HEALTH-RELATED BEHAVIORS AND THE EFFECT ON CHRONIC CONDITIONS

Tyler Billings

Abstract

A Poisson count model regression using the 2017 US National Health Interview Survey show that health related behaviors have varying levels of success in reducing the likelihood in the number of non-transmissible chronic conditions. By decreasing negative health related behaviors such as smoking and BMI, increasing positive health related behavior like an individual's physical activity, we find that the expected number of chronic conditions decreases. However, it was observed that alcohol consumption and light/moderate physical activity has a minimal effect on decreasing these chronic conditions. The results show that the health-related behaviors with the greatest effect on reducing the number of chronic conditions which should be the priority for a physician's recommendation is that individuals should stop smoking and to increase the frequency of vigorous physical activity.

1. Introduction and Review of the Literature

A disease or medical issues that persists for multiple months and likewise takes months to treat, if curable at all, is considered to be a chronic condition. Chronic conditions and diseases affected nearly 60% US of adults (18 years and older) and those with multiple chronic conditions make up 40% of the adult population demonstrated by Buttorff, Ruder, and Bauman (2017). This represents a significant increase from a study by Anderson and Horvath (2004) which found that in 2000, 45% of American adults with a chronic condition and 21% had more than one. These rates only increase for those individuals 65 years and older, with 80% afflicted by one chronic condition and 50% having more than one; these chronic conditions account for 61% of deaths in the elderly population in a year as discovered by the CDC (2007).

Health-related behaviors and access to preventative healthcare resources have long been established to be determinants of a person's health and life span. Humphreys, McLeod, and Ruseski (2014) show that the use of moderate physical activity as a reliable method for reducing symptoms of poor health has been widely recognized by healthcare professionals. Though there is research linking specific activities to chronic conditions and diseases, there was a lack of literature that looked at physical activity, diet, smoking, and drinking alcohol from a wholistic view. This requires further investigation to determine the effects of health-related behaviors and how they affect an individual's likelihood of suffering from non-transmissible chronic conditions.

It comes down to the doctors, who are responsible for providing advice and guidance on a person's well-being. Not necessarily treating one specific condition or disease, but best practices that can be used to prevent future illnesses. Oftentimes, a doctor will have a patient with which they have multiple recommendations on how they can improve their quality of life, their general health and functionality. These behaviors can be difficult to change, many of which are outright addictive which has been shown in eating habits by Fan and Jin (2014) and Richards (2007). If a doctor could identify the most detrimental or beneficial health-related behavior, they could then prioritize lifestyle changes. This course of action would allow patients to see the greatest improvements with the least disruptive adjustments.

Individuals can participate in health-related behaviors that can be classified as either positive or negative. The "positive" health-related behaviors are expected to reduce the number of chronic conditions that occur in an individual. Colman and Dhaval (2013) demonstrate that physical activity is inversely related to heart disease and hypertension. Sari (2014) had also shown that individuals who participate in moderate to vigorous physical activities recover quicker from hospital visits than those who do not. Specifically, individuals with chronic conditions saw a decrease in hospital stays by 24.8% or 22.5% for individuals participating in vigorous or moderate physical activity, respectively. Additionally, using the 2006 Canadian Community Health Survey (CCHS) Humphreys, McLeod, and Ruseski (2014) conducted a study of the effect of various levels of physical activity on high blood pressure, heart disease, diabetes, asthma, and arthritis which showed convincingly that individual's that participated in either type of physical activity have a smaller number of chronic conditions.

An individual can also participate in "negative" health-related behaviors which contribute towards an increase in the number of chronic conditions. Yang, Bishai, and Harman (2008) used the U.S. Cost and Use files from the Medicare Current Beneficiary Survey (MCBS) to show that BMI is overwhelmingly a contributing factor towards increasing the likelihood of chronic conditions, limited functionality, and an overall decrease in the quality of life. Himes (2000), while focusing on older populations, find that functional limitations increased significantly in individuals that are obese. Additionally, the US Department of Health and Human Services (2014) shows that smoking is related to several chronic diseases to include kidney disease, liver disease, cardiovascular disease, and raspatory diseases. Westman et al. (2015) links a shorter lifespan for those people who suffer from alcohol use disorders.

Other studies have reviewed people's lifestyles holistically, to try and consider several variables to better identify and model all significant negative health-related behaviors. Balia and Jones (2008) found that smoking, drinking, sleep patterns, obesity, and physical activities contributed significantly to increased premature mortality, accounting for 25.3% of occurrences while social class, education, occupational status, ethnicity and height accounted for less than 1% each.

While these journal articles have significant results which this study will use as the base and build off, they do leave gaps in our understanding of how health-related behaviors affect the occurrence of chronic conditions. Himes (2000) and Yang, Bishai, and Harman (2008) focused their sample population on the elderly (65 years or older). With the exception of Balia and Jones (2005), the other studies only focused on one dimension, either physical activity or obesity. While Balia and Jones (2008) used a multivariate model to include a wide range of health-related behaviors they primarily tested to see how mortality was affected and not various chronic conditions. There exists an opportunity to study how health-related behaviors affect the occurrence of non-genetic chronic conditions across adults (18 years or older). By understanding how the various positive and negative health-related behaviors affect chronic conditions in a holistic manner, doctors will be able to best craft a treatment plan or advice on which lifestyle changes should be made first to provide the largest benefit to the patient.

All the literature reviewed has shown results that align with the premise that when positive health-related behaviors are present there will be a decrease in the occurrence of chronic conditions or their symptoms, as well as negative health-related behaviors contributing to the likelihood of chronic conditions and their symptoms being reported. Not only do these studies demonstrate that just the presence of these health-related behaviors dictates the likelihood of chronic conditions, but the level at which the health-related behaviors occur is an important distinction. Both Humphreys, McLeod, and Ruseski (2012) and Sari (2014) observed in their results that there was a clear diminishing benefit to increasing physical activities beyond the moderate level, as it related to chronic conditions. Himes (2000) also found that that there was a difference between underweight, overweight, and obese individuals. With these demonstrations it will be important to not only measure the variables in the National Health Interview Survey (NHIS) in a binary fashion, but to distinguish degrees of occurrence when possible. Himes

(2000) also observed that there was a difference between men and women in the severity of the effects from obesity, it will be important to hold constant for the various socio-economic factors that are tracked within the NHIS. All the data sets from the literature reviewed above are from western cultures (e.g. Canada, the United Kingdom, and the United States) which have similar culture norms when it comes to health-related behaviors and useful in testing hypotheses which use a U.S. data set.

We find that these positive and negative health-related behaviors are significantly linked to an individual's quality of life in respect to the number of chronic conditions, with the only exception being moderate workouts. The impact of each health-related behavior varied significantly with smoking having the largest marginal effect and the frequency of vigorous physical activity having the greatest potential overall effect on decreasing the number of chronic conditions. Although moderate workouts proved to be not statistically significant in this study, this contradiction of prior literature could be explained by the NHIS inclusion of light physical activity in the variable along with an unclear definition of light, moderate, and vigorous workouts. Overall, my results support the findings in the literature that link the various health-related behaviors with one or more of the chronic conditions.

The rest of the paper is structured as follows. In the next section, we discuss the data in detail and describe the empirical specifications we use. The estimation results are analyzed in section 3 and concluding remarks follow in section 4.

2. Data and Empirical Model

As mentioned earlier, the data used in our analysis is the 2017 NHIS annual survey, the principle source of information on the health of civilians in the United States. This sample provides us with individual level data on 26,742 adults, gathered from 32,617 households across the United States. This analysis uses a cross-sectional methodology, the largest driving factor for this is the fact that the survey does not revisit individuals. Therefore, only benefit for assembling panel data from additional survey years is to increase the population size, which is unnecessary considering the large population sample. Additionally, the questions asked in the survey change over the years, which can change the dependent or independent variables, making the best cross-sectional data the best choice available.

The nature of some of the questions asked during the survey do not identify any timeframe for when an activity or condition took place. In order to reduce false positives from associating unrelated time periods to health-related behavior currently being taken, the variables have been chosen to include only those that have occurred within the past 12 months. Through this data set we will be looking at both negative and positive health-related behaviors tracked in the survey and comparing them to the number of identified chronic conditions. The general model being proposed has the dependent variable as the number of chronic conditions reported by an individual. This dependent variable will be a function of a set of independent variables that make up a person's health-related behaviors. Additionally, there will be included a set of control variables to account for sex, race, and age of an individual.

```
##
## Table 1. Descriptive Statistics (Tidied Data)
## ==========================================================================
## Statistic            N      Mean    St. Dev.  Min  Pctl(25) Pctl(75)  Max
## --------------------------------------------------------------------------
## NUMBER_CONDITIONS  25,818   0.662    0.905     0      0        1       7
## BMI                25,818 2,809.217 643.363  1,203  2,366    3,126   9,668
## ALCHOL_CONSUMPTION 25,818   0.225    0.418     0      0        0       1
## SMOKER             25,818   0.151    0.358     0      0        0       1
## MODERATE_WORKOUT   25,818   3.087    3.933     0      0        5      28
## VIGOROUS_WORKOUT   25,818   1.933    3.406     0      0        3      28
## --------------------------------------------------------------------------
```

The dependent variable (NUMBER_CONDITIONS) being studied will be the summation of negative health symptoms a person had identified as suffering from within the survey. If an individual answered "yes" to any of the conditions, then the dependent variable was increased by one. This gives us a dependent variable that is comprised of a few non-negative integers, which included zero. Our dependent variable, NUMBER_CONDITIONS, had 25,818 observations. This variable is 0 for 15,172 (59%) people interviewed in the sample and the maximum value we find is 7 with only 49 individuals having 5 or more chronic conditions. This comes out to the average of .662 chronic conditions a person has observed by the mean for NUMBER_CONDITIONS in Table 2. The conditions representing high cholesterol, hypertension, and pain limitations had missing values when compared to the entire population. On further investigation it was determined that they each had a preceding question which asked if they had ever had this symptom, if answered negatively then they were not asked similar questions around the previous twelve months. All those absences were treated as having answered as "no" to the relevant variables. The conditions were further narrowed down to only those that were identified as having been present in the past 12 months. The following chronic conditions have been identified in the prior literature as being related to one or more of the identified independent variables. HIGH_CHOLESTEROL (please refer to the data appendix for the NHS variable names) was counted if the person had answered that they had high cholesterol within the past 12 months, this has been linked to individuals that are obesity (BMI >30) in Martin, Schoeni, and Andreski's (2010). Humphrey, McLeod, and Ruseski (2014) found associations of physical inactivity with HYPERTENSION and so it was included if a person had hypertension within the past 12 months. The U.S. Department of Health and Human Services has found a connection with liver conditions and smoking, so CON_LIVER was included if a person had a liver condition within the past 12 months. Among several other chronic conditions Colman and Dave (2013) identify kidney conditions as being linked to physical activity and so, CON_KIDNEY was counted if a person had weak or failing kidneys within the past twelve months. Colombo, Rotondi, and Stanca found that physical activity had a dampening effect on if an individual had ulcers, so we include the variable ULCER if a person had an ulcer within the past 12 months. BRONCHITIS was counted if an individual had a chronic bronchitis within the

past 12 months, which has be linked with smoking by Harris and Harris (1996). Likewise, ASTHMA was counted if an individual suffered from an asthma attack within the past 12 months, similarly identified by Harris and Harris (1996).

In addition to the previous conditions, a second regression was performed that included timeframes that did not align with the previous variables' 12-month period. These new chronic conditions are still representative of an individual's health state and their health-related behavior because all the variables are still within a 12-month timeframe. JOINT_PAIN joint pain/aching had occurred within the past 30 days, LIMITING_PAIN was counted if an individual had enough pain that limited them in any way within the past 6 months, and FUNC_LIMIT if the individual had any functional limitations. These additional conditions are noted as being closely related to obesity through Himes' (2000) study.

The negative health-related behaviors will include several activities that, based on prior literature and the alternate hypothesis, will contribute to an overall increase in the number of chronic conditions a person suffers from. The first of these behaviors represented in the model is ALCHOL_CONSUMPTION equal to one, which denotes individuals that considers themselves a moderate or heavy drinker. If the participant smokes tobacco, some or each day, then the variable SMOKER is equal to one. Smoking and drinking have been found to be related to several chronic conditions and to mortality overall as we see in Balia and Jones' (2005) mortality study. The final negative health-related behavior is the individual's body mass index (BMI). The use of BMI is to represent an individual's diet choices, which are absent from the NHIS data. Sari (2014) and Yang, Bishai, and Harman (2008) link this variable as a cause for an individual's ability to recover from injuries and chronic conditions. Additionally, Himes (2000) identifies individuals who are overweight and obese are more likely to suffer from functional limitations and Humphreys, McLeod, and Ruseski (2014) link obesity as a risk factor that contribute to some chronic conditions and should therefore be considered as an independent variable. Looking at Table 1 we find the mean for each of these variables, which represent the percentage of the population that participates in the health-related behavior. We see that the mean BMI is 2,809, indicating that the average population is already considered overweight. Additionally, 22.5% identify as moderate or heavy drinks while only 15.1% of the sample population participate in smoking.

The other set of independent variables are positive health behaviors that are expected to be linked to a decrease in the occurrence of chronic conditions. The positive independent variables are the participation levels of physical activity at a weekly frequency with either moderate or vigorous intensity (MODERATE_WORKOUT and VIGOROUS_WORKOUT, respectively). These variables indicate the number of times they participate in either of the two types of workouts on a range between 0 and 28 times per week. Humpherys, McLeod, and Ruseski (2012) show a clear inverse relationship between numerous chronic conditions and various levels of physical activity, but there is a diminishing return to the intensity and frequency of the workouts preformed. Due to the diminishing returns we model them in a quadratic format, including the square of each variable. Reviewing the data in table 1 we find that on average

people participate in moderate workouts 3 times a week with the 75th percentile participating in it 5 times a week. Looking at vigorous workouts that average is slightly lower at 2 times a week with the 75th percentile still only participating in vigorous workouts 5 times a week.

Incorporated in the model are variables that we are controlling for, the sex of an individual, if an individual is elderly (65 years or older), and the race of the individual. Himes (2000) found that obesity has a greater effect on the occurrence of functional limitation in elderly individuals. Sari (2014) observed that there is a difference between males and females when it comes to recovering from chronic conditions during hospital stays. Boustan and Margo (2014) identified that there are still significant racial inequalities when it comes to individuals' health.

For a detailed breakdown of the variables and observations used within this study please refer to the Data Appendix.

```
                                   ## Figure 2. Correlation Matrix
## ================================================================================================
##                     NUMBER_CONDITIONS  BMI   ALCHOL_CONSUMPTION SMOKER MODERATE_WORKOUT VIGOROUS_WORKOUT
## ------------------------------------------------------------------------------------------------
## NUMBER_CONDITIONS           1         0.231        -0.064       0.040       -0.069           -0.132
##         BMI                0.231         1           -0.078       -0.022      -0.054           -0.076
## ALCHOL_CONSUMPTION       -0.064       -0.078          1           0.084       0.048            0.074
##       SMOKER              0.040       -0.022       0.084          1          -0.031           -0.044
##  MODERATE_WORKOUT        -0.069       -0.054       0.048        -0.031         1              0.285
##  VIGOROUS_WORKOUT        -0.132       -0.076       0.074        -0.044      0.285               1
## ------------------------------------------------------------------------------------------------
```

Looking at the correlation matrix above we see that there is a little to no correlation between the independent variables in our model. The greatest correlation is between moderate workouts and vigorous workouts but even that is only 0.285 This shows there is no significant multicollinearity, let alone perfect collinearity, and we will not need to adjust the model to account for interacting independent variables.

Research by Basu (2015) showed that childhood health had long term effects on the occurrence of chronic conditions later in life while using a poisson count model. Additionally, a more advanced, zero-inflated Poisson (ZIP) model, was used to count the number of days and individual stayed in a hospital based on their level of physical activities and prior chronic conditions in a study by Sari (2013). Similar to the prior research, our dependent variable only consists of small non-negative integers and zero.

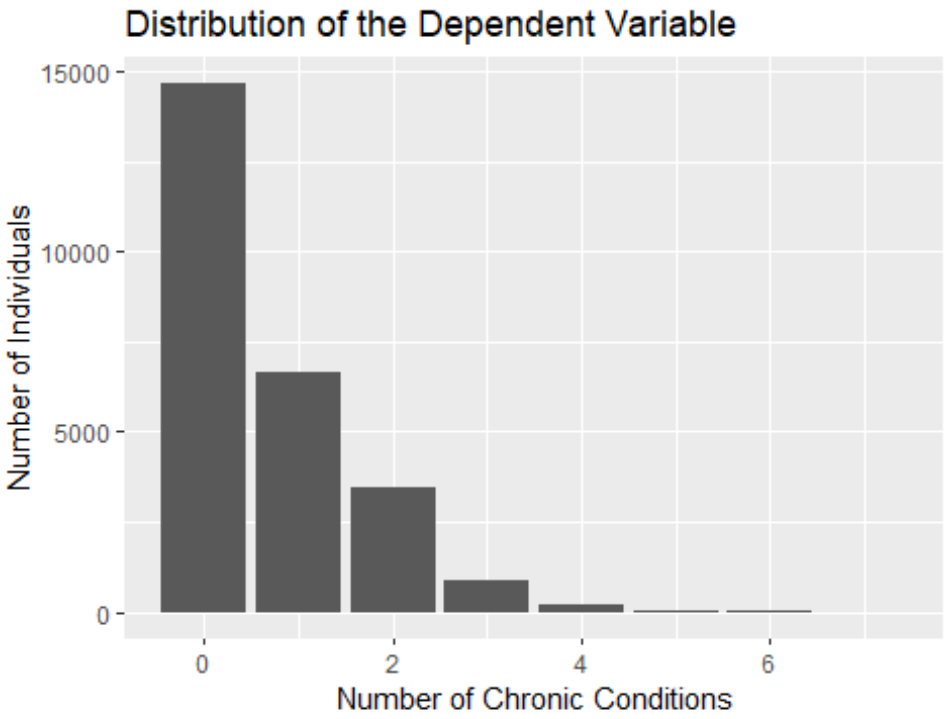

Figure 3.

Based on these facts and the work done by Wooldridge (2020) and Heiss (2016) we conclude that the best model to conduct our analysis is the Poisson count model.

$$E(NUMBER\_CONDITIONS|x_k) = \exp(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k)$$

$$\log[E(y|x_1, x_2, \ldots, x_k)] = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k$$

However, we know that all standard count data distributions exhibit heteroskedasticity discussed by Wooldridge (2020). In order to account for this, we use the quasi-maximum likelihood estimation which provides us with the robust standard errors. Furthermore, looking at the chi squared test and the log likelihood we find that we can reject the null hypothesis (that the model is not statistically significant) and we find that the model is significant at the 1% level.

3. Analysis of Results

The results of the Poisson count regression model are shown in Figure 4 which lists out the independent variables and the coefficient of their effect on the dependent variable, the number of chronic conditions. It is important to note that the Poisson regression results in log of the dependent variable, apart from BMI all the independent variables are interpreted in a log-level form. The table has four different columns, the first column contains the OLS estimates for the smaller set of chronic conditions. The second column is the Poisson estimate and all the chronic conditions that have the 12-month timeframe gathered from the NHIS data set. The third

column is the quasi-maximum likelihood estimation to account for heteroskedasticity of the smaller set of chronic conditions. The fourth and final column is another Poisson quasi-maximum likelihood estimation count regression, however, it includes the additional conditions, each of which have timeframes shorter than 12-months which are described in Table 4

Table 4. Regression Results

	Dependent variable:			
	NUMBER_CONDITIONS			
	OLS	Poisson	glm: quasipoisson link = log	
	(1)	(2)	(3)	(4)
ALCHOL_CONSUMPTION	-0.043***	-0.069***	-0.069***	-0.059***
	(0.013)	(0.020)	(0.021)	(0.014)
SMOKER	0.184***	0.280***	0.280***	0.319***
	(0.015)	(0.021)	(0.022)	(0.014)
log(BMI)	0.983***	1.388***	1.388***	1.021***
	(0.025)	(0.034)	(0.036)	(0.024)
MODERATE_WORKOUT	-0.014***	-0.019***	-0.019***	-0.008***
	(0.003)	(0.004)	(0.004)	(0.003)
MODERATE_WORKOUTsq	0.001***	0.001***	0.001***	0.001***
	(0.0001)	(0.0002)	(0.0002)	(0.0001)
VIGOROUS_WORKOUT	-0.037***	-0.072***	-0.072***	-0.072***
	(0.003)	(0.005)	(0.006)	(0.004)
VIGOROUS_WORKOUTsq	0.001***	0.002***	0.002***	0.002***
	(0.0002)	(0.0003)	(0.0003)	(0.0002)
MALE1	-0.057***	-0.057***	-0.057***	-0.170***
	(0.011)	(0.016)	(0.017)	(0.011)
ELDERLY1	0.587***	0.783***	0.783***	0.533***
	(0.012)	(0.016)	(0.017)	(0.011)
BLACK1	0.058***	0.085***	0.085***	-0.028
	(0.017)	(0.023)	(0.024)	(0.017)
ASIAN1	0.020	0.004	0.004	-0.169***
	(0.024)	(0.039)	(0.041)	(0.029)
AIAN1	0.029	0.060	0.060	-0.037
	(0.049)	(0.072)	(0.075)	(0.051)

```
##
## MULTIPLERACE1                   0.009              0.009      0.009    0.106***
##                                (0.037)            (0.056)    (0.059)    (0.037)
##
## OTHERRACE1                     -0.048             -0.066      -0.066    -0.205
##                                (0.116)            (0.174)    (0.183)    (0.129)
##
## Constant                      -7.198***          -11.626***  -11.626*** -7.495***
##                                (0.198)            (0.277)    (0.291)    (0.196)
##
## ------------------------------------------------------------------------------
## Observations                   25,818             25,818      25,818    25,818
## R2                              0.161
## Adjusted R2                     0.160
## Log Likelihood                                  -26,483.430
## Akaike Inf. Crit.                               52,996.860
## Residual Std. Error     0.829 (df = 25803)
## F Statistic         352.759*** (df = 14; 25803)
## ==============================================================================
## Note:                                              *p<0.1; **p<0.05; ***p<0.01
```

The required level of significance assigned to the model and the variables was set at 1%. The overall model significance came in with a very small value for the p-value, this being well within the assigned significance level of 1% we can reject the null hypothesis and state that the identified health-related behaviors and other factors do have explanatory power over the number of chronic conditions a person is suffering from.

```
## Figure 5

## Call:
## poissonmfx(formula = NUMBER_CONDITIONS ~ ALCHOL_CONSUMPTION +
##     SMOKER + log(BMI) + MODERATE_WORKOUT + MODERATE_WORKOUTsq +
##     VIGOROUS_WORKOUT + VIGOROUS_WORKOUTsq + MALE + ELDERLY +
##     BLACK + ASIAN + AIAN + MULTIPLERACE + OTHERRACE, data = Adult.Pop2.omit,
##     atmean = FALSE)
##
## Marginal Effects:
##                           dF/dx   Std. Err.         z     P>|z|
## ALCHOL_CONSUMPTION -0.04488154  0.01283825  -3.4959 0.0004724 ***
## SMOKER              0.20485337  0.01680440  12.1905 < 2.2e-16 ***
## log(BMI)            0.91958544  0.02382835  38.5921 < 2.2e-16 ***
## MODERATE_WORKOUT   -0.01243113  0.00268573  -4.6286 3.682e-06 ***
## MODERATE_WORKOUTsq  0.00052286  0.00013887   3.7650 0.0001665 ***
## VIGOROUS_WORKOUT   -0.04749826  0.00357963 -13.2690 < 2.2e-16 ***
## VIGOROUS_WORKOUTsq  0.00163160  0.00017550   9.2967 < 2.2e-16 ***
## MALE1              -0.03796787  0.01039586  -3.6522 0.0002600 ***
## ELDERLY1            0.59435040  0.01418186  41.9092 < 2.2e-16 ***
## BLACK1              0.05835474  0.01637467   3.5637 0.0003656 ***
## ASIAN1              0.00288868  0.02617638   0.1104 0.9121282
```

```
## AIAN1               0.04074843  0.05048135   0.8072 0.4195526
## MULTIPLERACE1       0.00565710  0.03760429   0.1504 0.8804193
## OTHERRACE1         -0.04238027  0.10820152  -0.3917 0.6952954
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## dF/dx is for discrete change for the following variables:
##
## [1] "ALCHOL_CONSUMPTION" "SMOKER"             "MALE1"
## [4] "ELDERLY1"           "BLACK1"             "ASIAN1"
## [7] "AIAN1"              "MULTIPLERACE1"      "OTHERRACE1"
```

```
## Figure 6

##Call:
## poissonmfx(formula = NUMBER_CONDITIONS ~ ALCHOL_CONSUMPTION +
##     SMOKER + log(BMI) + MODERATE_WORKOUT + MODERATE_WORKOUTsq +
##     VIGOROUS_WORKOUT + VIGOROUS_WORKOUTsq + MALE + ELDERLY +
##     BLACK + ASIAN + AIAN + MULTIPLERACE + OTHERRACE, data = Adult.Pop3.omit,
##     atmean = FALSE)
##
## Marginal Effects:
##                           dF/dx   Std. Err.         z     P>|z|
## ALCHOL_CONSUMPTION -0.11110380  0.02166763  -5.1276 2.934e-07 ***
## SMOKER              0.68083983  0.02841004  23.9648 < 2.2e-16 ***
## log(BMI)            1.95026591  0.03991755  48.8574 < 2.2e-16 ***
## MODERATE_WORKOUT   -0.01562576  0.00451939  -3.4575 0.0005452 ***
## MODERATE_WORKOUTsq  0.00100223  0.00023166   4.3263 1.516e-05 ***
## VIGOROUS_WORKOUT   -0.13684062  0.00598733 -22.8550 < 2.2e-16 ***
## VIGOROUS_WORKOUTsq  0.00446730  0.00029682  15.0507 < 2.2e-16 ***
## MALE1              -0.32125406  0.01754782 -18.3074 < 2.2e-16 ***
## ELDERLY1            1.12703542  0.02261298  49.8402 < 2.2e-16 ***
## BLACK1             -0.05242110  0.02662493  -1.9689 0.0489677 *
## ASIAN1             -0.29906664  0.03956219  -7.5594 4.049e-14 ***
## AIAN1              -0.06884912  0.07924715  -0.8688 0.3849621
## MULTIPLERACE1       0.21243114  0.06551802   3.2423 0.0011856 **
## OTHERRACE1         -0.35495665  0.16820277  -2.1103 0.0348334 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## dF/dx is for discrete change for the following variables:
##
## [1] "ALCHOL_CONSUMPTION" "SMOKER"             "MALE1"
## [4] "ELDERLY1"           "BLACK1"             "ASIAN1"
## [7] "AIAN1"              "MULTIPLERACE1"      "OTHERRACE1"
```

Looking at figures 5 & 6 we see the marginal effects of the modeled variables with figure 5 incorporating only the chronic conditions with a 12-month time frame and figure 6 including all of the chronic conditions listed in the data appendix. The relatively small impact of MODERATE_WORKOUT is contradictory from the prior literature such as Humphreys, McLeod, and Ruseski (2014) and Sari (2014) who both found moderate physical activity to be significant in various aspects of the quality of life. Furthermore, this regressor shrinks in its effectiveness when we view it from the larger set of chronic conditions as opposed to the smaller set (Figure 6 vs Figure 5). It may have lost its significant because the inclusion of the additional variables plays no role in reducing function limitations or joint pain am individual suffers and may even contribute to these conditions due to increased chance of long-term injury. The MODERATE_WORKOUT is likely to have a smaller significance because NHIS combined moderate and light physical activity into a single question. The addition of light activity could have decreased the measured effectiveness of the impact of moderate physical activity.

The first regressor that is identified as significant is ALCHOL_CONSUMPTION which indicated that, for all else equal, a person could expect to see a decrease in the number of chronic conditions by 4.49% for the smaller set of chronic conditions and 11.11% for the larger set. While this result goes against the initial hypothesis and categorization of alcohol consumption as being a negative health-related behavior, it is not a completely unexpected. Bali and Jones (2005) found that drinking contributed to premature mortality however Phillips and Zeckhauser (1996) concluded that a moderate level of alcohol consumption can be beneficial for some heart conditions while detrimental to other health aspects. The chronic conditions chosen for this study may react more positively towards alcohol consumption and explain for the negative marginal effect. Additionally, with a magnitude of 4.49% there would be a small change in a person's health and number of from a chronic conditions a person suffers from and likely should not be a priority for a physician's initial recommendations, nor should it be recommended without further research into further impacts of alcohol consumption.

The next regressor, SMOKER, aligned with our categorization of a negative health-related behavior. The results show that if an individual smokes, and all else remaining equal, it is expected that the number of chronic conditions would increase by 20.5% and 68.1% for models #3 and #4 respectively. This is by far the largest marginal effect of any of the regressors observed. To stop smoking to increase an individual's chance at improving their quality of life via decrease the number of chronic conditions should be one of the doctor's first and main recommendations.

The final negative health-related behavior observed is an individual's BMI. The resulting marginal effect was that for all else equal, increasing the BMI of an individual by 1 percentage point would result in an increase of the number of chronic conditions by 0.92% and 1.95% for figures 5 & 6, respectively. This agrees with our expectations as well as prior literature's findings that overweight and obese individuals suffer from more chronic conditions and premature mortality. While the magnitude for the marginal effect is relatively small compared to the other regressors there is a much greater range for a person's BMI, so this means that if a

person had a BMI of ten points higher we would see an equivalent effect as smoking, on the number of chronic conditions. Additionally, Himes (2000) as well as Yang, Bishai, and Harmon (2008) observed that obese individuals suffered from lower qualities of life when compared to overweight individuals, as such it is suspected that the marginal effect of BMI may increase as BMI increase, this is an area requiring further study. A healthy BMI should be one of the doctor's first recommendation.

The only significant positive health-related behavior was an individual's frequency of vigorous activity (VIGOROUS_WORKOUT). The coefficient for this regressor is a decrease in the number of chronic conditions by 4.59% and 13.23% for figure 5 and 6. However due to prior research by Humphreys, McLeod, and Ruseski (2012) as well as Sari (2014) we know there is diminishing returns for the number of high intensity workouts which was incorporated in the variable VIGOROUS_WORKOUTsq. Even with this diminishing return it is safe to say that this health-related behavior has one of the higher returns to a person's quality of life and should be one of the physician's initial recommendations.

In addition to the regressors it should also be noted that a person's sex was statistically significant to the number of chronic conditions. If an individual was female, then they can expect to have an increase of 3.8% and 32.1% chronic conditions for figured 5 and 6. This is indicative that females are more likely to suffer from specific chronic condition or suffer greater effects from health-related behaviors which is discussed by Himes (2000).

Likewise, in agreement with Boustan and Margo (2014) we find that certain ethnic groups are affected differently than others. In the first model, black individuals are more likely to have more chronic conditions than white individuals however this does not hold true in the second model. While Asian individuals will suffer from slightly fewer chronic conditions than white individuals in the expanded model. These shifts are likely due to these ethnicities only being affected by certain chronic conditions included in one model or the other.

Himes (2000) also showed that the elderly are much more likely to suffer from functional limitations which are further exasperated by being overweight and obese. If an individual is 65 years or older we can expect to see an increase in the number of chronic conditions increase by 59.4% and 112.7% for figures 5 and 6, respectively.

4. Conclusion

In this paper we examined individual's health-related behaviors and how they affected a person's quality of life by either increasing or decreasing the number of chronic conditions they suffered from. As opposed to earlier literature this paper attempts to view the person in a wholistic manner instead of how one health-related behavior is related to one chronic condition such as physical activity and heart disease. There is clear evidence that different health-related behaviors can vary in the effectiveness of improving or decreasing the quality of life for an individual via number of chronic conditions. The amount in which the effectiveness of the health-behaviors varied was unexpected and have some clear implications for doctors and how they should prioritize their guidance and recommendations. While the frequency of vigorous workouts could have a great effect of reducing the number of chronic conditions it was smoking that had the largest marginal effect which increased the likelihood of the number of chronic

conditions. These two health-related behaviors should be the priority for a physician's guidance on lifestyle improvements, followed by the reduction of BMI while reserving their recommendations on moderate alcohol consumption and moderate physical activity without consulting additional research.

While the analysis produced results that were both statistically and practically significant there are still limitations which could be addressed in future works to provide better understandings of the relationships between health-related behaviors and chronic conditions. First is the use of cross-sectional data, by using time-series data we could better correlate the regressors and dependent variable as well as track the impact of health-related behaviors over time. Another limitation is varying timeframes that were associated with the questionnaire, causing some chronic conditions to not be included or to provide an inconsistent picture of how the regressors are affecting the number of chronic conditions.

Data Appendix

Dependent Variable(s)

New Variable Name	Original Variable	Description
NUMBER_CONDITIONS	N/A	The summation of the occurrence of the chronic conditions identified below.
HIGH_CHOLESTEROL	CHLYR	This variable represents if a person had high cholesterol within the past 12 months. One is an affirmation, otherwise it is zero
HYPERTENSION	HYPYR1	This variable represents if a person had hypertension within the past 12 months. One is an affirmation, otherwise it is zero
CON_LIVER	LIVYR	This variable represents if a person had been told they had a liver condition within the past 12 months. One is an affirmation, otherwise it is zero.
CON_KIDNEY	KIDWKYR	This variable represents if a person had been told they had weak or failing kidneys in the past 12 months. One is an affirmation, otherwise it is zero.
ULCER	ULCYR	This variable represents if a person had an ulcer in the past 12 months. One is an affirmation, otherwise it is zero.
BRONCHITIS	CBRCHYR	This variable represents if a person had been told they had chronic bronchitis in the past 12 months. One is an affirmation, otherwise it is zero.
JOINT_PAIN*	JNTSYMP	This variable represents if a person had joint pain, aching, or stiffness in the past 30 days. One is an affirmation, otherwise it is zero.
LIMITING_PAIN*	PAINLMT	This variable represents if a person had pain that limited life or work activities in the past 6 months. One is an affirmation that it occurred on at least “some days”, otherwise it is zero.
FUNC_LIMIT*	FLA1AR	This variable represents if a person had any function limitation. One is an affirmation that they had been limited in any way, otherwise it is zero.

*These chronic conditions had differing timeframes from the rest. Two analyses were conducted, one for the top six conditions and another with all of the listed chronic conditions

Independent Variables

New Variable Name	Original Variable	Description
ALCOHOL_CONSUMPTION	ALCSTAT	This variable indicates if a person identifies as a current moderate or heavy drinker. If a person falls into one of these categories the variable is given the value of one, otherwise it is a zero
SMOKER	SMKSTAT2	This variable indicates if a person currently smokes tobacco products. If they smoke some or everyday the variable is given a value of one, otherwise it is a zero
BMI	BMI	This variable the measured BMI that they have.
MODERATE_WORKOUT	MODFREQW	This variable indicates if a person participates in light or moderate physical activity, and its value is the frequency in a week.
VIGOROUS_WORKOUT	VIGFREQW	This variable indicates if a person participates in vigorous activity and its value is the frequency in a week.

Dummy Variables

New Variable Name	Original Variable	Description
MALE1	SEX	This dummy variable is used to control for the sex of an individual. A one indicates that the person is male, zero if female.
ELDERLY1	AGE_P	This dummy variable is used to control if an individual is considered elderly. A one indicates that the person is 65 years or older, zero if they are younger.
REGION.NE **REGION.MW** **REGION.SO**	REGION	These dummy variables are used to control for the geographical location of an individual.
BLACK1 **AIAN1** **ASIAN1** **MULTIPLERACE1** **OTHERRACE**	RACEPI2	These dummy variables are used to control for the race of an individual.

References

Anderson, Gerard, and Jane Horvath. 2004. "The growing Burden of Chronic Disease in America." *Public Heath Reports* 119(3): 263-70.

Balia, Silvia, and Andrew M. Jones. 2008. "Mortality, Lifestyle and Socio-economic Status." *Journal of Health Economics* 27 (1): 1-26.

Basu, Rshmita. 2015. "Lasting Impacts of Childhood Health and Socioeconomic Circumstances on Adult Health Problems: Analysis of a Longitudinal Count Regression Model" *International Journal of Applied Economics* 12(1): 80-97.

Boustan, Leah, Robert A. Margo. 2012. "RACIAL DIFFERENCE IN HEALTH IN LONG-RUN PERSPECTIVE: A BRIEF INTRODUCTION" *National Bureau of Economic Research* Working Paper 20765.

Buttorff, Christine, Teague Ruder, and Melissa Bauman. 2017. "Multiple Chronic Conditions in the United States." Santa Monica, CA: RAND Corporation. https://www.rend.org/pubs/tools.TL221.html

Centers for Disease Control and Prevention and the Merck Company Foundation. 2007. "The State of Aging and Health in America." Whitehouse Station, NJ: The Merck Company Foundation. https://www.cdc.gov/aging/pdf/saha_2007.pdf

Colman, Gregory J., and Dhaval M. Dave. 2013. "Physical Activity and Health." *National Bureau of Economic Research* Working Paper 18858.

Fan, Maoyong, and Yanhong Jin. 2014b. "Obesity and Self-Control: Food Consumption, Physical Activity, and Weight-Loss Intention." *Applied Economic Perspectives and Policy* 36 (1): 125-45.

Harris, William T., and Lydia Harris. 1996. "The Decision to Quit Smoking: Theory and Evidence." Journal of Socio-Economics 25 (5): 601–18. doi:http://www.sciencedirect.com.leo.lib.unomaha.edu/science/journal/10535357.

Heiss, Florian. "Using R for Introductory Econometrics" Dusseldorf, Germany: 2016.

Himes, Christine L. 2000. "Obesity, Disease, and Functional Limitation in Later Life" *Demography* 37: 73-82.

Humphreys. Brad R., Logan McLeod, and Jane E. Ruseski. 2014. "Physical Activity and Health Outcomes: Evidence from Canada" *Health Economics* 23: 33-54.

National Center for Health Statistics. National Health Interview Survey, 2017. Public-use data file and documentation. https://www.cdc.gov/nchs/nhis/data-questionnaires-documentation.htm. 2018.

Phillips, Carl V., and Richard Zeckhauser. 1996. "Communicating the Health Effectis of Consumer Products: The Case of Moderate Alcohol Consumption and Coronary Heart Disease." *Managerial and Decision Economics* 17: 459-70.

Richards, Timothy J. 2007. "Obesity and Nutrient Consumption: A Rational Addition?" *Contemporary Economics Policy* 25 (3): 309-24.

Sari, Nazmi. 2014. "Sports, Exercise, and Length of Stay in Hospitals: Is There a Differential Effect for the Chronically Ill People" *Contemporary Economics Policy* 32 (2): 247-60.

U.S. Department of Health and Human Services. 2014. "The Health Consequences of Smoking – 50 Years of Progress" Rockville, MD: *Office of the Surgeon General*.

Westman J, K. Wahlbeck, T.M. Laursen, M. Gissler, M. Nordentoft, J. Hallgren, M. Arffman, U. Osby. 2015. "Mortality and life expectancy of people with alcohol use disorder in Denmark, Finland and Sweden" Acta Psychiatrica Scandinavica 131: 297-306.

Wooldridge, Jeffrey M. "Introductory Econometrics: A Modern Approach, Seventh Edition" Boston, MA: Cengage, 2020.

Yang, Zhou, David Bishai, and Jeffery Harman. 2008. "Convergrence of body mass with aging: The longitudinal interrelationship of health, weight, and survival." *Economics and Human Biology* 6: 469-481.