DebTrevedi Analysis

Introduction In this report the demand for medical care data which was studied by Deb and Trevedi in 1997 was analyzed. The original dataset consists of 19 variables and 4406 observations made on people older than 66 and covered by a public insurance program Medicare. The aim of this study was to study how access to private health insurance (privins) and medicare program (medicaid) impacts the number of physician office visits (ofp).

The inbuilt R dataset, DebTrevedi, was first read (db) and the variables, “ofnp”, “opp”, “opnp”, “emer” and “hosp” were removed as they were not used in our analysis (newdata). The data was then modeled using a negative binomial regression with ofp as the response variable and the rest of the variables as the explanatory variables (glm1). Backward elimination was then performed to arrive at a statistically significant model (step.lm). A separate negative binomial was then fitted with all possible interaction terms (glm.f). Significant terms from both models were combined and further backward and chi-square elimination was performed to obtain our final model with explanatory and interaction variables (final.lm). Lastly, models were compared and our final analysis was performed on the least complex model.

Results As we can see from Figure 1, the distribution of physician visits (ofp) is highly skewed to the right (figure 1). Our model shows that for those who that had access to private insurance and medicaid, the rate of ofp was higher by a factor of 1.72 and 1.22 (exp_estimates) when compared to those without private insurance and medicaid respectively .For those who were employed and had access to medicaid the rate of ofp was 2.76 (med_employed) higher than employed individuals without medicaid. Also, as the number of chronic conditions increased by 1 the rate of physician visits increased by a factor of 1.33, unless privately insured, in which case the rate of ofp increases by 1.25 (priv_numchron).

Compared to those who had average health, poor health individuals showed a increase rate of ofp by a factor 2.17, whereas, those in excellent health had a decreased rate of ofp by a factor of 0.57. Similarly males as well as married individuals had a decrease rate of ofp by a factor of 0.77 and 0.88 when compared to females and unmarried people respectively. However, males who were married were found to have an increase incidence of ofp by 18% (male_married) compared to married females. An increase in age and number of years of education (school) by 1 year was also found to increase ofp by 2.5% and 4.1% respectively.

Lastly, those with disabilities (adldiff) had the greatest impact on visits as the incidence of ofp in this group was 11.88 times greater compared to those without disabilities. Also, within this group an increase in age by 1, decreased the rate of ofp by 22% (adl_age).

Discussion Our analysis shows that access to both private insurance and medicaid significantly increases the rate of physician visits (ofp) by 72% and 22% respectively ($p<0.05$). Furthermore, for those who are employed and have access to medicaid, this rate (ofp) increases nearly 3 times (med_employed). Although our results were found to be statistically significant, our model also shows outliers, a curve pattern in the residuals (Figure 2) as well as large standardized residuals which could make our model and predictions unreliable. An attempt to transform the model using square root transformation on the response variable (ofp) was not successful as it resulted in a large degree of standard error (final.sqrt). Similarly, adding a quadratic or cubic term to ofp reduced the curvature in the residuals but also increased the standardized residuals and the AIC significantly (final.cubic). Results were therefore analyzed using the un-transformed model (final.lm), however an appropriate transformation and removal of outliers is necessary for more reliable analysis.

Apendix

Importing data.

```
library(MixAll)
```

```
## Loading required package: rtkore

## Loading required package: Rcpp

##
## Attaching package: 'rtkore'

## The following object is masked from 'package:Rcpp':
##
##     LdFlags
```

```
library(rtkore)

data(DebTrivedi)

db <- DebTrivedi

summary(db)
```

```
##       ofp              ofnp               opp                opnp
##  Min.   : 0.000   Min.   :  0.000   Min.   :  0.0000   Min.   :  0.0000
##  1st Qu.: 1.000   1st Qu.:  0.000   1st Qu.:  0.0000   1st Qu.:  0.0000
##  Median : 4.000   Median :  0.000   Median :  0.0000   Median :  0.0000
##  Mean   : 5.774   Mean   :  1.618   Mean   :  0.7508   Mean   :  0.5361
##  3rd Qu.: 8.000   3rd Qu.:  1.000   3rd Qu.:  0.0000   3rd Qu.:  0.0000
##  Max.   :89.000   Max.   :104.000   Max.   :141.0000   Max.   :155.0000
##       emer              hosp              health       numchron     adldiff
##  Min.   : 0.0000   Min.   :0.000   poor     : 554   Min.   :0.000   no :3507
##  1st Qu.: 0.0000   1st Qu.:0.000   average  :3509   1st Qu.:1.000   yes: 899
##  Median : 0.0000   Median :0.000   excellent: 343   Median :1.000
##  Mean   : 0.2635   Mean   :0.296                    Mean   :1.542
##  3rd Qu.: 0.0000   3rd Qu.:0.000                    3rd Qu.:2.000
##  Max.   :12.0000   Max.   :8.000                    Max.   :8.000
##      region          age         black        gender     married
##  midwest:1157   Min.   : 6.600   no :3890   female:2628   no :2000
##  noreast: 837   1st Qu.: 6.900   yes: 516   male  :1778   yes:2406
##  other  :1614   Median : 7.300
##  west   : 798   Mean   : 7.402
##                 3rd Qu.: 7.800
##                 Max.   :10.900
##      school          faminc        employed   privins    medicaid
##  Min.   : 0.00   Min.   :-1.0125   no :3951   no : 985   no :4004
##  1st Qu.: 8.00   1st Qu.: 0.9122   yes: 455   yes:3421   yes: 402
##  Median :11.00   Median : 1.6982
```

```
##  Mean   :10.29   Mean   : 2.5271
##  3rd Qu.:12.00   3rd Qu.: 3.1728
##  Max.   :18.00   Max.   :54.8351
```

```
str(db)
```

```
## 'data.frame':    4406 obs. of  19 variables:
##  $ ofp     : int  5 1 13 16 3 17 9 3 1 0 ...
##  $ ofnp    : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ opp     : int  0 2 0 5 0 0 0 0 0 0 ...
##  $ opnp    : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ emer    : int  0 2 3 1 0 0 0 0 0 0 ...
##  $ hosp    : int  1 0 3 1 0 0 0 0 0 0 ...
##  $ health  : Factor w/ 3 levels "poor","average",..: 2 2 1 1 2 1 2 2 2 2 ...
##   ..- attr(*, "contrasts")= num [1:3, 1:2] 1 0 0 0 0 1
##   .. ..- attr(*, "dimnames")=List of 2
##   .. .. ..$ : chr [1:3] "poor" "average" "excellent"
##   .. .. ..$ : chr [1:2] "poor" "excellent"
##  $ numchron: int  2 2 4 2 2 5 0 0 0 0 ...
##  $ adldiff : Factor w/ 2 levels "no","yes": 1 1 2 2 2 2 1 1 1 1 ...
##  $ region  : Factor w/ 4 levels "midwest","noreast",..: 3 3 3 3 3 3 1 1 1 1 ...
##  $ age     : num  6.9 7.4 6.6 7.6 7.9 6.6 7.5 8.7 7.3 7.8 ...
##  $ black   : Factor w/ 2 levels "no","yes": 2 1 2 1 1 1 1 1 1 1 ...
##  $ gender  : Factor w/ 2 levels "female","male": 2 1 1 2 1 1 1 1 1 1 ...
##  $ married : Factor w/ 2 levels "no","yes": 2 2 1 2 2 1 1 1 1 1 ...
##  $ school  : int  6 10 10 3 6 7 8 8 8 8 ...
##  $ faminc  : num  2.881 2.748 0.653 0.659 0.659 ...
##  $ employed: Factor w/ 2 levels "no","yes": 2 1 1 1 1 1 1 1 1 1 ...
##  $ privins : Factor w/ 2 levels "no","yes": 2 2 1 2 2 1 2 2 2 2 ...
##  $ medicaid: Factor w/ 2 levels "no","yes": 1 1 2 1 1 2 1 1 1 1 ...
```

Removing variables.

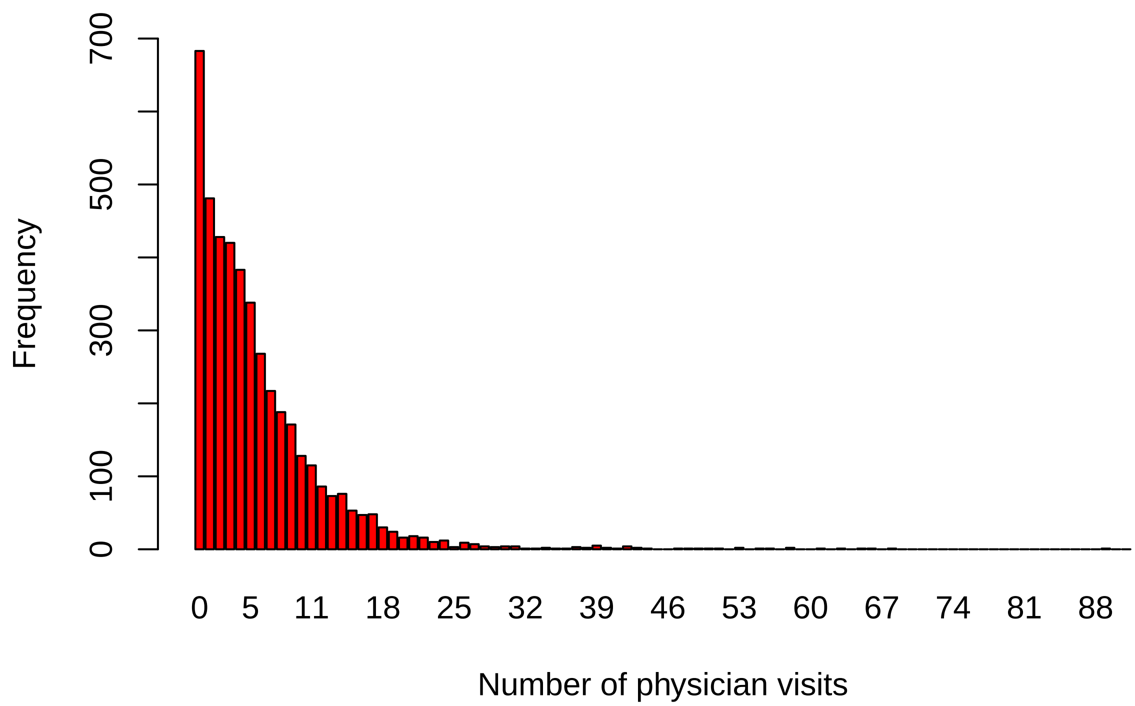
```
myvars <- names(db) %in% c("ofnp", "opp", "opnp", "emer", "hosp")

newdata <- db[!myvars]

attach(newdata)

str(newdata)
```

```
## 'data.frame':    4406 obs. of  14 variables:
##  $ ofp     : int  5 1 13 16 3 17 9 3 1 0 ...
##  $ health  : Factor w/ 3 levels "poor","average",..: 2 2 1 1 2 1 2 2 2 2 ...
##   ..- attr(*, "contrasts")= num [1:3, 1:2] 1 0 0 0 0 1
##   .. ..- attr(*, "dimnames")=List of 2
##   .. .. ..$ : chr [1:3] "poor" "average" "excellent"
##   .. .. ..$ : chr [1:2] "poor" "excellent"
##  $ numchron: int  2 2 4 2 2 5 0 0 0 0 ...
##  $ adldiff : Factor w/ 2 levels "no","yes": 1 1 2 2 2 2 1 1 1 1 ...
##  $ region  : Factor w/ 4 levels "midwest","noreast",..: 3 3 3 3 3 3 1 1 1 1 ...
```

```
## $ age     : num  6.9 7.4 6.6 7.6 7.9 6.6 7.5 8.7 7.3 7.8 ...
## $ black   : Factor w/ 2 levels "no","yes": 2 1 2 1 1 1 1 1 1 1 ...
## $ gender  : Factor w/ 2 levels "female","male": 2 1 1 2 1 1 1 1 1 1 ...
## $ married : Factor w/ 2 levels "no","yes": 2 2 1 2 2 1 1 1 1 1 ...
## $ school  : int  6 10 10 3 6 7 8 8 8 8 ...
## $ faminc  : num  2.881 2.748 0.653 0.659 0.659 ...
## $ employed: Factor w/ 2 levels "no","yes": 2 1 1 1 1 1 1 1 1 1 ...
## $ privins : Factor w/ 2 levels "no","yes": 2 2 1 2 2 1 2 2 2 2 ...
## $ medicaid: Factor w/ 2 levels "no","yes": 1 1 2 1 1 2 1 1 1 1 ...
```

Figure 1. Ofp barplot.

```
with(newdata, barplot(table(factor(ofp, levels = 0:91)),ylim = c(0, 700),col = 'red', xlab =
                           "Number of physician visits",ylab ="Frequency"))
```



Fitting first model with all terms.

```
library(MASS)

glm1 <- glm.nb(ofp ~ ., data = newdata)

summary(glm1)
```

```
##
## Call:
## glm.nb(formula = ofp ~ ., data = newdata, init.theta = 1.182249667,
##     link = log)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.6457  -0.9975  -0.3054   0.3081   5.3536
##
## Coefficients:
##                  Estimate Std. Error z value Pr(>|z|)
## (Intercept)      1.058308   0.213962   4.946 7.57e-07 ***
## healthpoor       0.322698   0.050295   6.416 1.40e-10 ***
## healthexcellent -0.384556   0.061806  -6.222 4.91e-10 ***
## numchron         0.189739   0.012174  15.586  < 2e-16 ***
## adldiffyes       0.096768   0.043183   2.241  0.02504 *
## regionnoreast    0.112397   0.046390   2.423  0.01540 *
## regionother      0.001141   0.040276   0.028  0.97739
## regionwest       0.137962   0.047548   2.902  0.00371 **
## age             -0.035642   0.026601  -1.340  0.18028
## blackyes        -0.066178   0.052755  -1.254  0.20968
## gendermale      -0.075075   0.034577  -2.171  0.02991 *
## marriedyes      -0.035570   0.036199  -0.983  0.32579
## school           0.027393   0.004695   5.835 5.38e-09 ***
## faminc          -0.001669   0.005638  -0.296  0.76720
## employedyes      0.025567   0.053042   0.482  0.62979
## privinsyes       0.345881   0.045353   7.626 2.41e-14 ***
## medicaidyes      0.276162   0.063160   4.372 1.23e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(1.1822) family taken to be 1)
##
##     Null deviance: 5665.6  on 4405  degrees of freedom
## Residual deviance: 5040.6  on 4389  degrees of freedom
## AIC: 24440
##
## Number of Fisher Scoring iterations: 1
##
##
##               Theta:  1.1822
##           Std. Err.:  0.0326
##
##  2 x log-likelihood:  -24404.3350
```

Backward step elmination of glm1.

```
step.lm<-stepAIC(glm1)
summary(step.lm)
```

Second glm with full interaction terms.

```
glm.f<-glm.nb(ofp ~  (medicaid+privins+employed+health+numchron+ black+adldiff+gender+age+married
                     +school+faminc+region)^2 , data = newdata  )

summary(glm.f)
```

Combined model of significant terms from glm1 and glm.f

```
mod.glm <- glm.nb(ofp ~  (medicaid+privins+numchron+region+health+gender+school+adldiff)^2 +medicaid*emp
                    employed*region+adldiff*age+gender*married, data = newdata  )

summary(mod.glm)
```

Backward step elmination of mod.glm.

```
step.modglm<-stepAIC(mod.glm)
summary(step.modglm)
```

dropterm sequence chisquare emlination of step.modglm

```
dropterm(step.modglm, test = "Chisq")

# drop medicaid:privins

summary(step.lm2 <- update(step.modglm, . ~ . - medicaid:privins))

dropterm(step.lm2, test = "Chisq")

# drop privins:health
summary(step.lm3 <- update(step.lm2, . ~ . - privins:health))


dropterm(step.lm3, test = "Chisq")

# drop medicaid:health

summary(step.lm4 <- update(step.lm3, . ~ . - medicaid:health))


dropterm(step.lm4, test = "Chisq")

# drop health:gender
summary(step.lm5 <- update(step.lm4, . ~ . - health:gender))

dropterm(step.lm5, test = "Chisq")
```

```r
summary(step.lm5)
```

Final model from step.lm5

```r
final.lm<- glm.nb(formula = ofp ~ medicaid + privins + numchron + region +
                    health + gender + school + adldiff + employed + age + married +
                    privins:numchron + privins:region + numchron:health + numchron:adldiff +
                    region:health + region:school + health:school + medicaid:employed +
                    health:employed + region:employed + adldiff:age + gender:married,
                  data = newdata )

summary(final.lm)
```

```
##
## Call:
## glm.nb(formula = ofp ~ medicaid + privins + numchron + region +
##     health + gender + school + adldiff + employed + age + married +
##     privins:numchron + privins:region + numchron:health + numchron:adldiff +
##     region:health + region:school + health:school + medicaid:employed +
##     health:employed + region:employed + adldiff:age + gender:married,
##     data = newdata, init.theta = 1.238004748, link = log)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.5380  -0.9942  -0.2932   0.3014   5.4726
##
## Coefficients:
##                                Estimate Std. Error z value Pr(>|z|)
## (Intercept)                    0.333109   0.272903   1.221 0.222233
## medicaidyes                    0.202779   0.063714   3.183 0.001459 **
## privinsyes                     0.540769   0.098743   5.477 4.34e-08 ***
## numchron                       0.283979   0.026008  10.919  < 2e-16 ***
## regionnoreast                  0.194126   0.173366   1.120 0.262821
## regionother                   -0.020327   0.140516  -0.145 0.884980
## regionwest                     0.211760   0.173470   1.221 0.222191
## healthpoor                     0.774269   0.173401   4.465 8.00e-06 ***
## healthexcellent               -0.553968   0.208778  -2.653 0.007969 **
## gendermale                    -0.261636   0.058763  -4.452 8.49e-06 ***
## school                         0.025290   0.009968   2.537 0.011174 *
## adldiffyes                     2.474735   0.423907   5.838 5.29e-09 ***
## employedyes                   -0.125300   0.105490  -1.188 0.234916
## age                            0.040169   0.031948   1.257 0.208639
## marriedyes                    -0.123187   0.042403  -2.905 0.003671 **
## privinsyes:numchron           -0.058837   0.026158  -2.249 0.024494 *
## privinsyes:regionnoreast       0.091090   0.128185   0.711 0.477324
## privinsyes:regionother        -0.175478   0.109112  -1.608 0.107783
## privinsyes:regionwest         -0.204271   0.128160  -1.594 0.110963
## numchron:healthpoor           -0.143627   0.031519  -4.557 5.19e-06 ***
## numchron:healthexcellent       0.097688   0.061432   1.590 0.111796
## numchron:adldiffyes           -0.087482   0.026939  -3.247 0.001165 **
## regionnoreast:healthpoor       0.381789   0.149587   2.552 0.010702 *
```

```
## regionother:healthpoor         0.176915   0.120664   1.466 0.142599
## regionwest:healthpoor         -0.101344   0.150742  -0.672 0.501393
## regionnoreast:healthexcellent  0.063610   0.185246   0.343 0.731313
## regionother:healthexcellent   -0.269609   0.164055  -1.643 0.100298
## regionwest:healthexcellent    -0.319998   0.169703  -1.886 0.059345 .
## regionnoreast:school          -0.021890   0.014663  -1.493 0.135484
## regionother:school             0.014545   0.011982   1.214 0.224774
## regionwest:school              0.013064   0.014024   0.932 0.351584
## healthpoor:school             -0.031551   0.012243  -2.577 0.009964 **
## healthexcellent:school         0.022537   0.016020   1.407 0.159478
## medicaidyes:employedyes        0.812072   0.370937   2.189 0.028579 *
## healthpoor:employedyes         0.866871   0.224777   3.857 0.000115 ***
## healthexcellent:employedyes    0.108788   0.153748   0.708 0.479212
## regionnoreast:employedyes      0.365795   0.158383   2.310 0.020913 *
## regionother:employedyes        0.027236   0.132790   0.205 0.837486
## regionwest:employedyes        -0.089882   0.153102  -0.587 0.557155
## adldiffyes:age                -0.288929   0.054466  -5.305 1.13e-07 ***
## gendermale:marriedyes          0.286417   0.072126   3.971 7.16e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(1.238) family taken to be 1)
##
##     Null deviance: 5843.2  on 4405  degrees of freedom
## Residual deviance: 5044.0  on 4365  degrees of freedom
## AIC: 24341
##
## Number of Fisher Scoring iterations: 1
##
##
##               Theta:  1.2380
##           Std. Err.:  0.0347
##
##  2 x log-likelihood:  -24256.5230
```

```
dropterm(final.lm, test = "Chisq")
```

```
## Single term deletions
##
## Model:
## ofp ~ medicaid + privins + numchron + region + health + gender +
##     school + adldiff + employed + age + married + privins:numchron +
##     privins:region + numchron:health + numchron:adldiff + region:health +
##     region:school + health:school + medicaid:employed + health:employed +
##     region:employed + adldiff:age + gender:married
##                     Df   AIC     LRT   Pr(Chi)
## <none>                 24339
## privins:numchron    1 24341  4.8681 0.0273573 *
## privins:region      3 24341  8.2942 0.0403069 *
## numchron:health     2 24357 22.2418 1.480e-05 ***
## numchron:adldiff    1 24347 10.0811 0.0014980 **
## region:health       6 24344 17.2586 0.0083785 **
## region:school       3 24341  8.9087 0.0305291 *
## health:school       2 24344  9.6699 0.0079469 **
```

```
## medicaid:employed  1 24342  5.8286 0.0157677 *
## health:employed    2 24352 17.3568 0.0001702 ***
## region:employed    3 24341  8.6879 0.0337409 *
## adldiff:age        1 24364 27.0078 2.026e-07 ***
## gender:married     1 24352 15.4201 8.607e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

exp_estimates

```
exp(exp_estimates <- cbind(Estimate = coef(final.lm), confint(final.lm)))
```

```
##                                  Estimate     2.5 %     97.5 %
## (Intercept)                     1.3952997 0.8145951  2.3875124
## medicaidyes                     1.2248022 1.0789190  1.3918687
## privinsyes                      1.7173267 1.4109835  2.0855728
## numchron                        1.3284050 1.2618892  1.3992152
## regionnoreast                   1.2142492 0.8677801  1.6996030
## regionother                     0.9798782 0.7476257  1.2819268
## regionwest                      1.2358506 0.8824187  1.7324499
## healthpoor                      2.1690060 1.5522603  3.0444963
## healthexcellent                 0.5746649 0.3776531  0.8762579
## gendermale                      0.7697910 0.6862456  0.8647204
## school                          1.0256127 1.0057985  1.0457544
## adldiffyes                     11.8785570 5.1114198 27.6564704
## employedyes                     0.8822326 0.7176969  1.0904190
## age                             1.0409868 0.9766550  1.1099016
## marriedyes                      0.8840986 0.8137359  0.9607660
## privinsyes:numchron             0.9428605 0.8944667  0.9934566
## privinsyes:regionnoreast        1.0953675 0.8513778  1.4095299
## privinsyes:regionother          0.8390559 0.6772670  1.0414395
## privinsyes:regionwest           0.8152412 0.6333275  1.0494156
## numchron:healthpoor             0.8662105 0.8127961  0.9234518
## numchron:healthexcellent        1.1026186 0.9670205  1.2624499
## numchron:adldiffyes             0.9162352 0.8681960  0.9670078
## regionnoreast:healthpoor        1.4649030 1.0907953  1.9712764
## regionother:healthpoor          1.1935293 0.9402974  1.5094505
## regionwest:healthpoor           0.9036222 0.6720213  1.2173553
## regionnoreast:healthexcellent   1.0656764 0.7405585  1.5396016
## regionother:healthexcellent     0.7636778 0.5531720  1.0525931
## regionwest:healthexcellent      0.7261505 0.5194764  1.0144465
## regionnoreast:school            0.9783482 0.9507752  1.0067161
## regionother:school              1.0146517 0.9913041  1.0386082
## regionwest:school               1.0131494 0.9856613  1.0413704
## healthpoor:school               0.9689416 0.9464000  0.9919571
## healthexcellent:school          1.0227934 0.9901313  1.0566528
## medicaidyes:employedyes         2.2525708 1.1531735  4.9909898
## healthpoor:employedyes          2.3794539 1.5550829  3.7921897
## healthexcellent:employedyes     1.1149254 0.8207146  1.5214157
## regionnoreast:employedyes       1.4416600 1.0558724  1.9736392
## regionother:employedyes         1.0276108 0.7901176  1.3344398
## regionwest:employedyes          0.9140393 0.6765372  1.2370582
```

```
## adldiffyes:age                 0.7490654 0.6718333  0.8351292
## gendermale:marriedyes          1.3316474 1.1550290  1.5337382
```

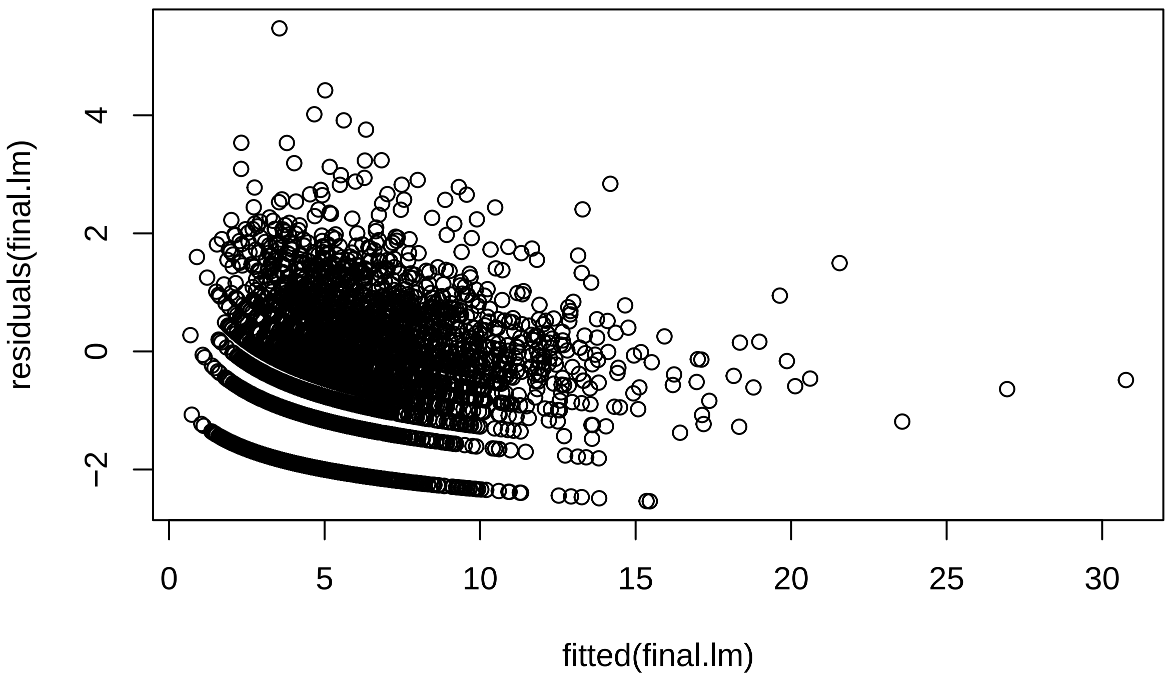
Analysis of interaction coeffecients

```
# The effect of chronic conditions on privately insured people.
priv_numchron<-exp(0.283979-0.058837)

#access to medicaid and employed

med_employed<-exp(0.812072+0.202779)

# male and married
male_married<-exp(0.286417-0.123187)


#disables adldiff and age
adl_age<-exp(0.040169-0.288929)

cbind(priv_numchron,med_employed,male_married,adl_age)
```

```
##      priv_numchron med_employed male_married   adl_age
## [1,]      1.252501     2.758952     1.177307 0.7797671
```

Figure 2 - Residuals vs fitted values for final.lm

```
plot(fitted(final.lm), residuals (final.lm))
```



Sqrt Transformed- Final sqrt

```
final.sqrt<- glm.nb(formula = ofp**0.5 ~ medicaid + privins + numchron + region +
                    health + gender + school + adldiff + employed + age + married +
                    privins:numchron + privins:region + numchron:health + numchron:adldiff +
                    region:health + region:school + health:school + medicaid:employed +
                    health:employed + region:employed + adldiff:age + gender:married,
                  data = newdata )

summary(final.sqrt)
```

```
##
## Call:
## glm.nb(formula = ofp^0.5 ~ medicaid + privins + numchron + region +
##     health + gender + school + adldiff + employed + age + married +
##     privins:numchron + privins:region + numchron:health + numchron:adldiff +
##     region:health + region:school + health:school + medicaid:employed +
##     health:employed + region:employed + adldiff:age + gender:married,
##     data = newdata, init.theta = 39628.4603, link = log)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.5990  -0.5934  -0.0166   0.4759   4.0007
```

```
##
## Coefficients:
##                                Estimate Std. Error z value Pr(>|z|)
## (Intercept)                   -0.365748   0.196414  -1.862 0.062585 .
## medicaidyes                    0.157966   0.044486   3.551 0.000384 ***
## privinsyes                     0.401732   0.075889   5.294 1.20e-07 ***
## numchron                       0.185773   0.017877  10.392  < 2e-16 ***
## regionnoreast                  0.140959   0.127929   1.102 0.270526
## regionother                    0.072161   0.104637   0.690 0.490428
## regionwest                     0.208437   0.126711   1.645 0.099973 .
## healthpoor                     0.464877   0.116127   4.003 6.25e-05 ***
## healthexcellent               -0.424916   0.164951  -2.576 0.009994 **
## gendermale                    -0.213977   0.043812  -4.884 1.04e-06 ***
## school                         0.021007   0.007168   2.931 0.003381 **
## adldiffyes                     1.342357   0.291506   4.605 4.13e-06 ***
## employedyes                   -0.078481   0.076337  -1.028 0.303907
## age                            0.038921   0.022613   1.721 0.085217 .
## marriedyes                    -0.049568   0.029513  -1.680 0.093048 .
## privinsyes:numchron           -0.045666   0.017871  -2.555 0.010608 *
## privinsyes:regionnoreast       0.049337   0.096277   0.512 0.608339
## privinsyes:regionother        -0.129624   0.081749  -1.586 0.112822
## privinsyes:regionwest         -0.110255   0.093919  -1.174 0.240421
## numchron:healthpoor           -0.085797   0.020245  -4.238 2.26e-05 ***
## numchron:healthexcellent       0.050149   0.043687   1.148 0.251004
## numchron:adldiffyes           -0.052308   0.017789  -2.941 0.003277 **
## regionnoreast:healthpoor       0.182719   0.097775   1.869 0.061655 .
## regionother:healthpoor         0.130383   0.080459   1.620 0.105127
## regionwest:healthpoor         -0.025188   0.099140  -0.254 0.799449
## regionnoreast:healthexcellent -0.086445   0.143008  -0.604 0.545526
## regionother:healthexcellent   -0.164916   0.125302  -1.316 0.188125
## regionwest:healthexcellent    -0.219560   0.129174  -1.700 0.089182 .
## regionnoreast:school          -0.015138   0.010409  -1.454 0.145827
## regionother:school             0.001048   0.008611   0.122 0.903097
## regionwest:school             -0.002205   0.009938  -0.222 0.824450
## healthpoor:school             -0.020994   0.008112  -2.588 0.009655 **
## healthexcellent:school         0.023601   0.012580   1.876 0.060646 .
## medicaidyes:employedyes        0.494619   0.216532   2.284 0.022355 *
## healthpoor:employedyes         0.411022   0.134752   3.050 0.002287 **
## healthexcellent:employedyes    0.123749   0.114272   1.083 0.278838
## regionnoreast:employedyes      0.154940   0.113141   1.369 0.170858
## regionother:employedyes        0.006788   0.096112   0.071 0.943692
## regionwest:employedyes        -0.034518   0.108636  -0.318 0.750679
## adldiffyes:age                -0.157714   0.037612  -4.193 2.75e-05 ***
## gendermale:marriedyes          0.200571   0.052723   3.804 0.000142 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(39628.46) family taken to be 1)
##
##     Null deviance: 4677.2  on 4405  degrees of freedom
## Residual deviance: 4030.5  on 4365  degrees of freedom
## AIC: 14140
##
## Number of Fisher Scoring iterations: 1
```

```
##
##
##               Theta:  39628
##           Std. Err.:  85249
## Warning while fitting theta: iteration limit reached
##
##  2 x log-likelihood:  -14055.95
```

Cube Transformed- final.cube

```
final.cube<- glm.nb(formula = ofp**3 ~ medicaid + privins + numchron + region +
                    health + gender + school + adldiff + employed + age + married +
                    privins:numchron + privins:region + numchron:health + numchron:adldiff +
                    region:health + region:school + health:school + medicaid:employed +
                    health:employed + region:employed + adldiff:age + gender:married,
                  data = newdata )

summary(final.cube)
```

```
##
## Call:
## glm.nb(formula = ofp^3 ~ medicaid + privins + numchron + region +
##     health + gender + school + adldiff + employed + age + married +
##     privins:numchron + privins:region + numchron:health + numchron:adldiff +
##     region:health + region:school + health:school + medicaid:employed +
##     health:employed + region:employed + adldiff:age + gender:married,
##     data = newdata, init.theta = 0.185757626, link = log)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.2205  -1.3494  -0.8543  -0.3760   8.7820
##
## Coefficients:
##                                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)                      4.64929    0.62215   7.473 7.84e-14 ***
## medicaidyes                      0.05729    0.14715   0.389 0.697024
## privinsyes                       1.20656    0.21721   5.555 2.78e-08 ***
## numchron                         0.60137    0.06004  10.016  < 2e-16 ***
## regionnoreast                    0.49793    0.39186   1.271 0.203840
## regionother                     -0.36649    0.31415  -1.167 0.243379
## regionwest                      -0.10186    0.39305  -0.259 0.795521
## healthpoor                       1.40520    0.41156   3.414 0.000639 ***
## healthexcellent                 -1.13395    0.44559  -2.545 0.010933 *
## gendermale                      -0.58115    0.13285  -4.374 1.22e-05 ***
## school                           0.05023    0.02273   2.210 0.027107 *
## adldiffyes                       8.09906    0.98941   8.186 2.71e-16 ***
## employedyes                     -0.67650    0.24022  -2.816 0.004860 **
## age                              0.03598    0.07348   0.490 0.624323
## marriedyes                      -0.64991    0.09827  -6.613 3.76e-11 ***
## privinsyes:numchron             -0.08856    0.06061  -1.461 0.143988
## privinsyes:regionnoreast        -0.09156    0.28698  -0.319 0.749705
## privinsyes:regionother          -0.80883    0.24413  -3.313 0.000923 ***
```

```
## privinsyes:regionwest         -0.40916    0.29026  -1.410 0.158641
## numchron:healthpoor           -0.30469    0.07556  -4.032 5.52e-05 ***
## numchron:healthexcellent       0.21994    0.13947   1.577 0.114813
## numchron:adldiffyes           -0.24125    0.06365  -3.790 0.000151 ***
## regionnoreast:healthpoor       1.33262    0.35890   3.713 0.000205 ***
## regionother:healthpoor         0.39905    0.28711   1.390 0.164565
## regionwest:healthpoor         -0.21229    0.35984  -0.590 0.555229
## regionnoreast:healthexcellent  3.12291    0.41119   7.595 3.08e-14 ***
## regionother:healthexcellent   -0.77580    0.35536  -2.183 0.029027 *
## regionwest:healthexcellent    -0.57814    0.37107  -1.558 0.119229
## regionnoreast:school          -0.04864    0.03375  -1.441 0.149459
## regionother:school             0.08689    0.02730   3.183 0.001457 **
## regionwest:school              0.07364    0.03219   2.287 0.022169 *
## healthpoor:school             -0.05838    0.02915  -2.003 0.045163 *
## healthexcellent:school        -0.01555    0.03411  -0.456 0.648465
## medicaidyes:employedyes        1.87061    0.90244   2.073 0.038187 *
## healthpoor:employedyes         3.62951    0.55363   6.556 5.53e-11 ***
## healthexcellent:employedyes   -0.09783    0.33880  -0.289 0.772760
## regionnoreast:employedyes      2.33139    0.36630   6.365 1.96e-10 ***
## regionother:employedyes        0.65935    0.30188   2.184 0.028950 *
## regionwest:employedyes         0.32159    0.35106   0.916 0.359641
## adldiffyes:age                -0.94701    0.12668  -7.476 7.67e-14 ***
## gendermale:marriedyes          1.04610    0.16423   6.370 1.89e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(0.1858) family taken to be 1)
##
##     Null deviance: 6719.0  on 4405  degrees of freedom
## Residual deviance: 5713.4  on 4365  degrees of freedom
## AIC: 55985
##
## Number of Fisher Scoring iterations: 1
##
##
##               Theta:  0.18576
##           Std. Err.:  0.00337
##
##  2 x log-likelihood:  -55901.38600
```