Where The Healthy People Are

Nathan Venos and Bradley Sawyer

Adjusted Community Rating Under ACA

- Since January 1, 2014, health insurance companies could only use Adjusted Community Rating (ACR) for pricing all health plans sold on the individual and small group insurance markets (link)

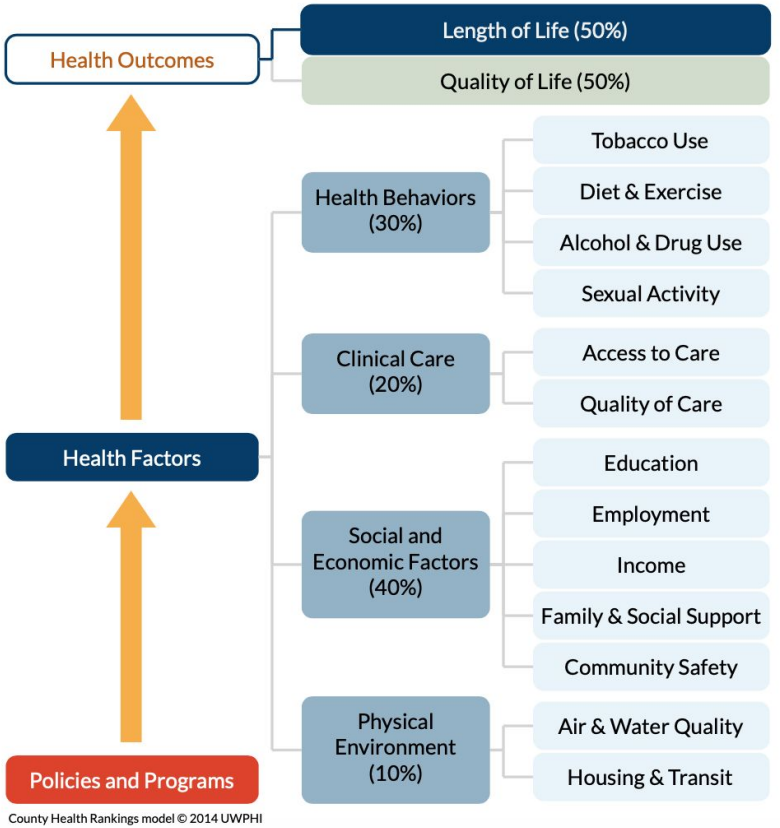
- The ACR provision allows for premium rates for a given benefits plan to vary based on only 4 criteria:
  - Plan size (individual or family)
  - Age (within a 3:1 ratio for older adults to younger adults)
  - Tobacco use (within a 1.5:1 ratio for smokers to non-smokers)
  - **Geographic area**

- Currently, 46 different US states and territories use counties as the ACR geographic boundary (link)

County Health Data

The Robert Wood Johnson Foundation and University of Wisconsin maintain the County Health Rankings & Roadmaps program which is:

"...based on a model of community health that emphasizes the many factors that influence how long and how well we live. The Rankings use more than 30 measures that help communities understand how healthy their residents are today (health outcomes) and what will impact their health in the future (health factors)." (link)



Business Question

- Generally: What are the counties within the United States that have the healthiest populations and therefore should have lower health insurance premiums for the succeeding plan year?

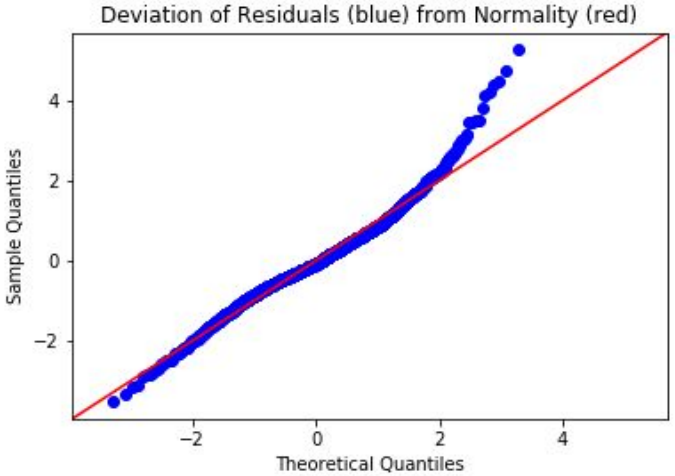
- Specifically: What county level phenomena are most associated with population health when measured in terms of Years of Potential Life Lost before age 75 per 100,000 population (aka Premature Death)?

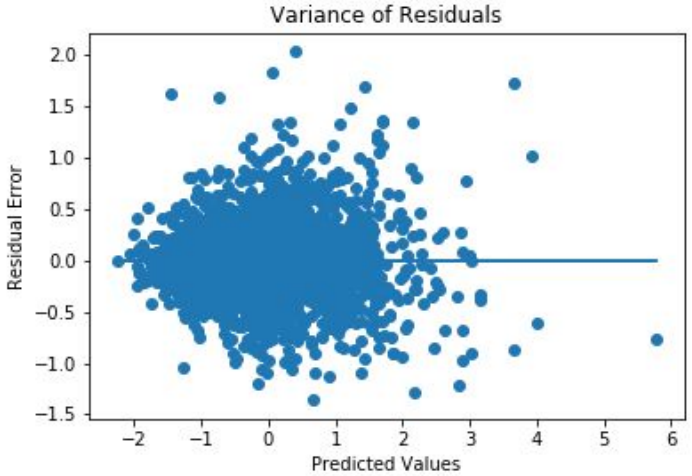
Linear Regression Model

- Initial regression model had 30 features and fit quality (adjusted R-square) of 0.848
  - 8 features were not statistically significant ($p<0.05$)
- Next step was to standardize data to account for difference in scale
- Used an iterative penalty estimation (LARS, L1) to eliminate un-predictive features
  - Total remaining features were 15
- Repeated the iterative process after including interactions between the remaining features.
- Final model of 13 features and 2 interactions predicted data with an adjusted R-squared of 0.849 and a value of 0.837 on new unseen data.

Robustness Checks

- Normality Assumption
  - Final model's errors are within acceptable range
- Homoscedasticity Assumption
  - Final model's errors are reasonably consistent
- Non-Collinearity Assumption
  - Final features variance inflation factors were all below 3 (above 5-10 is problematic)





Important Factors

Health Behavior:
Adult Smoking
Food Environment Index
Physical Inactivity
Excessive Drinking
STD Rate
Teen Birth Rate

Environmental:
Air Pollution - Particulate Matter
Driving Alone to Work
Long Commute of Driving Alone

Social/Economic:
Unemployment
Income Inequality
Children in Single Parent Household Rate
Injury Deaths

Interactions:
Physical Inactivity & Income Inequality
Teen Births & Injury Deaths

Takeaways

- County-level health can be accurately predicted with a relatively small subset of factors

- Many of these are not directly related to the local healthcare system, but socio-economic and behavioural phenomena

- Insurance providers should focus on using this information for identifying healthy and unhealthy populations for geographic premium rate adjustments

Questions?