0.003	0.001
	(0.010)	(0.011)	(0.004)	(0.005)
Share of Population 65+	**-0.108***	**-0.115****	**-0.022+**	-0.015
	(0.042)	(0.036)	(0.012)	(0.013)
Trade	**0.005****	**0.006*****	0.000	0.001
	(0.002)	(0.002)	(0.001)	(0.001)

Table 4.5 (continued) : Curative Expenditure - Decomposition by Income Groups.

	FE Control Curative Low Income (model 13)	FE Control Lag Curative Low Income (model 14)	FE Control Curative High Income (model 15)	FE Control Lag Curative High Income(model 16)
Savings Rate	0.007	0.007	**0.008***	**0.011***
	(0.009)	(0.008)	(0.004)	(0.004)
Short-term Interest Rate	-0.003	-0.000	**0.041*****	**0.028****
	(0.009)	(0.009)	(0.009)	(0.009)
Num.Obs.	240	230	408	391
R2	0.665	0.698	0.202	0.221
R2 Adj.	0.596	0.634	0.095	0.114
AIC	-340.2	-361.2	-765.2	-742.6
BIC	-305.4	-326.8	-725.1	-702.9
RMSE	0.11	0.11	0.09	0.09

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Heteroskedasticity-robust standard errors. Models 14, 16 use lagged versions of variables, identical row labels are used for readability; see model specifications for details or refer to the full table in the Appendix

In Figure 4.5 we show the marginal effect of lagged curative health spending on log GDP per capita, disaggregated by income group. We can see some contrast between high-income and low-income countries from our sample. For low-income economies, after crossing the turning point of 1.5% GDP, the marginal effect is positive. In contrast, the turning point for high-income countries is predicted to be located further on the line at around 8.79% of GDP. Since there is a U-shaped relationship for the high-income group, the marginal effect becomes negative after the turning point at 8.79. The difference shows us the importance of the economic context, particularly when designing policies aimed at maximizing the economic returns of healthcare investments.

4.1.1.4 Education spending level decomposition: extension of the FE model for curative expenditures

To explore the potential difference in the effect of curative expenditure based on education spending level, we created one more extension to the FE model. In this case, we included the interaction term between education spending as a percentage of GDP and the curative expenditure as a percentage of GDP, as well as its squared version. The results are presented in Table 4.6.

To investigate whether the effect of curative health expenditure on economic performance is moderated by differences in public education spending levels, we extend the fixed effects model by introducing interaction terms between curative health expenditure and education spending. This approach allows us to assess whether the effect of curative spending varies depending on a country's level of investment in education. The results are presented in Appendix A.4.

In the contemporaneous specification (model 17), we find that the baseline effect of curative expenditure is negative and significant (-0.279, $p < 0.05$), while the squared term is positive and highly significant (0.053, $p < 0.001$), indicating a U-shaped relationship. We should note that the interaction term Curative^2 * Education Spending is negative and significant (-0.003, $p < 0.05$), meaning that the U-shape in high-education countries, marginal returns to curative spending decline sooner. The lagged specification (model 18) also presents similar results; the coefficient sizes and directions are the same as in model 17. Therefore, we will not discuss them in detail.

As a robustness check, we also conducted the joint F-tests for interaction terms in both models. We observed that the effects are highly significant ($F = 11.35$ and $F = 13.50$, $p < 0.001$), therefore we believe that our estimations of moderating role of education are statistically robust.

Among the control variables, population size and savings rate remain significant and have expected signs across models. In contrast, the effects of life expectancy and short-term interest rates are not statistically significant. These patterns are consistent with earlier specifications, and we can see that macroeconomic fundamentals continue to play a role even when we account for the conditional effects of education on health spending outcomes. Overall, the results suggest that the relationship between curative healthcare spending and economic performance is also shaped by the level of education investment in the country.

Table 4.6 : Curative Expenditure - Decomposition by Education Expenditure.

	FE Control Curative by Education (model 17)	FE Control Curative by Education (Lag) (model 18)
Curative	**-0.279***	**-0.283***
	(0.125)	(0.114)
Curative²	**0.053*****	**0.056*****
	(0.013)	(0.013)
Curative * Education Spending	0.011	0.013
	(0.012)	(0.012)
Curative² * Education Spending	**-0.003***	**-0.004***
	(0.002)	(0.002)
Life Expectancy	-0.070	-0.080
	(0.046)	(0.051)
Population	**-0.036*****	**-0.036*****
	(0.009)	(0.009)
Share of Population 65+	-0.028	-0.027
	(0.019)	(0.018)

Table 4.6 (continued) : Curative Expenditure - Decomposition by Education Expenditure.

	FE Control Curative by Education (model 17)	FE Control Curative by Education (Lag) (model 18)
Trade	0.002	**0.003+**
	(0.002)	(0.001)
Savings Rate	**0.008****	**0.010****
	(0.003)	(0.003)
Short-term Interest Rate	-0.004	-0.005
	(0.006)	(0.006)
Num.Obs.	648	621
R2	0.331	0.367
R2 Adj.	0.263	0.302
AIC	-885.6	-904.9
BIC	-836.4	-856.2
RMSE	0.12	0.11
Joint F-test: Interactions (F, df1, df2, p)	F = 11.35, df = (2, 588), p = 0.0000	F = 13.50, df = (2, 562), p = 0.0000

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001. Heteroskedasticity-robust standard errors. Model 18 uses lagged versions of variables, and identical row labels are used for readability; see model specifications for details or refer to the full table in the Appendix A.4

4.1.2 Double machine learning estimation

4.1.2.1 Curative and lagged curative expenditure

This section presents results from applying the DML framework to estimate the causal effect of curative and lagged curative health expenditure on economic performance. Following Chernozhukov et al. (2024), DML is particularly well-suited to this setting, as it allows us to estimate low-dimensional target parameters while flexibly controlling for high-dimensional confounders using modern machine learning tools. We consider six machine learning (ML) learners for estimating the nuisance parameters (m and l): (i) Lasso, (ii) Ridge, (iii) Elastic Net, (iv) Random Forest, (v) Extreme Gradient Boosting (XGBoost), and (vi) Least Squares Model Averaging. We use root mean square error (RMSE) for predicting the outcome variable (Y) and the variable of interest (D) to identify the best-performing learners.

Table 4.7 summarizes the results for current curative expenditure. We report three specifications: All Initial X, Except X1998_Y, and No Initial X. These specifications refer to how initial values (from year 1998) are handled in the model: “All Initial X” includes all baseline covariates from 1998; “Except X1998_Y” excludes only the initial value of the outcome variable (Y) while keeping the rest; and “No Initial X” removes all initial (1998) variables from the set of controls. Across most specifications, including the best-performing model, which was XGBoost (Depth 4), estimates for both the linear and squared terms are small in magnitude and statistically insignificant. The only exception is the Ridge regression model, excluding initial values of X, which yields a statistically significant estimate of the linear term (0.1146) and a positive, but not significant, squared term (0.0057). Ridge suggests that there are increasing marginal returns to curative spending above a certain threshold. This result can be explained by how Ridge handles multicollinearity and coefficient shrinkage. Ridge regression applies ℓ^2 (L2) regularization, which penalizes large coefficients but does not eliminate variables from the model. Instead, it shrinks all coefficients toward zero in a smooth way. This behavior makes Ridge especially effective when the true signal is weak but spread across multiple correlated features.

However, drawing strong conclusions from this single specification would be premature. As emphasized by Chernozhukov et al. (2024), we should not rely on a single learner, but rather several learners. Therefore, the fact that we see the lack of significance in other learners—such as Lasso, Elastic Net, and tree-based methods—suggests sensitivity to the model structure. Given this instability, we conclude that there is no strong evidence for a contemporaneous effect of curative spending.

Table 4.7 : Curative Expenditure - DML Estimates.

Curative Expenditure – Double ML					
Model	Specification	Curative	SE	Curative ^{2}	SE
Lasso (CV)	All Initial X	0.0937	0.0849	-0.0048	0.0073
	Except X1998_Y	-0.0153	0.0915	0.0048	0.008
	No Initial X	0.0493	0.0998	0.0051	0.0094

Table 4.7 (continued) : Curative Expenditure - DML Estimates.

Model	Specification	Curative	SE	Curative^2	SE
Ridge (CV)	All Initial X	0.0268	0.0303	0.0007	0.0028
	Except X1998_Y	0.0521	0.034	0.0044	0.003
	No Initial X	**0.1146*** **	0.0336	0.0057	0.0032
Elastic Net (.5, CV)	All Initial X	0.0675	0.0813	-0.0059	0.0075
	Except X1998_Y	-0.0430	0.0886	0.0066	0.0082
	No Initial X	0.0325	0.0992	0.0037	0.009
Random Forest	All Initial X	-0.0386	0.0433	-0.0019	0.0027
	Except X1998_Y	-0.0278	0.0461	-0.0012	0.0031
	No Initial X	0.0083	0.0517	0.0022	0.0038
XGBoost (Depth 4)	All Initial X	0.1015	0.0925	0.0030	0.004
	Except X1998_Y	0.0260	0.114	0.0017	0.0054
	No Initial X	0.1630	0.0986	0.0069	0.0052
Best Performing Model	All Initial X	0.1015	0.0925	0.0030	0.004
	Except X1998_Y	0.0260	0.114	0.0017	0.0054
	No Initial X	0.1630	0.0986	0.0069	0.0052
Least Squares Avg.	All Initial X	0.0922	0.0761	0.0024	0.0054
	Except X1998_Y	0.0383	0.1104	0.0011	0.0069
	No Initial X	0.1561	0.1101	0.0070	0.0061

***p < 0.001, **p < 0.01, *p < 0.05

In contrast, results for lagged curative expenditure show greater consistency and robustness across learners (see Table 4.8). The Ridge model again produces statistically significant estimates for both the linear term (0.1141) and the squared term (0.0074). Interestingly, XGBoost, the highest-performing nonparametric learner, also identifies a significant convex effect (0.0069), even though the linear term is insignificant. The agreement between Ridge and XGBoost in the lagged specification increases our confidence in the presence of a real, convex causal effect. These results suggest that curative spending may influence economic growth with a temporal lag, reflecting delayed productivity gains from health improvements.

Table 4.8 : Lagged Curative Expenditure - DML Estimates.

Lagged Curative Expenditure – Double ML

Model	Specification	Curative	SE	Curative² Est.	Curative² SE
Lasso (CV)	No Initial X	0.0261	0.1056	0.0081	0.0095
Ridge (CV)	No Initial X	**0.1141*** **	0.0312	**0.0074** **	0.0030
Elastic Net (.5, CV)	No Initial X	-0.0120	0.0987	0.0075	0.0091
Random Forest	No Initial X	-0.0018	0.0464	0.0006	0.0038
XGBoost (Depth 4)	No Initial X	0.1013	0.0720	**0.0069*** **	0.0022
Best Performing Model	No Initial X	0.1013	0.0720	**0.0069*** **	0.0022
Least Squares Avg.	No Initial X	0.0738	0.0643	0.0000	0.0037

Table 4.9 - Table 4.11 present the model diagnostics in terms of RMSE for the outcome equation (Y), Curative, and squared Curative across our three specifications. Across all tables, XGBoost consistently shows the best predictive performance, with the lowest RMSE values for all three components. XGBoost is the only learner for which we observe a decrease in RMSE_Y when we move towards the “No Initial X” specification. The second-best learner is Random Forest, which shows performance similar to XGBoost; however, it is less successful at predicting the Curative squared term.

If we focus only on the “No Initial X” specification, XGBoost achieves an average RMSE_Y of approximately 0.08, significantly outperforming other learners, such as Lasso and Ridge. Similarly, in the estimating Curative, XGBoost achieves the lowest RMSE of 0.08, while Ridge regression performs the worst, with RMSE_D around 0.25. For the squared Curative term, we may notice even more divergence: Ridge regression produces RMSE values above 2.35, while XGBoost has RMSE_D² around 0.61. This indicates that Ridge struggles to accurately model the Curative and its square, despite producing statistically significant estimates in some specifications. In contrast, XGBoost is the most effective at capturing both outcome and Curative Expenditure variation, though it does not yield significant coefficient estimates. These results indicate that better predictive performance in nuisance models does not necessarily lead to more significant or stable causal estimates, and model choice should consider both fit and robustness in the final treatment effect estimation. Overall, on average, tree-based methods seem to be more efficient in our case.

Table 4.9 : RMSE_Y Comparison For Curative Models.

RMSE_Y Comparison Across Specifications			
Model	All Initial X	Except X1998_Y	No Initial X
Lasso (CV)	0.1458	0.1672	0.2058
Ridge (CV)	0.1949	0.2224	0.2453
Elastic Net (.5, CV)	0.1473	0.1680	0.2113
Random Forest	0.0987	0.1025	0.1097
XGBoost	0.4126	0.0748	0.0805

Note: in the table, we provide the average RMSE for 2-step estimation

Table 4.10 : RMSE_D Comparison for Curative Models.

RMSE_Curative Comparison Across Specifications			
Model	All Initial X	Except X1998_Y	No Initial X
Lasso (CV)	0.1033	0.1068	0.2058
Ridge (CV)	0.2775	0.2772	0.2459
Elastic Net (.5, CV)	0.1082	0.1082	0.2087
Random Forest	0.1210	0.1201	0.1102
XGBoost	0.0334	0.0335	0.0804

Table 4.11 : RMSE_D_squared Comparison for Curative Models

RMSE_Curative_squared Comparison Across Specifications			
Model	All Initial X	Except X1998_Y	No Initial X
Lasso (CV)	1.0537	1.0644	1.0541
Ridge (CV)	2.3504	2.3535	2.3507
Elastic Net (.5, CV)	1.0485	1.0361	1.0849
Random Forest	1.5475	1.5381	1.5382
XGBoost	0.6122	0.6103	0.6104

Now let us move to the model diagnostics for Lagged Curative DML estimation (see Table 4.12). The model diagnostics confirm the superior predictive performance of XGBoost across all components. It achieves the lowest RMSE for the outcome variable at 0.0780, outperforming both linear and tree-based alternatives. In estimating the lagged curative expenditure and its square, XGBoost again leads with RMSEs of 0.0391 and 0.7780, respectively. These results reflect its good performance in capturing both linear and nonlinear patterns in our data. In contrast, Ridge regression shows the weakest performance across all metrics, particularly for Lagged Curative and Lagged Curative², with RMSEs exceeding 0.28 and 2.36. Lasso and Elastic Net perform better than Ridge but are clearly outperformed by XGBoost.

Table 4.12 : RMSE for Lagged Curative Models.

	RMSE_Y	RMSE_Lag Curative	RMSE_Lag Curative_squared
Model	Specification - No Initial X		
Lasso (CV)	0.1989	0.1068	1.1051
Ridge (CV)	0.2428	0.2846	2.3684
Elastic Net (.5, CV)	0.2043	0.1063	1.0812
Random Forest	0.1134	0.1281	1.8206
XGBoost	0.0780	0.0391	0.7780

4.1.2.2 Preventive and lagged preventive expenditure

Table 4.13 provides the estimates from DML for preventive models across various specifications. Overall, we can observe a significant negative squared coefficient for preventive expenditure, for which most of the models agree. In particular, we can see that Lasso and Elastic Net deliver robust and statistically significant estimates, where the preventive coefficient is significantly positive and the preventive square coefficient is significantly negative. For instance, under the "No Initial X" specification, Lasso estimates a coefficient of 0.7839 ($p < 0.001$) for the linear term and –0.4362 ($p < 0.01$) for the squared term, which indicates a concave relationship—consistent with the theoretical expectation of diminishing returns. Ridge regression only detects statistically significant linear coefficients and smaller effect sizes. Random Forest does not detect any significant results, being an exception. XGBoost only provides a significant negative quadratic coefficient. These findings strongly support the hypothesis of an optimal level of preventive investment and diminishing returns on economic performance of preventive expenditure.

Table 4.13 : Preventive Expenditure - DML estimates.

Preventive Expenditure – Double ML					
Model	Specification	Preventive	SE	Preventive²	SE
Lasso (CV)	All Initial X	**0.7782*** **	0.1599	**-0.3066***	0.1254
	Except X1998_Y	**0.6518*** **	0.137	**-0.2662***	0.109
	No Initial X	**0.7839*** **	0.1356	**-0.4362** **	0.1486

Table 4.14 (continued) : Preventive Expenditure - DML estimates.

Model	Specification	Preventive	SE	Preventive²	SE
Ridge (CV)	All Initial X	0.2809	0.147	-0.0403	0.0998
	Except X1998_Y	0.2717	0.1527	-0.0774	0.0992
	No Initial X	**0.612*** **	0.1467	-0.181	0.1057
Elastic Net (.5, CV)	All Initial X	**0.7933*** **	0.1646	**-0.2826***	0.1343
	Except X1998_Y	**0.5369*** **	0.1185	-0.1409	0.1059
	No Initial X	**0.7785*** **	0.1095	**-0.4065*** **	0.1195
Random Forest	All Initial X	0.0491	0.0932	0.0124	0.0463
	Except X1998_Y	0.0508	0.0888	0.0098	0.0472
	No Initial X	0.1543	0.1481	0.0214	0.0943
XGBoost (Depth 4)	All Initial X	-0.0908	0.1511	-0.089	0.0464
	Except X1998_Y	-0.0581	0.1129	**-0.0738***	0.0319
	No Initial X	-0.1842	0.1586	**-0.1315****	0.0507
Best Performing Model	All Initial X	-0.0908	0.1511	-0.089	0.0464
	Except X1998_Y	-0.0581	0.1129	**-0.0738***	0.0319
	No Initial X	-0.1842	0.1586	**-0.1315****	0.0507
Least Squares Avg.	All Initial X	0.1202	0.3779	0.0436	0.1721
	Except X1998_Y	0.4406	0.4238	0.0588	0.1533
	No Initial X	0.2385	0.4484	0.0412	0.1951

***p < 0.001, **p < 0.01, *p < 0.05

Now let us proceed to the lagged Preventive models. Despite using the same specification ("No Initial X"), the DML estimates for lagged preventive expenditure show substantial variation across learners (see Table 4.14). Lasso and Ridge suggest an inverted U-shaped relationship, and if we compute the turning point, it will be approximately around 0.504% of GDP. This appears to be in line with the theory of diminishing returns. However, the best learner, according to RMSE (XGBoost), suggests an increasing convex curve, which is not what we would expect. One possible explanation for this result is that preventive spending is clustered at very low levels in our sample, the model may be capturing only the left side of a broader inverted U-shape.

Table 4.14 : Lagged Preventive Expenditure - DML estimates.

Model	Specification	D Est.	D SE	D² Est.	D² SE
Lasso (CV)	No Initial X	**1.0669*** **	0.1451	**-1.0583*** **	0.1927
Ridge (CV)	No Initial X	**0.5417*** **	0.1774	-0.1168	0.1937
Elastic Net (.5, CV)	No Initial X	**1.0736*** **	0.1476	**-0.9589*** **	0.1728
Random Forest	No Initial X	**0.3877*** **	0.0771	**0.2380*** **	0.0481
XGBoost (Depth 4)	No Initial X	**0.3763*** **	0.0472	**0.2166*** **	0.0089
Best Performing Model	No Initial X	**0.3763*** **	0.0472	**0.2166*** **	0.0089
Least Squares Avg.	No Initial X	0.2491	0.3882	0.0940	0.1011

***p < 0.001, **p < 0.01, *p < 0.05

Now we continue to model diagnostics and compare RMSEs for different learners (see Table 4.15 - Table 4.17). Similarly to curative models, XGBoost again takes the place of the best-performing learner with the lowest RMSEs. For example, for predicting the outcome variable, XGBoost shows an RMSE of 0.0744, while Random Forest achieves 0.10, and penalized regressions have more than 0.16 RMSE. When we focus on the performance of our models for predicting preventive expenditure, XGBoost shows a very low RMSE of 0.015, whereas other models struggle with RMSEs of around 0.04-0.05. Lastly, for squared preventive prediction, XGBoost is still the best, but the gap between its performance and other learners' performance is not that significant (0.03 vs 0.05-0.06).

When we analyze how the performance of our models changes across specifications, on average, penalized regressions perform worse at predicting the outcome variable as we exclude the initial X. For instance, Lasso's RMSE on Y increases from 0.1644 (with initial X) to 0.2161 (no initial X). This suggests that initial values may carry valuable predictive information about future outcomes. However, it is not the case for predicting preventive and preventive squared, as we do not see significant changes in RMSE across specifications for both nonparametric and penalized learners.

Overall, XGBoost's estimates are the most efficient and stable across specifications.

Table 4.15 : RMSE_Y Comparison For Preventive Models.

RMSE_Y Comparison Across Specifications			
Model	All Initial X	Except X1998_Y	No Initial X
Lasso (CV)	0.1644	0.1490	0.2161
Ridge (CV)	0.2237	0.1948	0.2478
Elastic Net (.5, CV)	0.1671	0.1484	0.2168
Random Forest	0.1019	0.0980	0.1095
XGBoost	0.0744	0.0721	0.0806

Note: in the table we provide average RMSE for 2 steps estimation

Table 4.16 : RMSE_D Comparison for Preventive Models.

RMSE_Preventive Comparison Across Specifications			
Model	All Initial X	Except X1998_Y	No Initial X
Lasso (CV)	0.0458	0.0459	0.0536
Ridge (CV)	0.0510	0.0511	0.0527
Elastic Net (.5, CV)	0.0479	0.0465	0.0533
Random Forest	0.0401	0.414	0.0397
XGBoost	0.0147	0.0147	0.0147

Table 4.17 : RMSE_D_squared Comparison for Preventive Models

RMSE_Preventive_squared Comparison Across Specifications			
Model	All Initial X	Except X1998_Y	No Initial X
Lasso (CV)	0.0475	0.0471	0.0473
Ridge (CV)	0.0513	0.0514	0.0511
Elastic Net (.5, CV)	0.0478	0.0482	0.0474
Random Forest	0.0620	0.0623	0.0616
XGBoost	0.0316	0.0316	0.0315

Table 4.18 summarizes RMSE for lagged preventive models. As we can see, predicting the outcome variable remains the hardest task for our models compared to predicting treatment variables. RMSE_Y is around 0.2 compared to RMSE_Lag_Curative and RMSE_Lag-Curative_squared, around 0.3-0.4. The second-best performer after XGBoost is Random Forest, although for predictive squared penalized learners, it provides lower RMSEs. From this we can conclude that for the preventive squared it is better to rely on estimates of Lasso and Ridge, and hence preventive expenditure will have an inverted U-shaped relationship.

Table 4.18 : RMSE for Lagged Preventive Models.

	RMSE_Y	RMSE_Lag Preventive	RMSE_Lag Preventive_squared
Model		No Initial X	
Lasso (CV)	0.2070	0.0429	0.0321
Ridge (CV)	0.2466	0.0433	0.0346
Elastic Net (.5, CV)	0.2088	0.0421	0.0324
Random Forest	0.1131	0.0302	0.0452
XGBoost	0.0785	0.0161	0.0301

4.1.3 Comparison of FE models and Double ML results

Table 4.19 summarizes the comparison between FE and DML estimates for curative and lagged curative expenditure. As we can see, both approaches agree on the presence of a nonlinear relationship, consistently identifying a significant and positive squared term. We can say that this supports the existence of increasing returns beyond a threshold level of spending. However, the size effect differs with the DML estimate being 4.6 times smaller than the FE estimate for squared contemporaneous and lagged curative terms. When it comes to the linear curative terms, our models are uncertain and yield only marginally significant coefficients (10% significance level); at the same time, the effect direction differs: FE models suggest a negative coefficient, while DML yields a positive coefficient. However, when we observe the lagged specification, the significance of the DML estimates disappears. These differences may have a root in how the two methods handle confounding, functional form assumptions, and model flexibility. FE models rely on strict parametric structures and may be more sensitive to omitted variable bias, while DML accounts for high-dimensional controls and interactions, capturing more complex underlying patterns.

Table 4.19 : Curative and Lagged Curative Coefficients Comparison.

	Curative		Curative^2	
	FE Models	Double ML	FE Models	Double ML
Coefficient	**-0.195+**	**0.1630+**	**0.032****	**0.0069*****
	Lag Curative		Lag Curative^2	
	FE Models	Double ML	FE Models	Double ML
Coefficient	**-0.175+**	0.1013	**0.030****	**0.0069*****

***p < 0.001, **p < 0.01, *p < 0.05, +p < 0.1

In Table 4.20, we compare coefficients for preventive and lagged preventive health expenditure. The results show clear disagreement between the two approaches, both in the sign and the direction of the coefficients. FE models suggest a weak and statistically insignificant relationship. In contrast, DML's best-performing model identifies a significant and positive effect for lagged preventive expenditure, including a strong positive squared term, indicating a convex (U-shaped) relationship. This suggests that economic benefits from preventive spending may emerge only at higher levels or with time lags. Again, we should note the DML's greater flexibility in modeling nonlinearities and controlling for complex interactions, while FE models may be limited by collinearity, low within-country variation in preventive spending. We think that these findings imply that traditional models may underestimate the delayed and nonlinear impact of preventive health investment.

Table 4.20 : Preventive and Lagged Preventive Coefficients Comparison.

	Preventive		Preventive^2	
	FE Models	Double ML	FE Models	Double ML
Coefficient	-0.317	-0.1842	0.378	**-0.1315****
	Lag Preventive		Lag Preventive^2	
	FE Models	Double ML	FE Models	Double ML
Coefficient	-0.445	**0.3763*****	0.672	**0.2166*****

***p < 0.001, **p < 0.01, *p < 0.05, +p < 0.1

Note: for Double ML only best learner's estimates are displayed.

5. CONCLUSIONS

In this thesis, we examined the economic impact of curative and preventive health expenditures across 29 OECD countries from 1998 to 2021. Using fixed effects and DML models, we assessed four key hypotheses.

In H1, we argued that curative health spending has a stronger and more direct effect on economic performance than preventive spending. This hypothesis is supported: curative expenditure showed a statistically significant U-shaped relationship with the real GDP per capita, but preventive spending showed no significant impact on economic growth, probably due to limited variation and data inconsistencies.

For the next hypothesis, H2, which proposed a nonlinear relationship, we can state that it is partially supported. Nonlinear effects were clearly observed for curative spending using both methodological approaches, and evidence for preventive spending was model-dependent. Only some DML models suggested an inverted U-shaped pattern, but these results were rather inconsistent across models.

In our H3, we expected the growth effects of health spending to be stronger in high-income countries, and we can conclude that it was partially supported. The positive effect of curative spending turned out to be statistically significant only in high-income OECD countries. However, the threshold and shape of the relationship can vary in different income groups, so it requires further formal testing.

In H4, we supposed that education spending would amplify the positive effect of curative health spending, and we conclude that it is not fully supported. Interaction models showed that countries with higher education investment experience marginal returns from curative health spending which are decreasing faster.

These findings lead to several policy implications. First of all, policymakers should be oriented to the efficiency thresholds to guide public investment in curative health. Next, there should be a strategic alignment between health and education policy for sustainable growth. Thirdly, countries that have currently underfunded health expenditures should first consider investment in education and then expand their health expenditure to see the maximum benefit for their economies.

Naturally, this study is not free of some limitations. Endogeneity still may remain a concern, as we did not apply instrumental variables or GMM methods, which can be considered in future research. Also, in the preventive expenditure data, there could be inconsistencies and low variation, which is something future research can focus on. Better classification and tracking of preventive health expenditures are needed for obtaining reliable estimates. Next, as our study was based on the OECD sample, there is limited generalizability to low-income or non-OECD countries. In future works, it is possible to expand the sample beyond OECD countries. Additionally, in the future, one may test other sources of heterogeneity, such as institutional quality, inequality, or health system type.

Overall, the main conclusions from this thesis are as follows. Curative health spending supports the economic performance, but only beyond a certain threshold and with a time lag. Its effectiveness on the income and education settings in the country.

REFERENCES

Aghion, P., Howitt, P., & Murtin, F. (2010). The relationship between health and growth: When Lucas meets nelson-phelps. *Review of Economics and Institutions, Università Di Perugia, 2 (1)*. https://doi.org/10.3386/w15813

Atilgan, E., Ertuğrul, H. M., Baycan, O., & Ulucan, H. (2024). Health-led growth hypothesis and health financing systems: An econometric synthesis for OECD countries. *Frontiers in Public Health*, *12*. https://doi.org/10.3389/fpubh.2024.1437304

Baldacci, E., Cui, Q., Clements, B. J., & Gupta, S. (2004). Social spending, human capital, and growth in developing countries: Implications for achieving the mdgs. *IMF Working Papers*, *04*(217), 1. https://doi.org/10.5089/9781451875140.001

Baltagi, B. H. (2005). *Econometric analysis of panel data*. John Wiley & Sons Inc.

Barro, R. J., & Sala-i-Martin, X. (2004). *Economic growth*. MIT Press.

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to Education*. Columbia Univ. Press i distr.

Beylik, U., Cirakli, U., Cetin, M., Ecevit, E., & Senol, O. (2022). The relationship between Health Expenditure Indicators and economic growth in OECD countries: A Driscoll-Kraay Approach. *Frontiers in Public Health*, *10*. https://doi.org/10.3389/fpubh.2022.1050550

Chernozhukov, V. & Hansen, C. & Kallus, N. & Spindler, M. & Syrgkanis, V. (2024): *Applied Causal Inference Powered by ML and AI*. CausalML-book.org; arXiv:2403.02467.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, *21*(1). https://doi.org/10.1111/ectj.12097

Cohen, J. T., Neumann, P. J., & Weinstein, M. C. (2008). Does preventive care save money? Health Economics and the Presidential candidates. *New England Journal of Medicine*, *358*(7), 661–663. https://doi.org/10.1056/nejmp0708558

Eurostat. (2025). *Real GDP per capita* [Data table]. European Commission. https://ec.europa.eu/eurostat/databrowser/view/sdg_08_10/default/table?lang=en

Gmeinder, M., Morgan, D., & Mueller, M. (2017). How much do OECD countries spend on prevention? *OECD Health Working Papers*. https://doi.org/10.1787/f19e803c-en

Grossman, M. (1972). On the concept of Health Capital and the demand for Health. *Journal of Political Economy*, *80*(2), 223–255. https://doi.org/10.1086/259880

Hu, Q., & Wang, L. (2024). Economic growth effects of public health expenditure in OECD countries: An empirical study using the dynamic panel threshold model. *Heliyon*, *10*(4). https://doi.org/10.1016/j.heliyon.2024.e25684

Maciosek, M. V., Coffield, A. B., Flottemesch, T. J., Edwards, N. M., & Solberg, L. I. (2010). Greater use of preventive services in U.S. health care could save lives at little or no cost. *Health Affairs*, *29*(9), 1656–1660. https://doi.org/10.1377/hlthaff.2008.0701

Moreno-Serra, R., & Smith, P. C. (2011). (working paper). *The effects of health coverage on population outcomes: a country-level panel data analysis*. Washington, DC: Results for Development Institute.

Mushkin, S. J. (1962). Health as an investment. *Journal of Political Economy*, *70*(5, Part 2), 129–157. https://doi.org/10.1086/258730

OECD, Eurostat, & World Health Organization. (2011). (rep.). *A System of Health Accounts 2011: 2011 Edition*. Paris: OECD Publishing.

Piabuo, S. M., & Tieguhong, J. C. (2017). Health Expenditure and Economic Growth - A review of the literature and an analysis between the Economic Community for Central African States (CEMAC) and selected African countries. *Health Economics Review*, *7*(1). https://doi.org/10.1186/s13561-017-0159-1

Prince, M. J., Wu, F., Guo, Y., Gutierrez Robledo, L. M., O'Donnell, M., Sullivan, R., & Yusuf, S. (2015). The burden of disease in older people and implications for Health Policy and practice. *The Lancet*, *385*(9967), 549–562. https://doi.org/10.1016/s0140-6736(14)61347-7

Russell, L. B. (1986). *Is prevention better than cure: Studies in Social Economics*. Brookings institution.

Sharma, R. (2018). Health and economic growth: Evidence from dynamic panel data of 143 years. *PLOS ONE*, *13*(10). https://doi.org/10.1371/journal.pone.0204940

Suhrcke, M., McKee, M., Stuckler, D., Sauto Arce, R., Tsolova, S., & Mortensen, J. (2006). The contribution of health to the economy in the European Union. *Public Health*, *120*(11), 994–1001. https://doi.org/10.1016/j.puhe.2006.08.011

von Haldenwang, C., & Ivanyna, M. (2010). *Assessing the tax performance of developing countries* (Discussion Paper No. 20/2010). German Institute of Development and Sustainability (IDOS). https://hdl.handle.net/10419/199355

Wang, F. (2018). The roles of preventive and curative health care in economic development. *PLOS ONE*, *13*(11). https://doi.org/10.1371/journal.pone.0206808

Wang, F., & Wang, J.-D. (2021). Investing preventive care and economic development in ageing societies: Empirical evidences from OECD countries. *Health Economics Review*, *11*(1). https://doi.org/10.1186/s13561-021-00321-3

Wang, F., Huang, W.-Y., Chang, Y.-T., Yang, S.-C., & Wang, J.-D. (2024). Effective prevention in clinical practice may save human capital loss: Real-world evidence from Taiwan's National Health Insurance. *Journal of the Formosan Medical Association*, *123*. https://doi.org/10.1016/j.jfma.2024.10.011

Wang, F., Wang, J.-D., & Huang, Y.-X. (2016). Health Expenditures spent for prevention, economic performance, and social welfare. *Health Economics Review*, *6*(1). https://doi.org/10.1186/s13561-016-0119-1

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. The MIT Press.

World Bank. (n.d.). *GDP growth (GDP per capita growth)* [Metadata glossary entry]. World Bank DataBank. https://databank.worldbank.org/metadataglossary/statistical-capacity-indicators/series/5.51.01.10.gdp

APPENDICES

This appendix presents the complete regression outputs for the fixed effects (FE) models used in the thesis, with no omitted lines results. Four main tables are included:

Appendix Table A.1 shows the full FE estimation results for curative health expenditure, including OLS, contemporaneous and lagged FE models, both with and without control variables.

Appendix Table A.2 provides the corresponding results for preventive health expenditure using the same model structure.

Appendix Table A.3 demonstrates FE models for curative health spending by income level (low-income and high-income OECD countries).

Appendix Table A.4 presents models with interaction terms between curative health expenditure and education spending.

APPENDIX A

Appendix A.1 : Full Results for FE Regressions for Curative Expenditure.

	OLS Curative	FE Curative	FE Control Curative	OLS lag Curative	FE lag Curative	FE lag Control lag Curative
Intercept	8.365***			8.335***		
	(0.199)			(0.197)		
Curative	0.590***	-0.061	-0.195+			
	(0.058)	(0.156)	(0.117)			
Curative²	-0.034***	0.006	0.032**			
	(0.004)	(0.011)	(0.011)			
Lagged Curative				0.607***	-0.056	-0.175+
				(0.058)	(0.150)	(0.106)
Lagged Curative²				-0.035***	0.005	0.030**
				(0.004)	(0.011)	(0.011)
Life Expectancy			-0.067			
			(0.046)			
Education Spending			-0.040			
			(0.030)			

	OLS Curative	FE Curative	FE Control Curative	OLS lag Curative	FE lag Curative	FE lag Control lag Curative
Population			-0.032***			
			(0.009)			
Share of Population 65+			-0.018			
			(0.018)			
Trade			0.002			
			(0.002)			
Savings Rate			0.008**			
			(0.003)			
Short-term Interest Rate			-0.003			
			(0.007)			
Lagged Life Expectancy						-0.073
						(0.051)
Lagged Education Spending						-0.055
						(0.034)
Lagged Population						-0.031***
						(0.009)
Lagged Share of Population 65+						-0.017

	OLS Curative	FE Curative	FE Control Curative	OLS lag Curative	FE lag Curative	FE lag Control lag Curative
						(0.017)
Lagged Trade						0.002
						(0.002)
Lagged Savings						0.009**
						(0.003)
Lagged Short-term Interest Rate						-0.005
						(0.006)
Num.Obs.	648	648	648	621	621	621
R2	0.244	0.007	0.295	0.253	0.007	0.331
R2 Adj.	0.242	-0.078	0.225	0.251	-0.081	0.264
AIC	1054.1	-646.3	-853.7	1001.5	-640.8	-872.7
BIC	1072.0	-632.9	-809.0	1019.2	-627.5	-828.4
Log.Lik.	-523.067			-496.754		
RMSE	0.54	0.15	0.12	0.54	0.14	0.12
Std.Errors	model_1	model_2	model_3	model_4	model_5	model_6

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Appendix A.2 : Full Results FE regressions for Preventive Expenditure.

	OLS Preventive	FE Preventive	FE Control Preventive	OLS lag Preventive	FE lag Preventive	FE lag Control lag Preventive
Intercept	10.002***			9.996***		
	(0.080)			(0.092)		
Preventive	1.971***	-0.576	-0.317			
	(0.385)	(0.616)	(0.460)			
Preventive²	-1.340***	0.493	0.378			
	(0.380)	(0.424)	(0.350)			
Lagged Preventive				2.115***	-0.741	-0.445
				(0.516)	(0.629)	(0.514)
Lagged Preventive²				-1.499*	0.649	0.672
				(0.635)	(0.532)	(0.522)
Life Expectancy			-0.032			
			(0.063)			
Education Spending			-0.043			
			(0.035)			
Population			-0.013			

	OLS Preventive	FE Preventive	FE Control Preventive	OLS lag Preventive	FE lag Preventive	FE lag Control lag Preventive
			(0.009)			
Share of Population 65+			-0.003			
			(0.021)			
Trade			0.001			
			(0.002)			
Savings Rate			0.007*			
			(0.003)			
Short-term Interest Rate			0.005			
			(0.007)			
Lagged Life Expectancy						-0.035
						(0.070)
Lagged Education Spending						-0.061
						(0.040)
Lagged Population						-0.013
						(0.010)

	OLS Preventive	FE Preventive	FE Control Preventive	OLS lag Preventive	FE lag Preventive	FE lag Control lag Preventive
Lagged Share of Population 65+						-0.001
						(0.019)
Lagged Trade						0.002
						(0.002)
Lagged Savings						0.009*
						(0.004)
Lagged Short-term Interest Rate						0.002
						(0.008)
Num.Obs.	648	648	648	621	621	621
R2	0.056	0.019	0.149	0.056	0.025	0.191
R2 Adj.	0.053	-0.065	0.065	0.053	-0.061	0.109
AIC	1198.6	-653.9	-732.2	1147.1	-652.1	-754.3
BIC	1216.5	-640.5	-687.4	1164.8	-638.8	-710.0
Log.Lik.	-595.319			-569.537		
RMSE	0.61	0.15	0.14	0.61	0.14	0.13

	OLS Preventive	FE Preventive	FE Control Preventive	OLS lag Preventive	FE lag Preventive	FE lag Control lag Preventive
Std.Errors	model_7	model_8	model_9	model_10	model_11	model_12

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001

Appendix A.3 : Full Results for Curative Expenditure - Decomposition by Income Groups.

	FE Control Curative Low Income	FE Control Lag Curative Low Income	FE Control Curative High Income	FE Control Lag Curative High Income
Curative	-0.055		0.129	
	(0.132)		(0.088)	
Curative²	0.023		-0.004	
	(0.020)		(0.005)	
Lagged Curative		-0.075		0.211*
		(0.149)		(0.100)
Lagged Curative²		0.025		-0.012+
		(0.022)		(0.007)
Life Expectancy	-0.101*		-0.009	
	(0.040)		(0.045)	

	FE Control Curative Low Income	FE Control Lag Curative Low Income	FE Control Curative High Income	FE Control Lag Curative High Income
Education Spending	0.076*		-0.042*	
	(0.035)		(0.019)	
Population	-0.064***		-0.003	
	(0.010)		(0.004)	
Share of Population 65+	-0.108*		-0.022+	
	(0.042)		(0.012)	
Trade	0.005**		0.000	
	(0.002)		(0.001)	
Savings Rate	0.007		0.008*	
	(0.009)		(0.004)	
Short-term Interest Rate	-0.003		0.041***	
	(0.009)		(0.009)	
Lagged Life Expectancy		-0.107*		-0.010
		(0.045)		(0.043)
Lagged Education Spending		0.060		-0.067**

	FE Control Curative Low Income	FE Control Lag Curative Low Income	FE Control Curative High Income	FE Control Lag Curative High Income
		(0.037)		(0.024)
Lagged Population		-0.063***		0.001
		(0.011)		(0.005)
Lagged Share of Population 65+		-0.115**		-0.015
		(0.036)		(0.013)
Lagged Trade		0.006***		0.001
		(0.002)		(0.001)
Lagged Savings		0.007		0.011*
		(0.008)		(0.004)
Lagged Short-term Interest Rate		-0.000		0.028**
		(0.009)		(0.009)
Num.Obs.	240	230	408	391
R2	0.665	0.698	0.202	0.221
R2 Adj.	0.596	0.634	0.095	0.114
AIC	-340.2	-361.2	-765.2	-742.6

	FE Control Curative Low Income	FE Control Lag Curative Low Income	FE Control Curative High Income	FE Control Lag Curative High Income
BIC	-305.4	-326.8	-725.1	-702.9
RMSE	0.11	0.11	0.09	0.09
Std.Errors	model_21	model_22	model_23	model_24

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix A.4 : Full Results for Curative Expenditure - Decomposition by Education Spending.

	FE Control Curative by Education	FE Control Curative by Education (Lag)
Curative	-0.279*	
	(0.125)	
Curative²	0.053***	
	(0.013)	
Curative * Education Spending	0.011	
	(0.012)	
Curative² * Education Spending	-0.003*	
	(0.002)	
Lagged Curative		-0.283*

	FE Control Curative by Education	FE Control Curative by Education (Lag)
		(0.114)
Lagged Curative²		0.056***
		(0.013)
Lagged Curative * Education Spending		0.013
		(0.012)
Life Expectancy	-0.070	
	(0.046)	
Population	-0.036***	
	(0.009)	
Share of Population 65+	-0.028	
	(0.019)	
Trade	0.002	
	(0.002)	
Savings Rate	0.008**	
	(0.003)	
Short-term Interest Rate	-0.004	
	(0.006)	

	FE Control Curative by Education	FE Control Curative by Education (Lag)
Lagged Life Expectancy		-0.080
		(0.051)
Lagged Population		-0.036***
		(0.009)
Lagged Share of Population 65+		-0.027
		(0.018)
Lagged Trade		0.003+
		(0.001)
Lagged Savings		0.010**
		(0.003)
Lagged Short-term Interest Rate		-0.005
		(0.006)
Num.Obs.	648	621
R2	0.331	0.367
R2 Adj.	0.263	0.302
AIC	-885.6	-904.9
BIC	-836.4	-856.2

	FE Control Curative by Education	FE Control Curative by Education (Lag)
RMSE	0.12	0.11
Std.Errors	model_24	model_27
Joint F-test: Interactions (F, df1, df2, p)	F = 11.35, df = (2, 588), p = 0.0000	F = 13.50, df = (2, 562), p = 0.0000

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

CURRICULUM VITAE

Name Surname : **Anna Pirogova**

EDUCATION :

- **B.Sc.** : 2023, Higher School of Economics, Management, International Business and Management
- **Exchange program** : 2022-2023, Bogazici University, Faculty of Business Administration and Management
- **M.A.** : 2025, Istanbul Technical University, Graduate School, Economics, *Dissertation Stage*

PROFESSIONAL EXPERIENCE AND REWARDS:

- **Business Analyst / Business Process Specialist**
  PASV Programming Academy, Silicon Valley (Ed-Tech Startup, California, USA)
  (05/2021 – 05/2025)
  Managed sales teams; built SQL/Zoho dashboards; analyzed funnel metrics; automated CRM workflows; integrated payment systems.
- **Teaching Assistant**
  NRU Higher School of Economics, Saint Petersburg, Russia
  (09/2020 – 06/2021)
  Assisted in undergraduate courses in Microeconomics, Macroeconomics, Mathematics, and Economic Statistics.

PUBLICATIONS, PRESENTATIONS AND PATENTS ON THE THESIS:

- **Anna Pirogova** 2025: The Roles Of Preventıve And Curatıve Health Care In Economıc Development. 7th International Black Sea Modern Scientific Research Congress, June 24-26, 2025, Artvin, Turkey. (Presentation Instance)

OTHER PUBLICATIONS, PRESENTATIONS AND PATENTS:

- **Pirogova A., Taraskina E. V., Zazdravnykh E.A.** 2022**.** Access to Healthcare and Digitalization**.** GSOM emerging markets conference 2022: CONFERENCE BOOK, Saint-Petersburg, 05–08 October 2022. – Saint-Petersburg: St. Petersburg State University Graduate School of Management, 2022. – P. 240-246. – EDN MORNEB