2016-05-26

The Effect of Health Insurance on Death Rates

Khorben Boyer
Western Oregon University, kboyer12@mail.wou.edu

Luke Schnee
Western Oregon University

Follow this and additional works at: https://digitalcommons.wou.edu/aes

Recommended Citation
https://digitalcommons.wou.edu/aes/52

This Presentation is brought to you for free and open access by the Student Scholarship at Digital Commons@WOU. It has been accepted for inclusion in Academic Excellence Showcase Proceedings by an authorized administrator of Digital Commons@WOU. For more information, please contact digitalcommons@wou.edu, kundas@mail.wou.edu, bakersc@mail.wou.edu.
The Effect of Health Insurance on Death Rates

Presented by Khorben Boyer and Luke Schnee
Overview of strategy

- **Question**: to determine whether the percentage of the U.S. population not covered by health insurance affects the age-adjusted death rate of the population.

- **Method**: To accomplish this research objective, cross-sectional data on all of the nation’s states was collected for a regression analysis.
Overview of strategy

- **Factors**: The data collected involved the age-adjusted death rate and the percentage of the population that is uninsured as the primary variables of interest. Whereas, the percentage of population that graduated high school, the percentage of adults who smoke, and the percentage of adults who are overweight or obese served as controls.

- **Purpose**: the results of this study would be expected to have direct implications for policymakers especially in government with respect to the current healthcare system and associated coverage.
Prior Research on similar topics and issues

- Issue has received significant attention in the last few decades.

- “Health Insurance Coverage and Mortality Revisited” by Richard Kronick examined the relation between insurance coverage and mortality.

- “Does Health Insurance Matter? Health beyond Universal Coverage” by Stephen H. Gorin reviewed the statistical and economic significance of insurance provided health care on mortality.
Prior Research on similar topics and issues

- Both papers concluded that the percentage of the U.S. population not covered by health insurance does not measurably affect the age-adjusted death rate of the general population.

- Recommended researching other possible factors such as poverty or smoking rates as more promising avenues to reducing mortality.
Data and Theory

Sample Definition

➤ Our cross-section based method utilized sample data collected from all fifty states of the U.S.

➤ This served the purpose of determining whether there were differences or variations between states with respect to key characteristics of their populations.
Data and Theory

Variable Definitions

- **Dependent Variable**: the age-adjusted death rate in deaths per 100,000 people.

- **Primary Independent Variable**: The percentage of the population not covered by health insurance.

- **Second Independent Variable** (control): The percentage of the adult population that graduated from high school.

- **Third Independent Variable** (control): The percentage of adults who smoke.

- **Fourth Independent Variable** (control): The percentage of adults who are overweight or obese.
### Descriptive Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age_Adjusted_Death_Rate_2011</td>
<td>50</td>
<td>584.8</td>
<td>956.2</td>
<td>759.784</td>
<td>85.2892</td>
</tr>
<tr>
<td>Pop_Not_Insured_2012</td>
<td>50</td>
<td>4.3</td>
<td>24.3</td>
<td>14.480</td>
<td>4.0164</td>
</tr>
<tr>
<td>Graduated_High_School_2012</td>
<td>50</td>
<td>81.4</td>
<td>92.8</td>
<td>88.000</td>
<td>3.1550</td>
</tr>
<tr>
<td>Adults_Who_Smoke_2012</td>
<td>50</td>
<td>10.6</td>
<td>28.3</td>
<td>19.836</td>
<td>3.6252</td>
</tr>
<tr>
<td>Adults_Overweight_Or_Obese_2012</td>
<td>50</td>
<td>55.7</td>
<td>69.5</td>
<td>63.880</td>
<td>3.0677</td>
</tr>
<tr>
<td>Valid N (listwise)</td>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Data and Theory

Simple Scatterplots

➢ Graph #1
Simple Scatterplots

Graph #2

5 and 43 are California and Texas
Simple Scatterplots

➢ Graph #3

44 is Utah
Data and Theory

Simple Scatterplots

➢ Graph #4
Data and Theory

Scatterplot Analysis Summary

- Overall, a linear regression analysis was performed for all of the independent variables in question due to properties of the linear model.

- Linear model is more parsimonious

- Easier to interpret implications.

- Possesses satisfactory fit and predictive results.

- Models for estimating both curves and lines simultaneously difficult to implement.
## Results

### Initial Regression Results

#### Model Summary

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R Square</th>
<th>Adjusted R Square</th>
<th>Std. Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.896a</td>
<td>0.803</td>
<td>0.786</td>
<td>39.4612</td>
</tr>
</tbody>
</table>

#### Casewise Diagnostics

<table>
<thead>
<tr>
<th>Case Number</th>
<th>Std. Residual</th>
<th>Age_Adjusted_Death_Rate_2011</th>
<th>Predicted Value</th>
<th>Residual</th>
</tr>
</thead>
<tbody>
<tr>
<td>44</td>
<td>3.111</td>
<td>699</td>
<td>576.232</td>
<td>122.7678</td>
</tr>
</tbody>
</table>
Initial Regression Results

- An $R^2$ value of .803 indicated that 80.3% of the variation in the dependent y-variable (age-adjusted death rate) was explained by the variation in the independent x variables (lack of insurance, smoking, etc.)

- The F-test had a statistical significance of less than .001 indicating a probability of all the x variables having no effect on the y variable of less than .1%
Initial Regression Results

- Regarding X outliers, we found via the mahalanobis test that no states qualified though the states of California, Colorado, and Massachusetts were partially outside the normal range.

- Regarding Y outliers, we found via case-wise diagnostics and studentized deleted residual tests that observation 44 which was the State of Utah was a clear outlier of this type.

- Furthermore, the visual test of graphing the studentized deleted residuals vs. the calculated Cook’s distance for each observation clearly indicated that Utah had singularly extensive leverage compared to all of the other states.
Results

Initial Regression Results

➢ Graph #5

44 is Utah
Outlier analysis conclusion

- Decided to remove Utah from the data set due to its strong outlier status and proceeded to perform a second regression on the remaining data to improve results.

- Utah’s outlier status and inappropriate inclusion in our data is well explained by its particular characteristics.

- Strong prohibitions against smoking.
Supplementary Regression Results (done without Utah Outlier)

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R Square</th>
<th>Adjusted R Square</th>
<th>Std. Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.923a</td>
<td>0.852</td>
<td>0.839</td>
<td>34.428</td>
</tr>
</tbody>
</table>

**ANOVA**

<table>
<thead>
<tr>
<th>Model</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>300515.2</td>
<td>4</td>
<td>75128.8</td>
<td>63.384</td>
<td>.0009</td>
</tr>
<tr>
<td>Residual</td>
<td>52152.68</td>
<td>44</td>
<td>1185.288</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>352667.9</td>
<td>48</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Supplementary Regression Results

### Coefficients

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
<th>Collinearity Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Constant)</td>
<td>607.426</td>
<td></td>
<td>2.317</td>
<td>0.025</td>
<td></td>
</tr>
<tr>
<td>Pop_Not_Insured_2012</td>
<td>1.231</td>
<td>0.058</td>
<td>0.8</td>
<td>0.428</td>
<td>0.633 1.579</td>
</tr>
<tr>
<td>Graduated_High_School_2012</td>
<td>-5.923</td>
<td>-0.218</td>
<td>-2.886</td>
<td>0.006</td>
<td>0.588 1.701</td>
</tr>
<tr>
<td>Adults_Who_Smoke_2012</td>
<td>16.836</td>
<td>0.669</td>
<td>7.043</td>
<td>0</td>
<td>0.372 2.685</td>
</tr>
<tr>
<td>Adults_Overweight_Or_Obese_2012</td>
<td>4.991</td>
<td>0.173</td>
<td>1.772</td>
<td>0.083</td>
<td>0.353 2.834</td>
</tr>
</tbody>
</table>
Supplementary Regression Results

- An even higher $R^2$ value of .852 indicated that 85.2% of the variation in the dependent y-variable (age-adjusted death rate) was explained by the variation in the independent x variables (lack of insurance, smoking, etc.)

- Again, The F-test had a statistical significance of less than .001 indicating a probability of all the x variables having no effect on the y variable of less than .1%
Results

Independent Variable Regression Results

- The percentage of population not covered by health insurance variable was found to not be statistically significant and thus the null hypothesis of no effect on the age-adjusted death rate could not be rejected. Also, even if the effect was found to be statistically significant, it would be a small response and the smallest relative to the other independent variables.

- The percentage of adult population graduated from high school variable was found to be statistically significant at the 1% level and thus the null hypothesis of no effect on the age-adjusted death rate was rejected. The effect was moderate but it is the second largest relative to the other independent variables.
Results

Independent Variable Regression Results

- The percentage of adult population who smoke variable was found to be statistically significant at the 1% level and thus the null hypothesis of no effect on the age-adjusted death rate was rejected. The effect was large and the largest relative to the other independent variables.

- The percentage of adult population who are overweight or obese was found to be statistically significant at the 10% level and thus the null hypothesis of no effect on the age-adjusted death rate was rejected. The effect was small and the second smallest relative to the other independent variables.
Results

Possible Regression flaws or assumption failures?

- **Omitted Variable bias**
  - Frequency with which policy holders used their insurance and how it compared to people who do not use their insurance could not be determined.
  - Research only examined individuals with insurance at the time of the incident and not as a long term program in line with the definition given for the insurance coverage variable.

- **Multicollinearity**
  - Collinearity diagnostic indicated that multicollinearity is not a significant factor in the overall regression analysis.

- **Sample Selection Bias**
  - All fifty states were included in the initial regression analysis thereby allowing a full representation of the entire nation by state.

- **Simultaneous Equations**
  - Analysis of the relation between dependent and independent variables did not indicate that the possible simultaneous relations to have much significance in terms of the total effect on the variables concerned. Specifically, it was concluded that the selected x variables each affect the y variable much more strongly than the y variable affects any of the x variables.
Results

Heteroskedasticity

➢ Graph #1
Results

Heteroskedasticity

Graph #2

![Graph](image-url)
Results

Heteroskedasticity

➢ Graph #3
Results

Heteroskedasticity

➤ Graph #4
Heteroskedasticity: remark on cluster

- Note that the above two graphs share the same tight, offset cluster of five observations at the upper right with labels of 1, 24, 34, 36, and 41.

- These are the states of Alabama, Mississippi, North Dakota, Oklahoma, and South Dakota respectively.

- This anomaly may be due to particular cultural features, differences in tobacco regulation or taxation, the dominance of fried food consumption, etc.
Results

Normality of Residuals

Graph #5

Histogram

Dependent Variable: Age_Adjusted_Death_Rate_2011

Mean = 2.97E-15
Std. Dev. = 0.957
N = 49

The five states
Results

Conclusion

➢ It was found that the percentage of the population not covered by health insurance did not affect the age-adjusted death rate in a statistically significant manner much less in an economically significant fashion.

➢ At least one potential problem was identified in the analysis involving the possibility of omitted variables bias with respect to our health insurance variable. Furthermore, the histogram of the residuals demonstrated a possible rightward skew that was considered to be negligible owing to our small sample size.

➢ Otherwise, the result for our health insurance coverage rate variable is comparable to what was found by the other studies covered in the literature review and does not suggest the need for further research.

➢ In contrast, it was found that the graduated high school rate, adult smoking rate, and adult overweight or obese rate variables all had both statistical and economic significance.

➢ For policy makers and the government, this implies that reductions in the age-adjusted death rate would best be served by addressing these other variables of interest rather the current extent of health insurance coverage throughout the U.S.
Important Works Cited


