New_StatsFinal

Michelle Roman

8/23/2021

Introduction

What is the relationship between self-reported health, class and the continuous variable "helpsick" which asks whether the government should help pay for medical care?

I propose that the health of the respondent would necessarily effect how high their medical bills are and consequently they might be more likely to support more government medical assistance. As such they should be negatively correlated, as the quality of the health of the respondent increases the belief in government assistance from medical care should decrease. The control variable class relates this association in that the extremes of the income brackets could have set views on how the government should be assisting citizens with medical bills despite their own health. A working-class respondent then, could be in excellent health and believe that the government should assist with medical care, while an upper-class respondent with poor health and high medical bills wouldn't necessarily think the government should help possibly because they have the resources to take care it themselves.

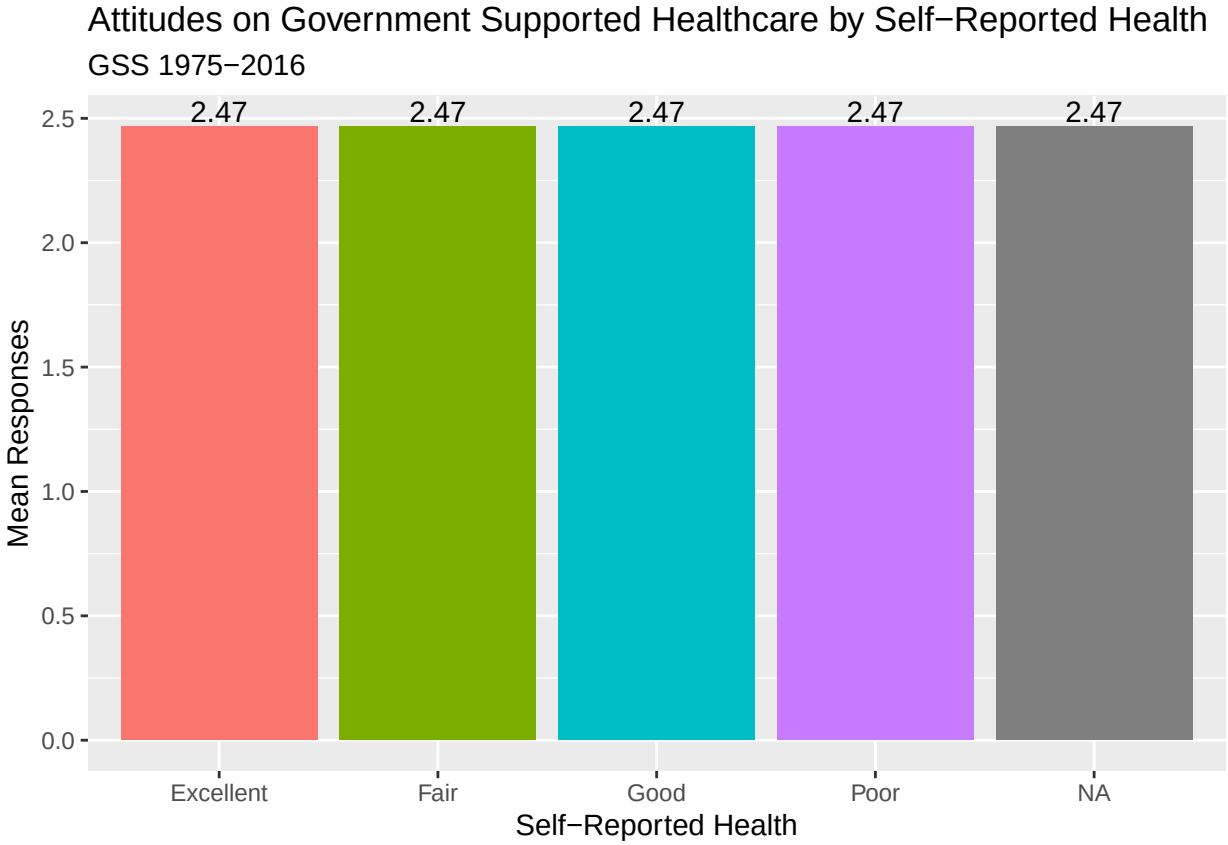
I'll be using the full set of observations from 1972-2016 and I've excluded all missing values from each variable in my analysis. The audience for this report is all professionals involved in public health policy.

Summary

The helpsick variable has a normal distribution for each possible response, with a mean of 2.47 and a median of 3 for the entire variable. Three being agree with both that the government should help cover medical costs and that people should take care of the costs themselves, with a range of possible answers from 1-5 (Govt. should help-People should care for themselves).The health variable has a strong left skew, with most respondents self-identifying as having "excellent" or "good" health. The class variable has varying distributions per the different categories.

1. This plot shows the mean responses for the helpsick question on a scale of 1-5, one being the government should help pay for medical expenses, three being agree with both, and 5 being people should take care of their own medical expenses. These means are then organized by the level of health that was self-reported by participants, ranging from Excellent to Poor. There is not a wide range of means responses for this question when self-reported health is brought into the discussion, which means there is possibly another factor that can help explain this lack of diversity across health status.

```
## Warning: `guides(<scale> = FALSE)` is deprecated. Please use `guides(<scale> =
## "none")` instead.
```



2. I've created the variable disagree within this sample to account for the proportion of respondents who believe that people should take care of their own medical expenses. I chose this variable because the bar plot revealed that responses where more in favor of government intervention. I wanted to account for the other side of the data to find a reason for the strong skew. Only 17% of respondents in this sample believed that people should be solely responsibility for their medical expenses.

```
##
##    FALSE     TRUE
## 0.820313 0.179687
```

```
##
##     FALSE      TRUE
## 0.3334337 0.6665663
```

3. I've also created a variable called excellent that only includes those who self-report having excellent health yet still believe that the government should help pay for medical expenses. 71% of the respondents in this sample fit into this category. Suggesting that my earlier hypotheses that personal medical conditions played a large role in this story, is true however it skewed opposite to my prediction.

4. This table above compares the distribution of both variables while holding class constant.What is most interesting to know about this is that middle and working class respondents have such similar responses across their income levels.

Inference

The null hypothesis is that self-reported health and responses to the helpsick question are dependent and strongly associated, also that the true difference in mean responses across health categories differs. The alternative hypotheses would be that they are independent of each other and are weakly associated, or perhaps not associated at all, and that the true difference in mean respondes across health categories is 0.

```
##
##  Welch Two Sample t-test
##
## data:  gss$helpsick[gss$health == "Excellent"] and gss$helpsick[gss$health == "Poor"]
## t = 10.229, df = 1378, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.3568843 0.5262483
## sample estimates:
## mean of x mean of y
##  2.582683  2.141117
```

I used the Welch Two sample t-test to test if the mean of the variable helpsick differs significantly at the .95 alpha level across different health categories, in this case, `Excellent` and `Poor`. The results provided a test statistic of `10.23` with a p-value that is less than `2.2e-16` and a confidence interval of `2.58-2.14`.

Then I used the Gamma Test to determent whether these ordered variables were independent of each other. The results showed a weak negative association between the responses given by responses and their self-reported class status. The test statistic was `3.5` which shows that these two factors are independent variables which allows us to reject the null hypotheses that these two are dependent and allows for other variables to help explain the phenomenon.

Regression

After regressing helpsick on self-reported health of respondents I found the following results. Using respondents with excellent health as a reference, with every change in health status from 1-4 (Excellent-Poor), the predicted response to the helpsick question decreases. For the respondents in Good health there is a .08 decrease in their response to the helpsick question, therefore the worse the health status of the respondent the more likely they are to respond with lower numbers on the helpsick scale on which 1 is coded for "Government should take action." The worse the health of the respondent the more likely they are to respond that the government should be helping cover medical expenses. All of the coefficients across health categories are significant because their p-values are less than .05 and their t-values are all more extreme than (-)1.96. Respondents with Fair health saw a .24 decrease in the likelihood that their response would be higher on the 1-5 scale, while those with Poor health saw a probability of negative .44 for going higher up on the scale. The two predicted values if someone selected Government should pay was `2.14` for those with Poor health and `2.34` for those with Fair health.

Holding class constant, we still see a negative probability that responses for the helpsick variable would increase across health categories. However, as self-reported class increases from Lower to Upper class the predicted responses for helpsick increase. All of the coefficients across health categories are significant because their p-values are less than .05 and their t-values are all more extreme than (-)1.96. A lower-class respondent with Excellent health would have a predicted response 2.17 for the helpsick question. The predicted response for respondents with Excellent health in the working class was `2.49`, for those in the middle class it was `2.66`, and for those in the upper class it is `2.77`.

When I added an interaction between health status and class rank only a few variables remained significant. The middle-class respondent with poor health and the working–class respondent with fair health each saw

Table 1: Government Spending on Medical Expenses

	Beliefs About Goverment Spending		
	1	2	3
Health	−.245***	−.190***	.042
	(.027)	(.027)	(.122)
Working	−.083***	−.050*	.132
	(.022)	(.022)	(.117)
Middle	−.442***	−.347***	−.149
	(.043)	(.043)	(.132)
Upper		.488***	.689***
		(.040)	(.103)
Health X Working		.600***	.756***
		(.064)	(.124)
Health X Middle		.318***	.472***
		(.040)	(.104)
Age X Upper			−.210
			(.129)
healthGood:classMiddle			−.232
			(.121)
healthPoor:classMiddle			−.368*
			(.149)
healthFair:classUpper			−.071
			(.196)
healthGood:classUpper			−.189
			(.163)
healthPoor:classUpper			−.266
			(.350)
healthFair:classWorking			−.263*
			(.128)
healthGood:classWorking			−.139
			(.122)
healthPoor:classWorking			−.066
			(.148)
Constant	2.583***	2.168***	1.993***
	(.017)	(.042)	(.101)
Observations	18,184	18,067	18,067
Notes:	*P < .05		
	**P < .01		
	***P < .001		

predicted decreases in possible responses of `.26` and `.37` respectively. A lower-class person with excellent health would have a predicted response of `1.99` on a scale of 1-5, a slight decrease from earlier models.

Conclusions

My hypothesis was confirmed. The association between helpsick and health was significant and as predicted upper classes were predicted to have a greater belief in people should be taking care of their own medical expenses without government expenses. However, the similarity of predicted responses leads me to believe that this is not a variable that sees much variation across class or health status lines. The highest amount of responses were for the answer agree with both (that the govt. should helpa dn people should also pay) which could possibly mean that across all of these different factors people are leaning towards government help but they could also be agreeing with both. Additional variables that I would add to this assignment would be political leanings to see whether controlling for that would explain more of the relationship between individual health and belief in government spending on health.The limitations of the current sample are that some data is unavailable for helpsick so limiting the years to gather data from would provide more accuracy. Possible directions could be to explore the ways in which class and individual health tie into other avenues of government spending and possibly in comparison to national security, risk, safety, belief in the effectiveness of vaccines, etc.