Impact of Health Care Spending on Health Outcomes*

Ethan Maddy†

May 10, 2020

Abstract

A fixed-effect model to determine the impact, if any, health spending as a percent of Gross National Income (GNI) has on health outcomes. Data sourced for this research is from the World Bank's Health Nutrition and Population Statistics DataBank. The data used in this research includes information from 259 unique countries from 2001-2020.

*Special thanks to Dr. Tyler Ransom for his feedback and guidance throughout this process.
†Department of Economics, University of Oklahoma. E-mail address: ethanmaddy@ou.edu

1 Introduction

Empirical evidence has been inconclusive about the strength of the connection between health care spending and health outcomes. This paper uses the Health Nutrition and Population Statistics data from the World Bank to model the correlation between health expenditure percent of Gross National Income (GNI) and death rate, incidence of HIV, life expectancy at birth, malnutrition prevalence, mortality from cardiovascular disease, cancer, diabetes, or chronic respiratory disease, number of infant deaths, number of maternal deaths, number of people who are undernourished, and percent of the population that use at least basic sanitation services.

Around the world, countries and their respective governments decide how to allocate their wealth through government spending and reallocation. Most developed countries spend between 3-10% of their budget on health related expenditures. However, some countries spend less while others spend much more. For example, the United States spends 17% of GDP on health, yet does not see returns on its investment like some developed countries do.

In this paper, linear regression analysis is applied to determine the relationship between health spending per capita and various health outcomes. This research paper consists of brief literature review regarding health spending, description of the data used, relevant variables, and underlying assumptions. Later, statistical regression models are used to determine results, followed by a discussion of said results and a brief conclusion. All references, figures and tables can be found on pages 13 to 16.

2 Literature Review

While it may seem intuitive for health spending to be significantly correlated with health outcomes, Martin and Smith (2008) find that ”Empirical evidence has hitherto been inconclusive about the strength of the link between health care spending and health outcomes.” In countries like the United States, increased health spending is not always associated with better health outcomes. Due to a highly complex and fragmented payment system that weakens the demand side of the health sector and entails high administrative costs, the United States spends more per capita of health than any other country in the world, adjusted as a percentage of GDP- Reinhardt and Anderson (2004).

Recent literature has been more conclusive regarding the link between health spending and health outcomes. For example, Farag and Erbil (2013) finds that health spending has a significant effect on reducing infant and under-5 child mortality with an elasticity of 0.13 to 0.33 for infant mortality and 0.15 to 0.38 for under-5 child mortality in models estimated using fixed effects methods. Hu and Mendoza (2013) strive to improve upon previous empirical studies by conducting a more careful analysis of the determinants controlling for possible endogeneity. Hu and Mendoza (2013) also find that both public spending on healthcare and the quality of governance matter for the reduction of child mortality rates. They go on to say that their mixed results on the interaction of governance with public spending cast some doubt on the conclusiveness of previous empirical studies.

Through this paper, I hope to provide a small addition to the field of health by briefly qualifying the conclusion that health spending is a worthwhile investment and is an important measure to yield positive health outcomes around the world.

3 Data

The data source for this research is the World Bank's Health Nutrition and Population Statistics DataBank. DataBank is an analysis and visualisation tool that contains collections of time series data on a variety of topics. The Health Nutrition and Population Statistics data is classified as Public under the Access to Information Classification Policy. The data used includes information from 259 unique countries from 2001-2020. 5,185 observations were taken from these 259 countries and were included in a fixed-effect [1] analysis.

All variables are defined in Table 1 and summary statistics are provided in Table 2. The independent variable used in this paper is Health Spending per capita divided by Gross National Income (GNI). This variable was created to show the percentage of GNI that a country spends on health. With this variable, smaller countries are compared equally to larger countries by using a percentage of spending rather than an absolute value.

Gross National Income and Gross Domestic Product are similar but have distinct differences.

- GNI: Gross National Income - total income received by the country from its residents and businesses regardless of whether they are located in the country or abroad.
- GDP: Gross Domestic Product - total market value of all finished goods and services produced within a country in a set time period.

[1] Fixed-effects models are a class of statistical models in which the levels of independent variables are assumed to be fixed and only the dependent variable changes in response to the levels of independent variables. In this paper, 'Country Name' is fixed in the regression models.

Before running the analysis, multiple variables were mutated and changed to factor or numeric variables. Log was used when the variable was reported as an absolute and could not be accurately interpreted.

- *HealthSpending* = 'Current health expenditure per capita (current US$)'
- *GNI* = 'GNI per capita, Atlas method (current US$)'
- *log.DeathRate* = log of 'Death rate, crude (per 1,000 people)'
- *log.HIV* = log of 'Incidence of HIV, all (per 1,000 uninfected population)'
- *log.LifeExpect* = log of 'Life expectancy at birth, total (years)'
- *log.Malnutrition* = log of 'Malnutrition prevalence, weight for age (% of children under 5)'
- *Mortality* = 'Mortality from CVD, cancer, diabetes or CRD between exact ages 30 and 70 (%)'
- *log.InfantDeaths* = log of 'Number of infant deaths'
- *log.MaternalDeaths* = log of'Number of maternal deaths'
- *log. Undernourished* = log of 'Number of people who are undernourished'
- *Sanitation* = 'People using at least basic sanitation services (% of population)'

Gauss-Markov Assumptions:

1. Linear in parameters – The model consists of a simple linear regression that is linear in parameters and follows the equation below:

$$Y = \beta_0 + \beta_1 + \beta_1 + ... + u \tag{1}$$

2. Random sample: The data includes approximately 5,185 observations 289 countries. Considering this, the sampling can be assumed as random.

3. No perfect collinearity: There is no perfect collinearity between the variables in the model since there is only one X variable.

4. Exogeneity $E(u|x) = 0$: It is difficult to assume this assumption for this paper. Considering only one independent variable is used, there is almost certainly Omitted Variable Bias (OVB).

5. Homoskedasticity: The model and variables constructed from the data are assumed to have constant variance of u's.

4 Empirical Methods

Unlike most econometric models, this research will explore how one independent variable, health care spending as a percent of GNI, affects many dependent variables, health outcomes. The empirical model used is depicted by the equation below:

$$Y_i = \beta_0 + \beta_1 \times X_1 + u \tag{2}$$

Where Y_i is a dependent health outcome variable and X_1 is health spending as a percent of Gross National Income. The parameters of interest are Y_i and are listed below. Nine different naive models are run with the equation above.

X_1 = Health Spending percent of GNI

Y_1 = log of Death Rate

Y_2 = log of HIV

Y_3 = log of Life Expectancy

Y_4 = log of Malnutrition

Y_5 = Mortality

Y_6 = log of Infant Deaths

Y_7 = log of Maternal Deaths

Y_8 = log of Undernourished

Y_9 = Sanitation

5 Research Findings

The main results are reported in Table 3. Nine regression models are presented and include coefficients, number of observations, R2, R2 adjusted, and R2 within. When interpreting the table, focus should be given to R2 within[2]. Since this model is looking at how the dependent variable changes for each of the panel units, R2 within is the most appropriate to look at for interpretation.

Model 1, Death Rate, in Table 3 shows the relationship between independent variable ”health spending” and dependent variable log of ”Death Rate”. According to 3 and using log-level interpretation, we see that if health spending increases by 1%, death rate will increase by 0.013%. With a standard error of 0.006, health spending has a significant, positive impact on death rate. The R2 within is measured as 0.244, showing that health spending accounts for 24.4% of the variance within the relevant panel units. This conclusion is counter-intuitive, and is certainly not desirable. At the end of this section, discussion about omitted variable bias and what health spending as a percentage of GNI truly measures.

Model 2, HIV, in Table 3 shows the relationship between independent variable ”health spending” and dependent variable log of ”HIV”. According to 3, there is no significant relationship between health spending and incidence of HIV due to the large standard error (0.014) that includes 0 in its interval.

Model 3, Life Expectancy, in Table 3 shows the relationship between independent variable ”health spending” and dependent variable log of ”Life Expectancy”. According to 3 and using log-level interpretation, we see that if health spending increases by 1%, life expectancy at birth will decrease by 0.002%. With a standard error of 0.001, health spending has a small but significant,

[2]R2 within - The amount of variance *within* the panel units that the model accounts for

negative impact on life expectancy. The R2 within is measured as 0.587, showing that health spending accounts for 58.7% of the variance within the relevant panel units. This conclusion is again counter-intuitive and undesirable.

Model 4, Malnutrition, in Table 3 shows the relationship between independent variable ”health spending” and dependent variable log of ”Malnutrition”. According to 3, there is no significant relationship between health spending and malnutrition due to the large standard error (0.012) and 0.000 coefficient.

Model 5, Mortality, in Table 3 shows the relationship between independent variable ”health spending” and dependent variable ”Mortality”. According to 3, there is no significant relationship between health spending and mortality due to the large standard error (0.057) and that includes 0 in its interval.

Model 6, Infant Deaths, in Table 3 shows the relationship between independent variable ”health spending” and dependent variable log of ”Infant Deaths”. According to 3 and using log-level interpretation, we see that if health spending increases by 1%, number of infant deaths will decrease by 0.007%. With a standard error of 0.004, health spending has a small but significant, negative impact on infant deaths. The R2 within is measured as 0.592, showing that health spending accounts for 59.2% of the variance within the relevant panel units. This conclusion is the first positive outcome showing the benefit of health spending worldwide.

Model 7, Maternal Deaths, in Table 3 shows the relationship between independent variable ”health spending” and dependent variable log of ”Maternal Deaths”. According to 3 and using log-level interpretation, we see that if health spending increases by 1%, number of maternal deaths will increase by 0.007%. With a standard error of 0.006, health spending has a small but significant, positive impact on maternal deaths. The R2 within is measured as 0.452, showing that health spending accounts for 45.2% of the variance within the relevant panel units. This conclusion is the another negative outcome and will be discussed at the end of this section.

Model 8, Undernourished, in Table 3 shows the relationship between independent variable "health spending" and dependent variable log of "Undernourished". According to 3 and using log-level interpretation, we see that if health spending increases by 1%, the number of people who are undernourished will decrease by 0.010%. With a standard error of 0.009, health spending has a small but significant, negative impact on undernourished. The R2 within is measured as 0.095, showing that health spending accounts for 9.5% of the variance within the relevant panel units, considerably smaller than the other models. This conclusion is the second positive outcome showing the benefit of health spending worldwide.

Model 9, Sanitation, in Table 3 shows the relationship between independent variable "health spending" and dependent variable "Sanitation". According to 3, there is no significant relationship between health spending and sanitation due to the large standard error (0.129) and that includes 0 in its interval.

In this analysis, all Gauss-Markov Assumptions were assumed to give the best linear unbiased estimate (BLUE). However, the fourth assumption, Exogeneity $E(u|x) = 0$, is almost certainly not confirmed. There are several different reasons why this model may not be yielding optimal results, two of which are discussed below.

1. Omitted Variable Bias (OVB)[3]: In this paper, one independent X variable, health spending as a percent of GNI, was used to find relationships between a variety of health

[3] Omitted variable bias is the bias in the OLS estimator that arises when the regressor, X, is correlated with an omitted variable

outcomes, as seen in Table 3. There are two ways in which OVB may play a part in influencing the model.

(a) X is correlated with the omitted variable

(b) The omitted variable is a determinant of the dependent variable Y

Both of these are almost certainly at play in this paper. First, health spending is correlated directly with how much wealth a country has. This problem was assumed to be fixed when health spending was divided by GNI, thus eliminating the absolute aspect of the variable. However, Gross National Income is not a panacea. When controlling for a country's wealth, it fails to take into account the type of governance in the country, how much health aid is given, and a variety of other aspects of government spending. Second, omitted variables are likely also determinants of health outcomes. For example, Yemen (a country included in this analysis), has had many different types of governments in the past decades. These different types of governance have likely had an adverse impact on the health of Yemen's population. While it may not be as apparent in other countries, there are countless variables that impact the health of a population, suggesting that health spending alone is the sole determinant is naive.

2. What health care spending per capita actually mean? In short, it means how much is spent, per person, on health care. In different areas of the world, these are very different things. For example, in the United States, per capita spending for the elderly remained about five times higher than spending for children (Lassman and Catlin (2014)). In developing countries, health spending is extremely limited and unfocused, fighting random fires rather than working at prevention. Additionally, many developing countries

do not keep track of their health budgets accurately, so the data provided by the World Bank may be flawed (Raciborska and Glassman (2008)). Overall, health care spending itself is not a silver bullet due to its intrinsic volatile nature.

6 Conclusion

The main goal of this paper was to contribute statistical analysis to the health economics literature. After running nine separate fixed-effect models, five were shown to have a significant relationship with health spending as a percent of Gross National Income. Several of these significant outcomes were undesirable (e.g. death rate having a positive relationship with health spending).

After considering these false outcomes and the variety of problems with naive modeling, the results can be interpreted as inconclusive. Future analysis will center around defining health spending, diminishing its correlates, and increasing the amount of variables to decrease omitted variable bias.

References

Farag, Nandakumar A. K. Wallack S. Hodgkin D. Gaumer G., M. and C. Erbil. 2013. "Health expenditures, health outcomes and the role of good governance." *International journal of health care finance and economics* :33–52.

Hu, Bingjie and Ronald U. Mendoza. 2013. "Public health spending, governance and child health outcomes: revisiting the links." *Journal of Human Development and Capabilities* :285–311.

Lassman, Hartman M. Washington B. Andrews K., D. and A. Catlin. 2014. "US health spending trends by age and gender: selected years 2002–10." *Health Affairs* :815–822.

Martin, Rice N., S. and P. C. Smith. 2008. "Does health care spending improve health outcomes? Evidence from English programme budgeting data." *Journal of Health Economics* :826–842.

Raciborska, Hernández P., D. A. and A. Glassman. 2008. "Accounting for health spending in developing countries." *Health Affairs* :1371–1380.

Reinhardt, Peter S. Hussey, Uwe E. and Gerard F. Anderson. 2004. "US health care spending in an international context." *Health Affairs* :10–25.

WHO. 2019. "Health Expenditure Report."

Figures and Tables

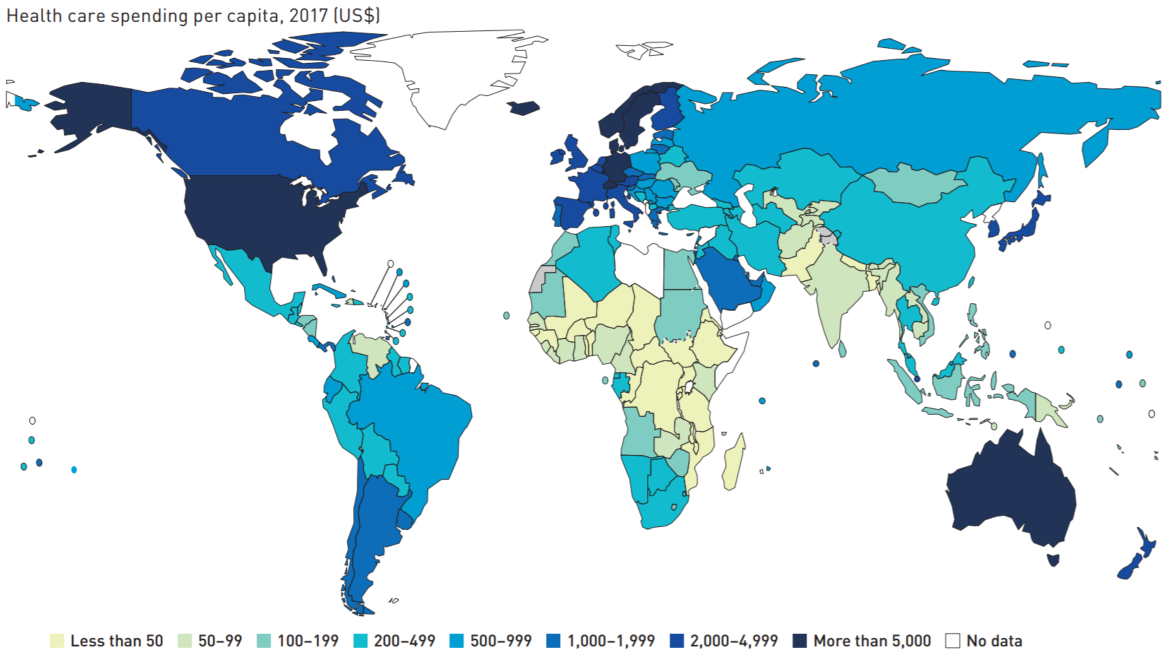

Figure 1: Health care spending per capita, 2017 (US$)
WHO (2019)

Table 1: Variable Definitions

Variable	Description
HealthSpending	Current health expenditure per capita (current US$)
GNI	GNI per capita, Atlas method (current US$)
DeathRate	Death rate, crude (per 1,000 people)
HIV	Incidence of HIV, all (per 1,000 uninfected population)
LifeExpect	Life expectancy at birth, total (years)
Malnutrition	Malnutrition prevalence, weight for age (% of children under 5)
Mortality	Mortality from CVD, cancer, diabetes or CRD between exact ages 30 and 70 (%)
InfantDeaths	Number of infant deaths
MaternalDeaths	Number of maternal deaths
Undernourished	Number of people who are undernourished
Sanitation	People using at least basic sanitation services (% of population)

Table 2: Summary Statistics

	Unique (#)	Missing (%)	Mean	SD	Min	Median	Max
GNI	2551	15	12048.6	17523.0	110.0	4403.1	121890.0
HealthSpending	4065	21	892.9	1566.2	4.5	252.2	10623.8
log.DeathRate	3154	15	2.0	0.4	0.1	2.0	3.1
log.HIV	680	51	-1.3	1.7	-4.6	-1.6	3.0
log.LifeExpect	4097	16	4.2	0.1	3.7	4.3	4.4
log.Malnutrition	293	87	2.2	1.1	-1.6	2.5	3.9
Mortality	380	83	20.0	5.5	7.8	20.2	39.0
log.InfantDeaths	3417	14				8.3	15.7
log.MaternalDeaths	471	26				5.2	13.0
log.Undernourished	606	49	14.8	2.4	11.5	14.5	20.6
Sanitation	3928	18	72.4	29.0	3.6	84.9	100.0

Table 3: Health Spending on Y_i

	Death Rate	HIV	Life Expectancy
HealthSpendPCNT	0.013	0.009	-0.002
	(0.006)	(0.014)	(0.001)
Num.Obs.	3900	2303	3878
R2	0.946	0.961	0.970
R2 Adj.	0.942	0.959	0.968
R2 Within	0.244	0.161	0.587
FE: ‘Country Name‘	X	X	X
Std. errors	Clustered (‘Country Name‘)	Clustered (‘Country Name‘)	Clustered (‘Country Name‘)

	Malnutrition	Mortality	Infant Deaths
HealthSpendPCNT	0.000	0.026	-0.007
	(0.012)	(0.057)	(0.004)
Num.Obs.	638	864	3971
R2	0.980	0.975	0.999
R2 Adj.	0.973	0.966	0.999
R2 Within	0.445	0.615	0.592
FE: ‘Country Name‘	X	X	X
Std. errors	Clustered (‘Country Name‘)	Clustered (‘Country Name‘)	Clustered (‘Country Name‘)

	Maternal Deaths	Undernourished	Sanitation
HealthSpendPCNT	0.007	-0.010	0.020
	(0.006)	(0.009)	(0.129)
Num.Obs.	3605	2548	3725
R2	0.998	0.990	0.986
R2 Adj.	0.998	0.989	0.985
R2 Within	0.452	0.095	0.410
FE: ‘Country Name‘	X	X	X
Std. errors	Clustered (‘Country Name‘)	Clustered (‘Country Name‘)	Clustered (‘Country Name‘)