

# Final Project

Bradley Rann

April 2023

## 1 Introduction

After Covid-19 the world, and particularly the US, reignited the link between politics and healthcare. The healthcare system in the US is unique for numerous reasons. The most distinctive and politically important is due to the for-profit nature throughout the entire system. Although there is nuance to what is for-profit, essentially all other developed countries have "free" government coverage (like the UK) or a way to purchase affordable insurance (like the Netherlands). Essentially private companies either can't or barely make profit on basic healthcare coverage in other countries, whereas US providers can make profit at every level.

The goal of this paper isn't a political, or moral, evaluation of different healthcare systems. The United States spends nearly double per capita than the next closest country on healthcare. It also spends a significant amount of its GDP on healthcare every year. This paper then, will have two goals. First, I will have a brief analysis of the spending and whether it is inefficient, especially compared to other countries. Second, I want to see what is driving the constant increase in healthcare expenditures in the US. From there I want to see if any of the factors like increasing obesity can predict what future healthcare expenditure might be holding other factors constant.

As I hinted at previously, healthcare is an extremely complicated and intertwined issue. No one factor determines the entire health of a country or its cost, Population dynamics, for example, is not only correlated with healthcare and who predominately needs it, but the economic development of a country as well. This is obviously an issue because economic development is also correlated with the quality of healthcare.

## 2 Literature Review

The wealth of research dedicated to this subject clarifies where some of these costs occur, but with any complex system question, there is no singular answer. Concerning for future price predictions of healthcare is that research papers going back to the early 70s have been studying the increasing cost of medical care. Most papers highlighting the increased care cost, from every decade they

were published, mention that healthcare prices are once again rising. This is obviously concerning as it indicates that prices have been consistently rising for decades. This is a figure that is easily checked and can be referred to in Figure 1 (CMS Source). One of the most popular culprits for cost increases in the US is technological innovation in the medical field. As I will detail, even this is subject to debate. Other factors include admin costs, increasing population age, and others.

There are numerous papers that claim that technological cost and the increased capabilities of the healthcare sector are a major, if not the driving, factor to increased healthcare cost. This is the contention of Harvard economist Joseph P. Newhouse who wrote prolifically on the subject in the 1990s. Newhouse proposes that there is a trade off to containing healthcare costs and continuing healthcare innovation (Newhouse 1993). The argument runs along typical economic logical lines: if the opportunity to gain profit is lost in translation while controlling prices, then there will be less incentive to innovate. Some papers produce figures that the increase in prices are anywhere from half to two-thirds of the total increase (Ginsburg 2008).

The two areas where the rate of costs have been increasing the fastest are in pharmaceutical drug prices and admin costs in private health insurance which both saw jumps from 11 to 16 percent in the 2000s (Bodenheimer 2005). Technological abilities increasing costs theoretically works with the increase price of pharmaceutical drugs, but not with the jump in increased admin cost. Moreover, Bodenheimer also argues that it is not technology, but technological diffusion that has some blame in the rising cost. There's two ways Bodenheimer argues that medical technology is diffused into the general populace, through direct to consumer advertisement (like with pharmaceutical drugs after 1988) and through specialists. The US has a greater proportion of medical specialists than most other countries. These specialists tend to advocate for bringing in new technologies at faster rates, and having better technologies can attract better specialists and hospital prestige. Bodenheimer claims this issue is much more prevalent in the United States than it is in Western Europe for example. This medical arms race is consequently raising the prices for medical care in the US. Unfortunately, there doesn't seem to be much introspection for the virtue of this technological expansion. Put another way, are patients coming out ahead of this quality of care to cost trade-off?

Continue with sources about increases in quality of care and obesity increasing costs.

### 3 Data

The data in this paper will come from multiple sources. The Centers for Medicare and Medicaid Services, a government agency, has an incredible data set charting national health expenditures since 1960. The data set not only charts total nation expenditure by the year, but subdivides it by department. Although not every department started by 1960, it details the expenditure from

categories like "total out of pocket" to departments like Veteran Affairs, Indian Health Services, workers compensation, Maternal/Child Health, and more.

Another data set I will be using is from the CDC. This data is the National Health and Nutrition Examination Survey (NHANES). This data is admittedly less neat than expenditure numbers. To start, there is no clean catalog of its numbers from NHANES I (starting in the 70s) to the current yearly survey the CDC conducts. In my final paper I will convert some of these earlier data sets to have more data, but for this rough draft I will only include the continuous data from the 2000s to now because it is easily available. NHANES is fairly massive in its scope and the data comes from different areas. The NHANES has different sections from demographic, dietary, examination, to lab data. The Demographic and dietary are surveys and estimates on how many adults are smoking, their levels of activity, and even what they eat and obviously ethnicity and age. The examination data is from doctor examinations of patients that go from height, weight, blood pressure and more. For our purposes this is the most important segment of the NHANES data. The lab data is simply urine, blood, and other medical lab related data.

## 4 Methods

In this paper I will use some of the simple techniques we have learned throughout this course. The data has been cleaned and combined using R. All code used to generate these results will be from R. To estimate the effect obesity has on national healthcare expenditure I used the following methodology: create a linear regression model for healthcare expenditure and use machine learning to estimate this cost. I believe the linear regression will be the most efficient way to estimate this as I expect a linear relationship between the rate of obesity and healthcare expenditures. As the population has higher rates of obesity, they will need more medical care for obesity-related illness.

$$Expenditure = B_0 + B_1 Obesity + E_I \quad (1)$$

The above is not the full model but a rough idea/placeholder of the final model equation.

## 5 Findings

This section will have the output of my results of my model and then discussion revolving around that. I will also write about how my results are similar, dissimilar, or unexpected considering the state of the literature. What could be added to the model to make it better and future research opportunities. I will also write about the validity of the model and other stats.

## 6 Conclusion

The conclusion will likely summarize the complicated nature of estimating figures like this in the medical/economic literature as there are too many factors for any model to account for. Discussion of my figures and results and overall success of the project will be summarized. I will also summarize why this research and future research are important along with other developments as I finish the project.

## 7 Figures

Figures will be generated from ggplot and equations to help visualize the data and then results.

## 8 Bibliography

My understanding is that figures and bibliography still need to be implemented at the end after the essay and is in conjunction with the .bib file we have to make in text citations in the paper, and those will go here when completed.