Truly Revolutionizing Healthcare: A Call to Transition from Public to Private Insurance

Conner Davis, Kushal Gourikrishna, Michael Jaweed, Dong Jian

Contents

Introduction

Access to healthcare and health insurance continues to be a hot button topic in the United States. While a segment of the population prefer the elective of private insurance market, others continue to advocate for a free and public health insurance system for all. It has been claimed that having health insurance decreases mortality rate. In 2021 a study was performed that concluded people who newly enrolled in health insurance after being uninsured saw reduced mortality over the next two year period.[1] Less has been investigated; however, in the impacts between private insurance and public insurance in terms of mortality rate. While being insured is shown to decrease mortality compared to the uninsured, is it enough for uninsured people to enroll in public healthcare or is there and impact due to the difference in mortality rate between public healthcare and private healthcare?

Politicians can influence the breakdown of insurance types in their specific county by either pushing for public insurance programs, advocating and assisting with private insurance enrollment, or if the impact is found to be insignificant, divert funds to other methods.

We would like to estimate the impact that a particular type of insurance has on mortality rate. Specifically, the data set we will be working with looks at mortality rate due to cancer. We are working with a dataset that reports the percentage of people within a particular county in the United States that has private insurance or public insurance along with a multitude of other socioeconomic factors. The model aims to estimate mortality rate due to cancer based on tuning the percentage of privately insured and publicly insured people.

Data and Methodology

The data used in the study contains health-related information and socioeconomic data aggregated from multiple sources including the American Community Survey, clinicaltrials.gov, and cancer.gov, spanning the years 2010-2016. The data is measured at the county-level for each feature, and is spread across a variety of counties across the U.S. — containing 3047 observations. We operationalized health insurance type by looking at the percentage breakdown of privately insured, publicly insured, and uninsured people within counties across the United States.

We chose cancer mortality rate as opposed to average number of cancer related deaths in a county as our dependent variable. Cancer mortality rate adjusts for population by measuring a mean death rate per capita (100,000 people). The data related to insurance types required some transformations before we could begin analysis. First, we created 2 variables— percentage of county that has both public and private, and percentage of county that is uninsured.

We estimated pct public and private by:

(1) subtracting pct of public insurance - pct of public insurance alone
(2) subtracting pct of private insurance - pct of private insurance alone
(3) averaging these two values when possible

Second, a significant portion of the variable for pct of private insurance alone was missing. We estimated these missing values by: (pct private insurance - pct public and private insurance)

Lastly, we estimated the pct of uninsured by: 1 - (pct private insurance alone + pct public insurance alone + pct public and private insurance)

Ultimately after these transformations, we based our analysis on the total percentages of private insurance, public insurance.

Exploratory plots and models were built using a 30% sub-sample of the data. Once the models and plots were finalized, the remaining 70%, totaling 2133 rows, was used to generate the final statistics and plots in this report. From exploratory plots, the distribution of cancer mortality rate was much more normal compared to the very skewed distribution average number of cancer related deaths. Hence, we continue with our analysis

[1] Jacob Goldin, Ithai Z Lurie, Janet McCubbin. "Health Insurance and Mortality: Experimental Evidence from Taxpayer Outreach" The Quarterly Journal of Economics(2021).

using the cancer mortality rate as our dependent variable. Figure 1 shows cancer mortality rate against the percent of privately insured and publicly insured people.

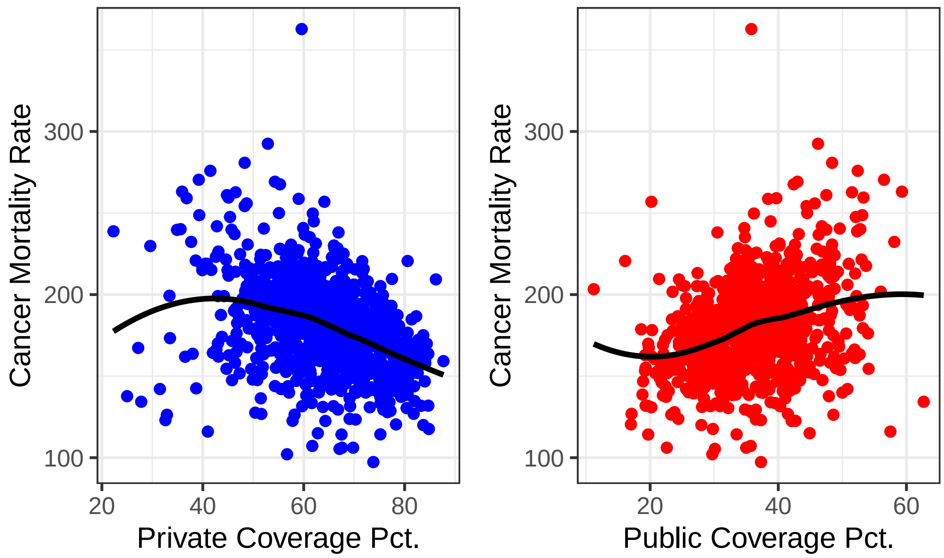

Figure 1: Cancer mortality rate as function of private and public insurance coverage

Upon initial review, with an increase in percent of privately insured people we see a decrease in cancer mortality rate and the opposite relationship is displayed for publicly insured people. Both relationships are shown to be roughly linear which allows us to continue to create some OLS regression models. We create a linear regression in the following form:

$$Cancer\widehat{Mortality}Rate = \beta_0 + \beta_1 \cdot (percentprivatecoverage) + \beta_2 \cdot (percentpubliccoverage) + \mathbf{F}\gamma + \epsilon$$

In this equation β_1 represents the change in cancer mortality (positive or negative) per 1% increase in privately insured people in the USA, β_2 represents the change in cancer mortality (positive or negative) per 1% increase in publicly insured people in the USA, $\mathbf{F}$ is a row vector of additional external factors such as poverty percentage, unemployment percentage,etc., γ is a column vector of associated coefficients and ϵ is the representation of error.

A model including the percent of uninsured people per county was tested on the exploratory data set, but due to high correlation between the three insurance percentages this model was difficult to interpret. For simplicity and interpretability, that term was dropped.

Results

Table 1 shows the results of three estimated regressions. Across all models, one of the key coefficients on private coverage percentage was statistically significant. Point estimates were relatively steady and range from -0.446 to -0.488. To provide some sense of scale, consider a county with 1,000,000 population, a 2% increase in private coverage would decrease cancer mortality rate by ~1 death per 100,000 people for a total of ~5 people in this example county.

In contrast, the key coefficient on public coverage percentage did not remain statistically significant. It seems as more and more variables were added to the model, both the statistical significance of percent public insurance and coefficient was reduced. This suggests that perhaps percent of public coverage is not one of the most important predictors on cancer mortality and that other external factors may in fact be more representative. While no statistical significance can be claimed, it is interesting to note that the coefficient for public coverage percentage transitioned from positive to negative indicating its impact on cancer mortality rate may have flipped with the introduction of more external factors.

Table 1: Estimated Regressions

	Output Variable: Cancer Mortality Rate		
	(1)	(2)	(3)
Private Coverage Pct.	-0.446^{***}	-0.468^{***}	-0.463^{***}
	(0.094)	(0.099)	(0.106)
Public Coverage Pct.	1.053^{***}	0.567^{***}	-0.033
	(0.119)	(0.104)	(0.099)
Cancer Diagnosis Rate		0.232^{***}	0.206^{***}
		(0.017)	(0.015)
Poverty Pct.		0.760^{***}	0.767^{***}
		(0.146)	(0.148)
Unemployment Pct.			0.791^{***}
			(0.225)
Pct. No High School			0.202^{**}
			(0.067)
Pct. High School (18-24)			1.051^{***}
			(0.087)
Pct. High School (Over 25)			-0.212^{**}
			(0.072)
Constant	168.667^{***}	71.079^{***}	58.361^{***}
	(9.922)	(11.547)	(11.569)
Observations	2,133	2,133	2,133
R^2	0.198	0.405	0.477
Adjusted R^2	0.197	0.404	0.475
Residual Std. Error	24.619 (df = 2130)	21.217 (df = 2128)	19.902 (df = 2124)
F Statistic	262.656^{***} (df = 2; 2130)	361.797^{***} (df = 4; 2128)	242.428^{***} (df = 8; 2124)

Note: HC_1 robust standard errors in parentheses. Note that cancer diagnosis rate is mean per capita (100,000) cancer diagnosis and cancer mortality rate is mean per capita (100,000) cancer deaths. Pct. No High School is representative of ages 18-24. Unemployment pct. is representative of people over 16.

Limitations

With large sample models, there are two assumptions that should be met: I.I.D data and having a unique BLP. The Independence assumption, which states that each data point should be independent of the others, may not hold due to potential geographic clustering or temporal autocorrelation within county-level data. In other words, neighboring counties or data from the same county across different time periods may not be truly independent. Also, one county could strategically decide on their insurance policy after seeing other counties policy implementation. On the other hand, the Identical distribution says that each data point comes from the same population distribution. The cancer mortality rate could be distributed differently in some regions than the others.

In order for a unique best linear predictor to exist, the explanatory variables must have a finite mean and a finite, non-zero variance, and if they do not have perfect colinearity. If there's heavy tails (the raw poverty percent variable has a right skew) then there's potential for non-finite variance. As long as the lm command runs, we know that we didn't end up having perfect colinearity. It's important to note that even though we gets a finite variance within our sample, the underlying distribution can still have non-finite variance.

As far as structural limitations, several omitted variables may bias this model's estimates. Omitted variables could include lifestyle factors (e.g. smoking or dietary), environmental factors (e.g. exposure to pollutants), genetic predispositions, the quality and accessibility of healthcare services, early detection practices, and the prevalence of specific types of cancers which have different survival rates, occoupation. All these factors could influence cancer mortality rates and might also be correlated with the type of health insurance (public or private). A strong limitation presented in this study describes a strong confounding relationship between private and public insurance. Upon incorporating the variable of individuals aged 25 and over possessing bachelor's degrees, the significance of public insurance in the model diminishes. This unexpected finding suggests that higher educational attainment may be interacting with insurance status in a way that was not anticipated in the study's design. The relationship between education and insurance coverage could be multifaceted and requires further investigation to understand fully. It highlights the need for a more nuanced analysis that may include considering socioeconomic factors, employment status, and regional disparities in healthcare access and education.

For example, if areas with high private insurance coverage also tend to have better access to healthcare services and we fail to include this in our model due to data unavailability, the direction of the bias is away from zero. The model might attribute the lower mortality rates to the private insurance when they are actually due to better access to healthcare.

Another example, if areas with more public insurance also have higher rates of smoking and smoking is not factored into our model, we might overestimate the mortality associated with public insurance because the model does not separate the effects of insurance type from the effects of smoking.

Conclusion

The goal of this study was to estimate the impacts to cancer mortality rate based on whether a person has private insurance or public insurance. For private insurance, our models predict that for every 2% increase in percentage of people on private insurance the cancer mortality rate will decrease by ~1 death per 100,000. In contrast, as we explore more external factors to our model, the impact of public insurance percentage did not remain statistically significant. Hence, we cannot claim in this study to draw any significant relationship between percent of public coverage and cancer mortality rate. Interestingly, however, the coefficient steadily decreased from. a positive value until it was negative. It may be worth pursuing this relationship in additional models to see if there is any significance to this sign change.

Through our research, we aspire to provide policy makers with insights into the real-world impact of insurance on individuals, specifically emphasizing the potential benefits of private insurance. Our findings lay the groundwork for recommendations aimed at truly revolutionizing healthcare. By transitioning from public to private insurance, and implementing strategies to make private insurance more accessible to those who may not ordinarily obtain it, we believe that our proposals can foster greater efficiency, personalization, and overall improvement in the healthcare system. Our goal is to help shape a future where health insurance is not just a function of coverage, but a pathway to enhanced well-being and societal progress.