

Section 2: Fitting PQI Models

PQIs of Interest

For these models, we will be looking at modeling the relationship between various PQIs at the HRR (hospital-referral region) level and Medicaid expansion status. For model selection, we started with the base model and added in potential covariates one by one from our list of potential covariates that did not have significant missingness in order to improve the model fit. We using likelihood ratio tests at each step to determine if adding the covariate was an appropriate addition to the model (checking the reduced model against the full model).

Exploratory analysis on covariates

First we quantify the missingness in the outcomes:

```
tbl <- hrr %>%
  select( `pq105_copd_or_asthma_in_older_adults_admission_rate_(age_40-64)`, `pq105_copd_or_asthma_in_o
  summarize_all(.funs = function(x) return(paste0(round(sum(is.na(x))/length(x), 3)*100, "%"))) %>% t()

names(tbl) <- c("% of Missing values")
kableExtra::kable(tbl)
```

	% of Missing values
pqi05_copd_or_asthma_in_older_adults_admission_rate_(age_40-64)	0.8%
pqi05_copd_or_asthma_in_older_adults_admission_rate_(age_65-74)	0.6%
pqi05_copd_or_asthma_in_older_adults_admission_rate_(age_75+)	0.6%
pqi07_hypertension_admission_rate_(age_<_65)	30.8%
pqi07_hypertension_admission_rate_(age_65-74)	20%
pqi07_hypertension_admission_rate_(age_75+)	5.6%
pqi08_chf_admission_rate_(age_65-74)	0.6%
pqi08_chf_admission_rate_(age_75+)	0.6%
pqi08_chf_admission_rate_(age_<_65)	1.4%

We have a very small percentage of missing data in our target covariates and 2 of the 3 PQIs. We believe PQI05 and PQI08 will be reasonable metrics to evaluate outpatient quality and it's relation to Medicaid expansion.

Check distribution of outcomes

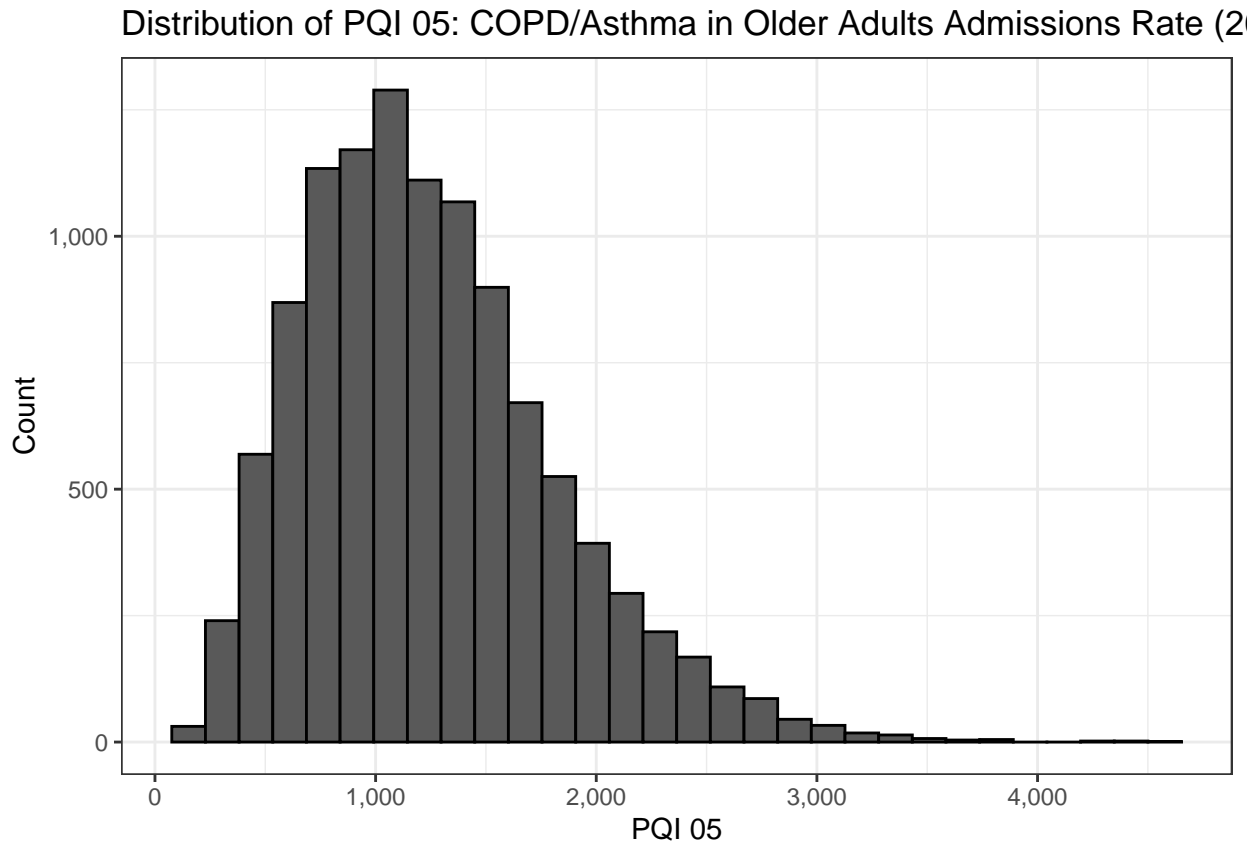
```
hrr %>%
  filter(!is.na(expansion_status)) %>%
  select(year, `pq105_copd_or_asthma_in_older_adults_admission_rate_(age_40-64)`, `pq105_copd_or_asthma
  gather(age, pqi05, -year) %>%
  mutate(
    age = str_sub(age, -10)
  ) %>% ggplot(aes( x = pqi05)) + geom_histogram(color = "black") +
  scale_x_continuous(labels = scales::comma) +
  scale_y_continuous(labels = scales::comma) +
  xlab("PQI 05") +
  ylab("Count") +
```

```
ggtitle("Distribution of PQI 05: COPD/Asthma in Older Adults Admissions Rate (2007-2018)") +
theme_bw()
```

PQI05

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
## Warning: Removed 4 rows containing non-finite values (stat_bin).
```



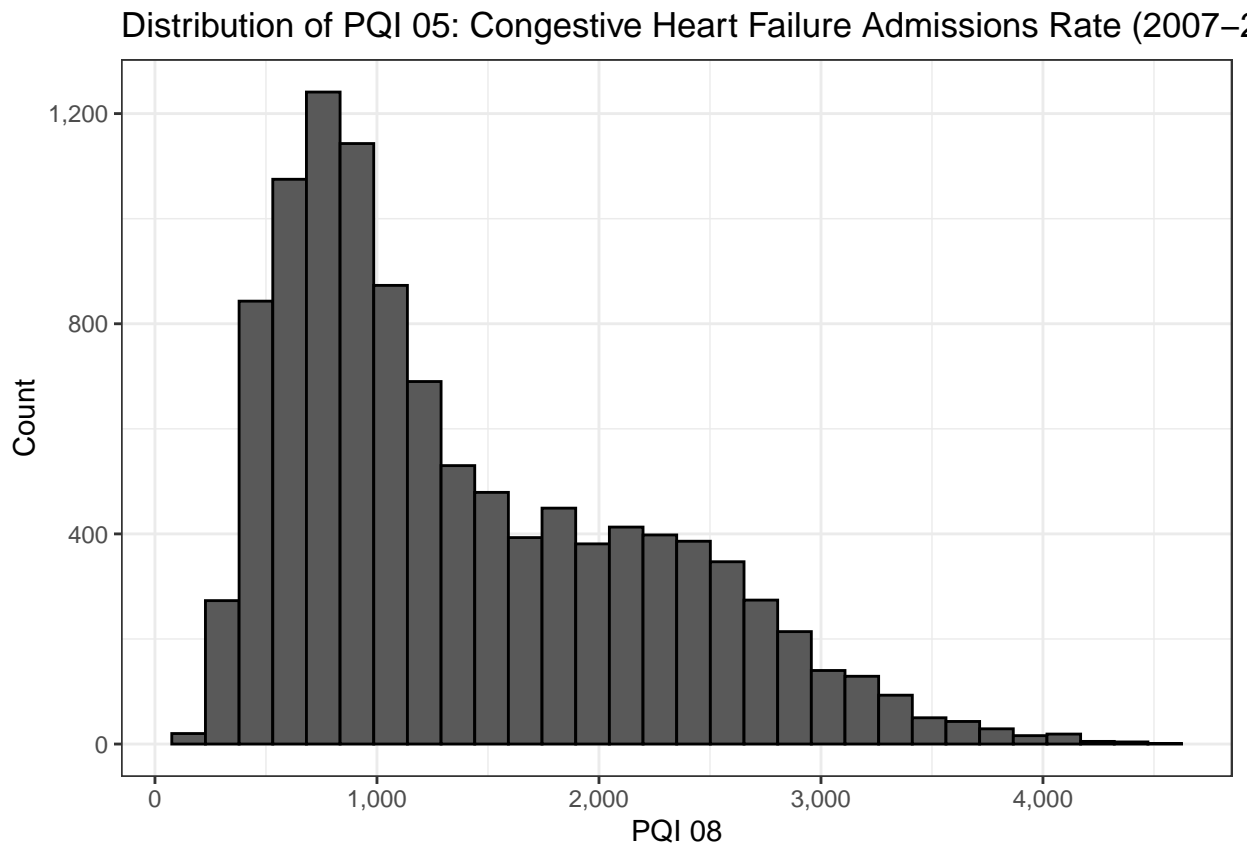
The distribution has a slight right skewe.

```
hrr %>%
  filter(!is.na(expansion_status)) %>%
  select(year, `pq108_chf_admission_rate_(age_65-74)`, `pq108_chf_admission_rate_(age_75+)`, `pq108_chf_
gather(age, pq108, -year) %>%
  mutate(
    age = str_sub(age, -10)
  ) %>% ggplot(aes( x = pq108)) + geom_histogram(color = "black") +
  scale_x_continuous(labels = scales::comma) +
  scale_y_continuous(labels = scales::comma) +
  xlab("PQI 08") +
  ylab("Count") +
  ggtitle("Distribution of PQI 05: Congestive Heart Failure Admissions Rate (2007-2018)") +
  theme_bw()
```

PQI08

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
## Warning: Removed 29 rows containing non-finite values (stat_bin).
```



This outcome is almost bi-modally normally distributed (or just has a large right skew).

Model Fitting

PQI05 Since these variables are rates, we will want to use Poisson models. There may be overdispersion, so I'm going to start fitting a negative binomial and then check if the poisson version has overdispersion.

```
pqi05_data <- hrr %>%  
  filter(!is.na(expansion_status)) %>%  
  select(year, hrr, state, expansion_status, percent_male, percent_african_american, `percent_hispanic`  
  gather(age, pqi05, c("pqi05_copd_or_asthma_in_older_adults_admission_rate_(age_40-64)", "pqi05_copd_o  
  mutate(age = str_extract(age, "(?<=age_).+(?=\\)"))
```

```
mod.nb_pqi05 <- MASS::glm.nb(pqi05 ~ `expansion_status` + offset(log(year)), data = pqi05_data)  
summary(mod.nb_pqi05)
```

Start with the base model

```
##  
## Call:  
## MASS::glm.nb(formula = pqi05 ~ expansion_status + offset(log(year)),  
##   data = pqi05_data, init.theta = 5.177273821, link = log)  
##  
## Deviance Residuals:  
##      Min       1Q   Median       3Q      Max  
## -4.0233  -0.8265  -0.1317   0.5245   3.4725
```

```
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -0.82680    0.01095  -75.47  <2e-16 ***
## expansion_status2  0.37158    0.01349   27.54  <2e-16 ***
## expansion_status3  0.46554    0.01482   31.41  <2e-16 ***
## expansion_status4  0.39108    0.01286   30.42  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(5.1773) family taken to be 1)
##
##      Null deviance: 12425  on 10975  degrees of freedom
## Residual deviance: 11328  on 10972  degrees of freedom
## (4 observations deleted due to missingness)
## AIC: 168003
##
## Number of Fisher Scoring iterations: 1
##
##              Theta:  5.1773
##             Std. Err.:  0.0681
##
## 2 x log-likelihood:  -167992.5890
```

We'll start adding in potential covariates and use likelihood ratio tests to determine if they should be added to the model.

```
mod.nb_pqi05_c1 <- MASS::glm.nb(pqi05 ~ `expansion_status` + percent_male + offset(log(year)), data = pqi05_data)
summary(mod.nb_pqi05_c1)
```

Percent Male

```
##
## Call:
## MASS::glm.nb(formula = pqi05 ~ expansion_status + percent_male +
##   offset(log(year)), data = pqi05_data, init.theta = 5.568802164,
##   link = log)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -4.0755  -0.8034  -0.1365   0.5030   3.7105
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    2.57558    0.11464   22.47  <2e-16 ***
## expansion_status2  0.35311    0.01304   27.09  <2e-16 ***
## expansion_status3  0.42303    0.01436   29.46  <2e-16 ***
## expansion_status4  0.35499    0.01247   28.46  <2e-16 ***
## percent_male    -7.52708    0.25176  -29.90  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(5.5688) family taken to be 1)
```

```
##
## Null deviance: 13360 on 10975 degrees of freedom
## Residual deviance: 11304 on 10971 degrees of freedom
## (4 observations deleted due to missingness)
## AIC: 167160
##
## Number of Fisher Scoring iterations: 1
##
##
## Theta: 5.5688
## Std. Err.: 0.0734
##
## 2 x log-likelihood: -167147.9930
```

```
anova(mod.nb_pqi05, mod.nb_pqi05_c1)
```

```
## Likelihood ratio tests of Negative Binomial Models
##
## Response: pqi05
##
## Model theta Resid. df
## 1 expansion_status + offset(log(year)) 5.177274 10972
## 2 expansion_status + percent_male + offset(log(year)) 5.568802 10971
## 2 x log-lik. Test df LR stat. Pr(Chi)
## 1 -167992.6
## 2 -167148.0 1 vs 2 1 844.5961 0
```

By the LRT, we add percent_male into the model.

```
mod.nb_pqi05_c2 <- MASS::glm.nb(pqi05 ~ `expansion_status` + percent_male + percent_african_american +
summary(mod.nb_pqi05_c2)
```

Percent African American

```
##
## Call:
## MASS::glm.nb(formula = pqi05 ~ expansion_status + percent_male +
## percent_african_american + offset(log(year)), data = pqi05_data,
## init.theta = 5.631183585, link = log)
##
## Deviance Residuals:
## Min 1Q Median 3Q Max
## -4.0396 -0.8006 -0.1413 0.4971 3.6948
##
## Coefficients:
## Estimate Std. Error z value Pr(>|z|)
## (Intercept) 2.05759 0.12253 16.79 <2e-16 ***
## expansion_status2 0.35165 0.01297 27.12 <2e-16 ***
## expansion_status3 0.40630 0.01437 28.27 <2e-16 ***
## expansion_status4 0.32730 0.01265 25.88 <2e-16 ***
## percent_male -6.45093 0.26736 -24.13 <2e-16 ***
## percent_african_american 0.58295 0.04963 11.75 <2e-16 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(5.6312) family taken to be 1)
```

```
##
## Null deviance: 13509 on 10975 degrees of freedom
## Residual deviance: 11301 on 10970 degrees of freedom
## (4 observations deleted due to missingness)
## AIC: 167033
##
## Number of Fisher Scoring iterations: 1
##
## Theta: 5.6312
## Std. Err.: 0.0743
##
## 2 x log-likelihood: -167019.1880
anova(mod.nb_pqi05_c1, mod.nb_pqi05_c2)

## Likelihood ratio tests of Negative Binomial Models
##
## Response: pqi05
##
## Model
## 1 expansion_status + percent_male + offset(log(year))
## 2 expansion_status + percent_male + percent_african_american + offset(log(year))
## theta Resid. df 2 x log-lik. Test df LR stat. Pr(Chi)
## 1 5.568802 10971 -167148.0
## 2 5.631184 10970 -167019.2 1 vs 2 1 128.805 0
```

By the LRT, we add in percent African American into the model.

```
mod.nb_pqi05_c3 <- MASS::glm.nb(pqi05 ~ `expansion_status` + percent_male + percent_african_american +
summary(mod.nb_pqi05_c3)
```

Percent Hispanic

```
##
## Call:
## MASS::glm.nb(formula = pqi05 ~ expansion_status + percent_male +
## percent_african_american + percent_hispanic + offset(log(year)),
## data = pqi05_data, init.theta = 5.641251527, link = log)
##
## Deviance Residuals:
## Min 1Q Median 3Q Max
## -4.0459 -0.8031 -0.1408 0.4974 3.7136
##
## Coefficients:
## Estimate Std. Error z value Pr(>|z|)
## (Intercept) 2.04810 0.12248 16.722 < 2e-16 ***
## expansion_status2 0.33723 0.01325 25.460 < 2e-16 ***
## expansion_status3 0.38883 0.01483 26.225 < 2e-16 ***
## expansion_status4 0.31815 0.01277 24.907 < 2e-16 ***
## percent_male -6.37993 0.26779 -23.824 < 2e-16 ***
## percent_african_american 0.56892 0.04966 11.455 < 2e-16 ***
## percent_hispanic -0.20229 0.04510 -4.486 7.27e-06 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## (Dispersion parameter for Negative Binomial(5.6413) family taken to be 1)
##
## Null deviance: 13533 on 10975 degrees of freedom
## Residual deviance: 11300 on 10969 degrees of freedom
## (4 observations deleted due to missingness)
## AIC: 167015
##
## Number of Fisher Scoring iterations: 1
##
## Theta: 5.6413
## Std. Err.: 0.0744
##
## 2 x log-likelihood: -166998.5920
anova(mod.nb_pqi05_c2, mod.nb_pqi05_c3)
```

```
## Likelihood ratio tests of Negative Binomial Models
##
## Response: pqi05
##
## 1 expansion_status + percent_male + percent_african_american + offset(log(year)) Model
## 2 expansion_status + percent_male + percent_african_american + percent_hispanic + offset(log(year))
## theta Resid. df 2 x log-lik. Test df LR stat. Pr(Chi)
## 1 5.631184 10970 -167019.2
## 2 5.641252 10969 -166998.6 1 vs 2 1 20.5957 5.672349e-06
```

By the LRT, we add in percent hispanic into the model.

Age category We add in the age category by which the data was originally split up.

```
mod.nb_pqi05_c4 <- MASS::glm.nb(pqi05 ~ `expansion_status` + percent_male + percent_african_american +
summary(mod.nb_pqi05_c4)
```

```
##
## Call:
## MASS::glm.nb(formula = pqi05 ~ expansion_status + percent_male +
## percent_african_american + percent_hispanic + age + offset(log(year)),
## data = pqi05_data, init.theta = 7.567523803, link = log)
##
## Deviance Residuals:
## Min 1Q Median 3Q Max
## -4.2129 -0.7692 -0.1276 0.5102 4.7410
##
## Coefficients:
## Estimate Std. Error z value Pr(>|z|)
## (Intercept) 2.125077 0.105965 20.054 < 2e-16 ***
## expansion_status2 0.343707 0.011449 30.019 < 2e-16 ***
## expansion_status3 0.400339 0.012816 31.239 < 2e-16 ***
## expansion_status4 0.322851 0.011042 29.238 < 2e-16 ***
## percent_male -6.060388 0.231438 -26.186 < 2e-16 ***
## percent_african_american 0.613227 0.042915 14.289 < 2e-16 ***
## percent_hispanic -0.172917 0.038978 -4.436 9.15e-06 ***
## age65-74 -0.536424 0.008529 -62.897 < 2e-16 ***
## age75+ -0.228711 0.008524 -26.832 < 2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(7.5675) family taken to be 1)
##
##      Null deviance: 18120  on 10975  degrees of freedom
## Residual deviance: 11222  on 10967  degrees of freedom
##      (4 observations deleted due to missingness)
## AIC: 163655
##
## Number of Fisher Scoring iterations: 1
##
##
##              Theta:  7.568
##              Std. Err.:  0.101
##
## 2 x log-likelihood:  -163634.999
anova(mod.nb_pqi05_c3, mod.nb_pqi05_c4)

## Likelihood ratio tests of Negative Binomial Models
##
## Response: pqi05
##
## 1      expansion_status + percent_male + percent_african_american + percent_hispanic + offset(log(y
## 2 expansion_status + percent_male + percent_african_american + percent_hispanic + age + offset(log(y
##      theta Resid. df    2 x log-lik.  Test      df LR stat. Pr(Chi)
## 1 5.641252      10969      -166998.6
## 2 7.567524      10967      -163635.0 1 vs 2      2 3363.594      0
```

By the LRT, we add age category into the model.

```
mod.nb_pqi05_c5 <- MASS::glm.nb(pqi05 ~ `expansion_status` + percent_male + percent_african_american +
summary(mod.nb_pqi05_c5)
```

Percent Eligible for Medicaid

```
##
## Call:
## MASS::glm.nb(formula = pqi05 ~ expansion_status + percent_male +
##      percent_african_american + percent_hispanic + age + percent_eligible_for_medicaid +
##      offset(log(year)), data = pqi05_data, init.theta = 8.35361968,
##      link = log)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -4.4650  -0.7770  -0.1415   0.5348   4.1859
##
## Coefficients:
##
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      2.423078   0.100970    24.00 <2e-16 ***
## expansion_status2    0.363912   0.010939    33.27 <2e-16 ***
## expansion_status3    0.401069   0.012203    32.87 <2e-16 ***
## expansion_status4    0.362050   0.010593    34.18 <2e-16 ***
## percent_male      -7.509902   0.222524   -33.75 <2e-16 ***
## percent_african_american  0.031092   0.044387    0.70  0.484
```



```
## percent_hispanic          -0.835947   0.041843  -19.98   <2e-16 ***
## age65-74                  -0.547949   0.008121  -67.47   <2e-16 ***
## age75+                    -0.232240   0.008116  -28.62   <2e-16 ***
## percent_eligible_for_medicaid  1.947953   0.055985   34.79   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(8.3536) family taken to be 1)
##
## Null deviance: 19987  on 10975  degrees of freedom
## Residual deviance: 11199  on 10966  degrees of freedom
## (4 observations deleted due to missingness)
## AIC: 162535
##
## Number of Fisher Scoring iterations: 1
##
##
##          Theta:  8.354
##        Std. Err.:  0.112
##
## 2 x log-likelihood:  -162512.753
```

```
anova(mod.nb_pqi05_c4, mod.nb_pqi05_c5)
```

```
## Likelihood ratio tests of Negative Binomial Models
```

```
##
```

```
## Response: pqi05
```

```
##
```

```
## 1 expansion_status + percent_male + percent_african_american + percent_hispanic + age + percent_eligible_for_medicaid
```

```
## 2 expansion_status + percent_male + percent_african_american + percent_hispanic + age + percent_eligible_for_medicaid_2 + offset(log(year))
```

```
##      theta Resid. df    2 x log-lik. Test      df LR stat. Pr(Chi)
```

```
## 1 7.567524    10967    -163635.0
```

```
## 2 8.353620    10966    -162512.8 1 vs 2      1 1122.246      0
```

By the LRT, we add in the percent eligible for Medicaid.

```
mod.nb_pqi05_1 <- MASS::glm.nb(pqi05 ~ `expansion_status` + percent_male + percent_african_american + percent_hispanic + age + percent_eligible_for_medicaid + percent_eligible_for_medicaid_2 + offset(log(year)), data = pqi05_data %>%
summary(mod.nb_pqi05_1)
```

Check for quadratic effects of percent_eligible_for_medicaid

```
##
```

```
## Call:
```

```
## MASS::glm.nb(formula = pqi05 ~ expansion_status + percent_male +
##   percent_african_american + percent_hispanic + age + percent_eligible_for_medicaid +
##   percent_eligible_for_medicaid_2 + offset(log(year)), data = pqi05_data %>%
##   mutate(percent_eligible_for_medicaid_2 = percent_eligible_for_medicaid^2),
##   init.theta = 8.483111512, link = log)
```

```
##
```

```
## Deviance Residuals:
```

```
##      Min       1Q   Median       3Q      Max
## -4.5275  -0.7872  -0.1388   0.5317   4.2223
```

```
##
```

```
## Coefficients:
```

```
##              Estimate Std. Error z value Pr(>|z|)
```

```
## (Intercept)                2.180570    0.101768   21.427   <2e-16 ***
## expansion_status2          0.383544    0.010965   34.978   <2e-16 ***
## expansion_status3          0.411703    0.012138   33.917   <2e-16 ***
## expansion_status4          0.373298    0.010552   35.376   <2e-16 ***
## percent_male               -7.685798    0.221278  -34.734   <2e-16 ***
## percent_african_american    0.032957    0.044056    0.748    0.454
## percent_hispanic           -0.524958    0.047027  -11.163   <2e-16 ***
## age65-74                   -0.549280    0.008059  -68.157   <2e-16 ***
## age75+                     -0.231903    0.008054  -28.794   <2e-16 ***
## percent_eligible_for_medicaid  4.658809    0.209431   22.245   <2e-16 ***
## percent_eligible_for_medicaid_2 -5.633921    0.418303  -13.469   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(8.4831) family taken to be 1)
##
## Null deviance: 20294 on 10975 degrees of freedom
## Residual deviance: 11196 on 10965 degrees of freedom
## (4 observations deleted due to missingness)
## AIC: 162363
##
## Number of Fisher Scoring iterations: 1
##
##
##          Theta:  8.483
##        Std. Err.:  0.113
##
## 2 x log-likelihood: -162339.028
```

```
anova(mod.nb_pqi05_c4, mod.nb_pqi05_1)
```

```
## Likelihood ratio tests of Negative Binomial Models
```

```
##
```

```
## Response: pqi05
```

```
##
```

```
## 1
```

```
expansion_status + percent_male +
```

```
## 2 expansion_status + percent_male + percent_african_american + percent_hispanic + age + percent_eligible_for_medicaid +
```

```
##      theta Resid. df      2 x log-lik. Test      df LR stat. Pr(Chi)
```

```
## 1 7.567524      10967      -163635
```

```
## 2 8.483112      10965      -162339 1 vs 2      2 1295.971      0
```

By the LRT, we add the quadratic effects of the percent eligible for medicaid.

Effect Modification We think there may be effect modification between % eligible for medicaid and expansion status, since in general Medicaid expansion means a higher percentage of the population qualifies for Medicaid.

```
mod.nb_pqi05_2 <- MASS::glm.nb(pqi05 ~ `expansion_status` + percent_male + percent_african_american + percent_hispanic + age + percent_eligible_for_medicaid + percent_eligible_for_medicaid_2 + expansion_status * percent_eligible_for_medicaid + offset(log(year)), data = pqi05_data %>% mutate(percent_eligible_for_medicaid_2 = percent_eligible_for_medicaid^2))
summary(mod.nb_pqi05_2)
```

```
##
```

```
## Call:
```

```
## MASS::glm.nb(formula = pqi05 ~ expansion_status + percent_male +
```

```
##      percent_african_american + percent_hispanic + age + percent_eligible_for_medicaid +
```

```
##      percent_eligible_for_medicaid_2 + expansion_status * percent_eligible_for_medicaid +
```

```
##      offset(log(year)), data = pqi05_data %>% mutate(percent_eligible_for_medicaid_2 = percent_eligible_for_medicaid^2))
```

```

##      init.theta = 8.568631037, link = log)
##
## Deviance Residuals:
##      Min        1Q      Median        3Q        Max
## -4.4280  -0.7852  -0.1479   0.5342   4.1484
##
## Coefficients:
##                                     Estimate Std. Error z value
## (Intercept)                      2.000326   0.103475  19.331
## expansion_status2                  0.308328   0.033096   9.316
## expansion_status3                  0.694193   0.041038  16.916
## expansion_status4                  0.300265   0.031712   9.469
## percent_male                     -7.231498   0.226164 -31.975
## percent_african_american          0.154790   0.045311   3.416
## percent_hispanic                 -0.582027   0.047471 -12.261
## age65-74                        -0.549445   0.008019 -68.517
## age75+                          -0.231721   0.008014 -28.916
## percent_eligible_for_medicaid     4.524902   0.237064  19.087
## percent_eligible_for_medicaid_2  -5.538923   0.421537 -13.140
## expansion_status2:percent_eligible_for_medicaid 0.342275   0.141841   2.413
## expansion_status3:percent_eligible_for_medicaid -1.357398   0.178489  -7.605
## expansion_status4:percent_eligible_for_medicaid 0.300946   0.131443   2.290
##                                     Pr(>|z|)
## (Intercept)                       < 2e-16 ***
## expansion_status2                  < 2e-16 ***
## expansion_status3                  < 2e-16 ***
## expansion_status4                  < 2e-16 ***
## percent_male                      < 2e-16 ***
## percent_african_american          0.000635 ***
## percent_hispanic                  < 2e-16 ***
## age65-74                         < 2e-16 ***
## age75+                           < 2e-16 ***
## percent_eligible_for_medicaid     < 2e-16 ***
## percent_eligible_for_medicaid_2   < 2e-16 ***
## expansion_status2:percent_eligible_for_medicaid 0.015818 *
## expansion_status3:percent_eligible_for_medicaid 2.85e-14 ***
## expansion_status4:percent_eligible_for_medicaid 0.022047 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(8.5686) family taken to be 1)
##
##      Null deviance: 20497  on 10975  degrees of freedom
## Residual deviance: 11194  on 10962  degrees of freedom
##      (4 observations deleted due to missingness)
## AIC: 162256
##
## Number of Fisher Scoring iterations: 1
##
##
##              Theta:  8.569
##             Std. Err.: 0.114
##
## 2 x log-likelihood: -162225.780

```

```
anova(mod.nb_pqi05_1, mod.nb_pqi05_2)
```

```
## Likelihood ratio tests of Negative Binomial Models
```

```
##
```

```
## Response: pqi05
```

```
##
```

```
## 1
```

```
expansion_status + percent_male + percent_african_american + percent_hispanic + age + percent_eligible
```

```
## 2 expansion_status + percent_male + percent_african_american + percent_hispanic + age + percent_eligible
```

```
##      theta Resid. df      2 x log-lik. Test      df LR stat. Pr(Chi)
```

```
## 1 8.483112      10965      -162339.0
```

```
## 2 8.568631      10962      -162225.8 1 vs 2      3 113.2482      0
```

By the LRT, we add the interaction term to the model.

```
AIC(mod.nb_pqi05)
```

Model Diagnostics

```
## [1] 168002.6
```

```
AIC(mod.nb_pqi05_c1)
```

```
## [1] 167160
```

```
AIC(mod.nb_pqi05_c2)
```

```
## [1] 167033.2
```

```
AIC(mod.nb_pqi05_c3)
```

```
## [1] 167014.6
```

```
AIC(mod.nb_pqi05_c4)
```

```
## [1] 163655
```

```
AIC(mod.nb_pqi05_c5)
```

```
## [1] 162534.8
```

```
AIC(mod.nb_pqi05_1)
```

```
## [1] 162363
```

```
AIC(mod.nb_pqi05_2)
```

```
## [1] 162255.8
```

The full model has the lowest AIC.

```
BIC(mod.nb_pqi05)
```

```
## [1] 168039.1
```

```
BIC(mod.nb_pqi05_c1)
```

```
## [1] 167203.8
```

```
BIC(mod.nb_pqi05_c2)
```

```
## [1] 167084.3
```

```
BIC(mod.nb_pqi05_c3)
```

```
## [1] 167073
```

```
BIC(mod.nb_pqi05_c4)
```

```
## [1] 163728
```

```
BIC(mod.nb_pqi05_c5)
```

```
## [1] 162615.1
```

```
BIC(mod.nb_pqi05_1)
```

```
## [1] 162450.7
```

```
BIC(mod.nb_pqi05_2)
```

```
## [1] 162365.3
```

The full model has the lowest AIC and BIC of all the models.

Check Poisson model As a last step, we'll check the model as a Poisson model, and check if there is truly overdispersion.

```
mod.pois_pqi05 <- glm(pqi05 ~ expansion_status + percent_male +  
  percent_african_american + percent_hispanic + age + percent_eligible_for_medicaid +  
  percent_eligible_for_medicaid_2 + expansion_status * percent_eligible_for_medicaid +  
  offset(log(year)), data = pqi05_data %>% mutate(percent_eligible_for_medicaid_2 = percent_eligible_
```

```
# Evaluate overdispersion
```

```
deviance(mod.pois_pqi05)/mod.pois_pqi05$df.residual # 138.7201
```

```
## [1] 138.7201
```

```
pearson.stat_pqi05 <- sum((na.omit(pqi05_data)$pqi05 - fitted(mod.pois_pqi05))^2/fitted(mod.pois_pqi05))  
pearson.stat_pqi05/mod.pois_pqi05$df.residual # 144.1579
```

```
## [1] 144.1579
```

```
# Evidence of overdispersion
```

```
mean(pqi05_data$pqi05, na.rm = T) # 1249.904
```

```
## [1] 1249.904
```

```
var(pqi05_data$pqi05, na.rm = T) # 325943.6
```

```
## [1] 325943.6
```

By these, we can say the negative binomial is the right choice.

```
exp(coef(mod.nb_pqi05_2))
```

Interpretation

```
## (Intercept)  
## 7.391464e+00  
## expansion_status2  
## 1.361147e+00  
## expansion_status3
```

```
##                2.002092e+00
##                expansion_status4
##                1.350216e+00
##                percent_male
##                7.234365e-04
##                percent_african_american
##                1.167413e+00
##                percent_hispanic
##                5.587647e-01
##                age65-74
##                5.772700e-01
##                age75+
##                7.931671e-01
##                percent_eligible_for_medicaid
##                9.228691e+01
##                percent_eligible_for_medicaid_2
##                3.930759e-03
## expansion_status2:percent_eligible_for_medicaid
##                1.408148e+00
## expansion_status3:percent_eligible_for_medicaid
##                2.573296e-01
## expansion_status4:percent_eligible_for_medicaid
##                1.351136e+00
```

```
exp(confint(mod.nb_pqi05_2))
```

```
## Waiting for profiling to be done...
```

```
##                2.5 %      97.5 %
## (Intercept)      6.034390e+00 9.054446e+00
## expansion_status2 1.276179e+00 1.451597e+00
## expansion_status3 1.847346e+00 2.169706e+00
## expansion_status4 1.270277e+00 1.435093e+00
## percent_male      4.629838e-04 1.130631e-03
## percent_african_american 1.064846e+00 1.280048e+00
## percent_hispanic  5.079822e-01 6.148085e-01
## age65-74          5.682550e-01 5.864281e-01
## age75+            7.807896e-01 8.057407e-01
## percent_eligible_for_medicaid 5.806953e+01 1.464244e+02
## percent_eligible_for_medicaid_2 1.709239e-03 9.064200e-03
## expansion_status2:percent_eligible_for_medicaid 1.066727e+00 1.859120e+00
## expansion_status3:percent_eligible_for_medicaid 1.811176e-01 3.657415e-01
## expansion_status4:percent_eligible_for_medicaid 1.049344e+00 1.739046e+00
```

PQI08 Since these variables are rates, we will want to use Poisson models. There may be overdispersion, so I'm going to start fitting a negative binomial and then check if the poisson version has overdispersion.

```
pqi08_data <- hrr %>%
  filter(!is.na(expansion_status)) %>%
  select(year, hrr, state, expansion_status, percent_male, percent_african_american, `percent_hispanic`
  gather(age, pqi08, c("pqi08_chf_admission_rate_(age_<_65)", "pqi08_chf_admission_rate_(age_65-74)", "
  mutate(age = str_extract(age, "(?<=age_).+(?=\\)"))

mod.nb_pqi08 <- MASS::glm.nb(pqi08 ~ `expansion_status` + `percent_male` + `percent_african_american` +
```

```
summary(mod.nb_pqi08)
```

```
##
## Call:
## MASS::glm.nb(formula = pqi08 ~ expansion_status + percent_male +
##   percent_african_american + percent_hispanic + percent_eligible_for_medicaid +
##   age + offset(log(year)), data = pqi08_data, init.theta = 12.04476341,
##   link = log)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -4.0260  -0.7317  -0.0961   0.5478   4.4076
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      0.842115   0.084365   9.982  <2e-16 ***
## expansion_status2  0.159481   0.009139  17.450  <2e-16 ***
## expansion_status3  0.256869   0.010189  25.211  <2e-16 ***
## expansion_status4  0.146708   0.008843  16.589  <2e-16 ***
## percent_male      -4.439616   0.185970 -23.873  <2e-16 ***
## percent_african_american  1.746731   0.037032  47.168  <2e-16 ***
## percent_hispanic   0.312481   0.034895   8.955  <2e-16 ***
## percent_eligible_for_medicaid 0.775592   0.046753  16.589  <2e-16 ***
## age65-74          -0.274105   0.006796 -40.332  <2e-16 ***
## age75+             0.819164   0.006779 120.847  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(12.0448) family taken to be 1)
##
##    Null deviance: 45950  on 10950  degrees of freedom
## Residual deviance: 11118  on 10941  degrees of freedom
## (29 observations deleted due to missingness)
## AIC: 159048
##
## Number of Fisher Scoring iterations: 1
##
##              Theta: 12.045
##            Std. Err.: 0.163
##
## 2 x log-likelihood: -159025.919
```

```
mod.nb_pqi08 <- MASS::glm.nb(pqi08 ~ `expansion_status` + offset(log(year)), data = pqi08_data)
summary(mod.nb_pqi08)
```

Start with the base model

```
##
## Call:
## MASS::glm.nb(formula = pqi08 ~ expansion_status + offset(log(year)),
##   data = pqi08_data, init.theta = 3.068864655, link = log)
##
```

```
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.7065  -0.9817  -0.3512   0.6556   2.7778
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -0.53722     0.01422  -37.783 < 2e-16 ***
## expansion_status2  0.13273     0.01752   7.574 3.62e-14 ***
## expansion_status3  0.30074     0.01925  15.627 < 2e-16 ***
## expansion_status4  0.19609     0.01669  11.751 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(3.0689) family taken to be 1)
##
##      Null deviance: 11800  on 10950  degrees of freedom
## Residual deviance: 11537  on 10947  degrees of freedom
## (29 observations deleted due to missingness)
## AIC: 174770
##
## Number of Fisher Scoring iterations: 1
##
##
##              Theta:  3.0689
##              Std. Err.:  0.0395
##
## 2 x log-likelihood:  -174759.5180
```

We'll start adding in potential covariates and use likelihood ratio tests to determine if they should be added to the model.

```
mod.nb_pqi08_c1 <- MASS::glm.nb(pqi08 ~ `expansion_status` + percent_male + offset(log(year)), data = pqi08_data)
summary(mod.nb_pqi08_c1)
```

Percent Male

```
##
## Call:
## MASS::glm.nb(formula = pqi08 ~ expansion_status + percent_male +
##   offset(log(year)), data = pqi08_data, init.theta = 3.179707276,
##   link = log)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.5886  -1.0103  -0.3650   0.6679   2.7211
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)     2.63963     0.15177  17.392 < 2e-16 ***
## expansion_status2  0.10928     0.01725   6.334 2.39e-10 ***
## expansion_status3  0.26118     0.01899  13.751 < 2e-16 ***
## expansion_status4  0.16276     0.01650   9.867 < 2e-16 ***
## percent_male     -7.02396     0.33330 -21.074 < 2e-16 ***
## ---
```



```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(3.1797) family taken to be 1)
##
## Null deviance: 12225  on 10950  degrees of freedom
## Residual deviance: 11517  on 10946  degrees of freedom
## (29 observations deleted due to missingness)
## AIC: 174344
##
## Number of Fisher Scoring iterations: 1
##
##
##          Theta:  3.1797
##        Std. Err.:  0.0410
##
## 2 x log-likelihood:  -174331.5630
anova(mod.nb_pqi08, mod.nb_pqi08_c1)
```

```
## Likelihood ratio tests of Negative Binomial Models
##
## Response: pqi08
##
##              Model      theta Resid. df
## 1      expansion_status + offset(log(year)) 3.068865      10947
## 2 expansion_status + percent_male + offset(log(year)) 3.179707      10946
##      2 x log-lik.  Test      df LR stat. Pr(Chi)
## 1      -174759.5
## 2      -174331.6 1 vs 2      1  427.955      0
```

By the LRT, we add percent_male into the model.

```
mod.nb_pqi08_c2 <- MASS::glm.nb(pqi08 ~ `expansion_status` + percent_male + percent_african_american +
summary(mod.nb_pqi08_c2)
```

Percent African American

```
##
## Call:
## MASS::glm.nb(formula = pqi08 ~ expansion_status + percent_male +
##   percent_african_american + offset(log(year)), data = pqi08_data,
##   init.theta = 3.33540608, link = log)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.5399  -0.9869  -0.4287   0.6859   2.8502
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      1.17537    0.15926   7.380 1.58e-13 ***
## expansion_status2    0.10407    0.01685   6.177 6.55e-10 ***
## expansion_status3    0.20454    0.01866  10.960 < 2e-16 ***
## expansion_status4    0.08606    0.01642   5.241 1.59e-07 ***
## percent_male      -3.97708    0.34749 -11.445 < 2e-16 ***
## percent_african_american 1.57386    0.06445  24.418 < 2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(3.3354) family taken to be 1)
##
##      Null deviance: 12822  on 10950  degrees of freedom
## Residual deviance: 11491  on 10945  degrees of freedom
##      (29 observations deleted due to missingness)
## AIC: 173771
##
## Number of Fisher Scoring iterations: 1
##
##
##              Theta:  3.3354
##              Std. Err.:  0.0431
##
## 2 x log-likelihood:  -173757.2660
anova(mod.nb_pqi08_c1, mod.nb_pqi08_c2)

## Likelihood ratio tests of Negative Binomial Models
##
## Response: pqi08
##
##                                     Model
## 1                                expansion_status + percent_male + offset(log(year))
## 2 expansion_status + percent_male + percent_african_american + offset(log(year))
##      theta Resid. df    2 x log-lik.  Test    df LR stat. Pr(Chi)
## 1 3.179707    10946      -174331.6
## 2 3.335406    10945      -173757.3 1 vs 2      1 574.2976      0
```

By the LRT, we add in percent African American into the model.

```
mod.nb_pqi08_c3 <- MASS::glm.nb(pqi08 ~ `expansion_status` + percent_male + percent_african_american +
summary(mod.nb_pqi08_c3)
```

Percent Hispanic

```
##
## Call:
## MASS::glm.nb(formula = pqi08 ~ expansion_status + percent_male +
##      percent_african_american + percent_hispanic + offset(log(year)),
##      data = pqi08_data, init.theta = 3.355400869, link = log)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.5246  -0.9809  -0.4301   0.6874   2.9193
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      1.22334    0.15886   7.701 1.35e-14 ***
## expansion_status2    0.13623    0.01718   7.932 2.16e-15 ***
## expansion_status3    0.24307    0.01922  12.649 < 2e-16 ***
## expansion_status4    0.10251    0.01655   6.194 5.86e-10 ***
## percent_male       -4.19865    0.34733 -12.088 < 2e-16 ***
## percent_african_american 1.61771    0.06436  25.134 < 2e-16 ***
## percent_hispanic     0.46665    0.05840   7.991 1.34e-15 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(3.3554) family taken to be 1)
##
##      Null deviance: 12898  on 10950  degrees of freedom
## Residual deviance: 11487  on 10944  degrees of freedom
##      (29 observations deleted due to missingness)
## AIC: 173702
##
## Number of Fisher Scoring iterations: 1
##
##
##              Theta:  3.3554
##             Std. Err.:  0.0434
##
## 2 x log-likelihood:  -173685.6210
anova(mod.nb_pqi08_c2, mod.nb_pqi08_c3)

## Likelihood ratio tests of Negative Binomial Models
##
## Response: pqi08
##
## 1                                expansion_status + percent_male + percent_african_american + offset(log(year))
## 2 expansion_status + percent_male + percent_african_american + percent_hispanic + offset(log(year))
##      theta Resid. df    2 x log-lik.  Test      df LR stat. Pr(Chi)
## 1 3.335406    10945      -173757.3
## 2 3.355401    10944      -173685.6 1 vs 2      1 71.64473      0
```

By the LRT, we add percent hispanic to the model.

```
mod.nb_pqi08_c4 <- MASS::glm.nb(pqi08 ~ `expansion_status` + percent_male + percent_african_american +
summary(mod.nb_pqi08_c4)
```

Age category

```
##
## Call:
## MASS::glm.nb(formula = pqi08 ~ expansion_status + percent_male +
##      percent_african_american + percent_hispanic + age + offset(log(year)),
##      data = pqi08_data, init.theta = 11.76069359, link = log)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -3.9405  -0.7316  -0.0925   0.5477   4.0962
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    0.728074   0.085302   8.535  <2e-16 ***
## expansion_status2  0.149780   0.009217  16.250  <2e-16 ***
## expansion_status3  0.257508   0.010309  24.979  <2e-16 ***
## expansion_status4  0.128954   0.008881  14.520  <2e-16 ***
## percent_male     -3.877336   0.186322 -20.810  <2e-16 ***
## percent_african_american 1.975012   0.034488  57.267  <2e-16 ***
```

```
## percent_hispanic      0.579847   0.031306  18.522   <2e-16 ***
## age65-74              -0.267488   0.006877 -38.898   <2e-16 ***
## age75+                0.824019   0.006859 120.134   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(11.7607) family taken to be 1)
##
## Null deviance: 44878  on 10950  degrees of freedom
## Residual deviance: 11122  on 10942  degrees of freedom
## (29 observations deleted due to missingness)
## AIC: 159312
##
## Number of Fisher Scoring iterations: 1
##
##
##          Theta:  11.761
##        Std. Err.:  0.159
##
## 2 x log-likelihood: -159292.229
```

```
anova(mod.nb_pqi08_c3, mod.nb_pqi08_c4)
```

```
## Likelihood ratio tests of Negative Binomial Models
```

```
##
```

```
## Response: pqi08
```

```
##
```

```
## 1      expansion_status + percent_male + percent_african_american + percent_hispanic + offset(log(y
```

```
## 2 expansion_status + percent_male + percent_african_american + percent_hispanic + age + offset(log(y
```

```
##      theta Resid. df      2 x log-lik.  Test      df LR stat. Pr(Chi)
```

```
## 1  3.355401      10944      -173685.6
```

```
## 2 11.760694      10942      -159292.2 1 vs 2      2 14393.39      0
```

By the LRT, we add in age category to the model.

```
mod.nb_pqi08_c5 <- MASS::glm.nb(pqi08 ~ `expansion_status` + percent_male + percent_african_american +
summary(mod.nb_pqi08_c5)
```

Percent Eligible for Medicaid

```
##
```

```
## Call:
```

```
## MASS::glm.nb(formula = pqi08 ~ expansion_status + percent_male +
##   percent_african_american + percent_hispanic + age + percent_eligible_for_medicaid +
##   offset(log(year)), data = pqi08_data, init.theta = 12.04476341,
##   link = log)
```

```
##
```

```
## Deviance Residuals:
```

```
##      Min      1Q   Median      3Q      Max
```

```
## -4.0260 -0.7317 -0.0961  0.5478  4.4076
```

```
##
```

```
## Coefficients:
```

```
##              Estimate Std. Error z value Pr(>|z|)
```

```
## (Intercept)      0.842115   0.084365   9.982   <2e-16 ***
```

```
## expansion_status2      0.159481   0.009139  17.450   <2e-16 ***
```

```
## expansion_status3          0.256869   0.010189  25.211   <2e-16 ***
## expansion_status4          0.146708   0.008843  16.589   <2e-16 ***
## percent_male               -4.439616   0.185970 -23.873   <2e-16 ***
## percent_african_american    1.746731   0.037032  47.168   <2e-16 ***
## percent_hispanic            0.312481   0.034895   8.955   <2e-16 ***
## age65-74                   -0.274105   0.006796 -40.332   <2e-16 ***
## age75+                     0.819164   0.006779 120.847   <2e-16 ***
## percent_eligible_for_medicaid 0.775592   0.046753  16.589   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(12.0448) family taken to be 1)
##
## Null deviance: 45950 on 10950 degrees of freedom
## Residual deviance: 11118 on 10941 degrees of freedom
## (29 observations deleted due to missingness)
## AIC: 159048
##
## Number of Fisher Scoring iterations: 1
##
## Theta: 12.045
## Std. Err.: 0.163
##
## 2 x log-likelihood: -159025.919
```

```
anova(mod.nb_pqi08_c4, mod.nb_pqi08_c5)
```

```
## Likelihood ratio tests of Negative Binomial Models
##
## Response: pqi08
##
## 1 expansion_status + percent_male + percent_african_american + percent_hispanic + age + percent_eligible_for_medicaid
## 2 expansion_status + percent_male + percent_african_american + percent_hispanic + age + percent_eligible_for_medicaid_2 + offset(log(year))
## theta Resid. df 2 x log-lik. Test df LR stat. Pr(Chi)
## 1 11.76069 10942 -159292.2
## 2 12.04476 10941 -159025.9 1 vs 2 1 266.3096 0
```

By the LRT we add in percent eligible for medicaid into the model.

```
mod.nb_pqi08_1 <- MASS::glm.nb(pqi08 ~ `expansion_status` + percent_male + percent_african_american + percent_hispanic + age + percent_eligible_for_medicaid + percent_eligible_for_medicaid_2 + offset(log(year)), data = pqi08_data %>%
summary(mod.nb_pqi08_1)
```

Check for quadratic effects of percent_eligible_for_medicaid

```
##
## Call:
## MASS::glm.nb(formula = pqi08 ~ expansion_status + percent_male +
## percent_african_american + percent_hispanic + age + percent_eligible_for_medicaid +
## percent_eligible_for_medicaid_2 + offset(log(year)), data = pqi08_data %>%
## mutate(percent_eligible_for_medicaid_2 = percent_eligible_for_medicaid^2),
## init.theta = 12.1405392, link = log)
##
## Deviance Residuals:
## Min 1Q Median 3Q Max
```

```
## -4.0739 -0.7263 -0.0963 0.5450 4.3951
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    0.700635   0.085325   8.211  <2e-16 ***
## expansion_status2 0.170850   0.009195  18.581  <2e-16 ***
## expansion_status3 0.263183   0.010173  25.872  <2e-16 ***
## expansion_status4 0.153197   0.008842  17.325  <2e-16 ***
## percent_male    -4.552656   0.185637 -24.525  <2e-16 ***
## percent_african_american 1.749337   0.036892  47.417  <2e-16 ***
## percent_hispanic 0.501013   0.039369  12.726  <2e-16 ***
## age65-74        -0.274766   0.006770 -40.587  <2e-16 ***
## age75+          0.819160   0.006752 121.321  <2e-16 ***
## percent_eligible_for_medicaid 2.397809   0.175574  13.657  <2e-16 ***
## percent_eligible_for_medicaid_2 -3.370766   0.350466  -9.618  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(12.1405) family taken to be 1)
##
## Null deviance: 46312 on 10950 degrees of freedom
## Residual deviance: 11117 on 10940 degrees of freedom
## (29 observations deleted due to missingness)
## AIC: 158962
##
## Number of Fisher Scoring iterations: 1
##
##              Theta: 12.141
##            Std. Err.: 0.164
##
## 2 x log-likelihood: -158937.946
```

```
anova(mod.nb_pqi08_c5, mod.nb_pqi08_1)
```

```
## Likelihood ratio tests of Negative Binomial Models
```

```
##
```

```
## Response: pqi08
```

```
##
```

```
## 1 expansion_status + percent_male + percent_african_american + per
```

```
## 2 expansion_status + percent_male + percent_african_american + percent_hispanic + age + percent_elig
```

```
##      theta Resid. df    2 x log-lik. Test    df LR stat. Pr(Chi)
```

```
## 1 12.04476    10941    -159025.9
```

```
## 2 12.14054    10940    -158937.9 1 vs 2      1 87.97355      0
```

By the LRT we add the quadratic effects of percent eligible for medicaid to the model.

Effect Modification We think there may be effect modification between % eligible for medicaid and expansion status, since in general Medicaid expansion means a higher percentage of the population qualifies for Medicaid.

```
mod.nb_pqi08_2 <- MASS::glm.nb(pqi08 ~ `expansion_status` + percent_male + percent_african_american + per
summary(mod.nb_pqi08_2)
```

```
##
```

```
## Call:
```

```

## MASS::glm.nb(formula = pqi08 ~ expansion_status + percent_male +
##   percent_african_american + percent_hispanic + age + percent_eligible_for_medicaid +
##   percent_eligible_for_medicaid_2 + expansion_status * percent_eligible_for_medicaid +
##   offset(log(year)), data = pqi08_data %>% mutate(percent_eligible_for_medicaid_2 = percent_eligib
##   init.theta = 12.16674904, link = log)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -4.0867  -0.7258  -0.0940   0.5477   4.3780
##
## Coefficients:
##                                     Estimate Std. Error z value
## (Intercept)                        0.634268   0.087106   7.282
## expansion_status2                   0.186611   0.027882   6.693
## expansion_status3                   0.404622   0.034528  11.719
## expansion_status4                   0.210895   0.026696   7.900
## percent_male                       -4.493580   0.190483 -23.590
## percent_african_american            1.786420   0.038095  46.893
## percent_hispanic                   0.495312   0.039903  12.413
## age65-74                          -0.274424   0.006763 -40.579
## age75+                             0.819672   0.006745 121.526
## percent_eligible_for_medicaid      2.528170   0.199480  12.674
## percent_eligible_for_medicaid_2   -3.254435   0.354462  -9.181
## expansion_status2:percent_eligible_for_medicaid -0.047076   0.119383  -0.394
## expansion_status3:percent_eligible_for_medicaid -0.643629   0.150071  -4.289
## expansion_status4:percent_eligible_for_medicaid -0.260566   0.110566  -2.357
##                                     Pr(>|z|)
## (Intercept)                        3.30e-13 ***
## expansion_status2                   2.19e-11 ***
## expansion_status3                   < 2e-16 ***
## expansion_status4                   2.79e-15 ***
## percent_male                       < 2e-16 ***
## percent_african_american            < 2e-16 ***
## percent_hispanic                   < 2e-16 ***
## age65-74                          < 2e-16 ***
## age75+                             < 2e-16 ***
## percent_eligible_for_medicaid      < 2e-16 ***
## percent_eligible_for_medicaid_2    < 2e-16 ***
## expansion_status2:percent_eligible_for_medicaid 0.6933
## expansion_status3:percent_eligible_for_medicaid 1.80e-05 ***
## expansion_status4:percent_eligible_for_medicaid 0.0184 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(12.1667) family taken to be 1)
##
## Null deviance: 46411 on 10950 degrees of freedom
## Residual deviance: 11116 on 10937 degrees of freedom
## (29 observations deleted due to missingness)
## AIC: 158944
##
## Number of Fisher Scoring iterations: 1
##
##

```

```
##              Theta: 12.167
##            Std. Err.: 0.164
##
## 2 x log-likelihood: -158913.782
anova(mod.nb_pqi08_1, mod.nb_pqi08_2)

## Likelihood ratio tests of Negative Binomial Models
##
## Response: pqi08
##
## 1                                     expansion_status + percent_male + percent_afric
## 2 expansion_status + percent_male + percent_african_american + percent_hispanic + age + percent_elig
##      theta Resid. df    2 x log-lik.  Test    df LR stat.    Pr(Chi)
## 1 12.14054    10940    -158937.9
## 2 12.16675    10937    -158913.8 1 vs 2      3 24.16406 2.30854e-05
```

By the LRT, we add the interaction term to the model.

```
AIC(mod.nb_pqi08)
```

Model Diagnostics

```
## [1] 174769.5
```

```
AIC(mod.nb_pqi08_c1)
```

```
## [1] 174343.6
```

```
AIC(mod.nb_pqi08_c2)
```

```
## [1] 173771.3
```

```
AIC(mod.nb_pqi08_c3)
```

```
## [1] 173701.6
```

```
AIC(mod.nb_pqi08_c4)
```

```
## [1] 159312.2
```

```
AIC(mod.nb_pqi08_c5)
```

```
## [1] 159047.9
```

```
AIC(mod.nb_pqi08_1)
```

```
## [1] 158961.9
```

```
AIC(mod.nb_pqi08_2)
```

```
## [1] 158943.8
```

```
BIC(mod.nb_pqi08)
```

```
## [1] 174806
```

```
BIC(mod.nb_pqi08_c1)
```

```
## [1] 174387.4
```



```
BIC(mod.nb_pqi08_c2)
```

```
## [1] 173822.4
```

```
BIC(mod.nb_pqi08_c3)
```

```
## [1] 173760
```

```
BIC(mod.nb_pqi08_c4)
```

```
## [1] 159385.2
```

```
BIC(mod.nb_pqi08_c5)
```

```
## [1] 159128.2
```

```
BIC(mod.nb_pqi08_1)
```

```
## [1] 159049.6
```

```
BIC(mod.nb_pqi08_2)
```

```
## [1] 159053.3
```

The full model has the lowest AIC and BIC of all the models.

```
mod.pois_pqi08 <- glm(pqi08 ~ `expansion_status` + `percent_male` + `percent_african_american` + percent
```

```
# Evaluate overdispersion
```

```
deviance(mod.pois_pqi08)/mod.pois_pqi08$df.residual # 98.75487
```

Check Poisson model

```
## [1] 98.75487
```

```
pearson.stat_pqi08 <- sum((na.omit(pqi08_data)$pqi08 - fitted(mod.pois_pqi08))^2/fitted(mod.pois_pqi08))
```

```
pearson.stat_pqi08/mod.pois_pqi08$df.residual # 99.92519
```

```
## [1] 99.92519
```

```
# Evidence of overdispersion
```

```
mean(pqi08_data$pqi08, na.rm = T) # 1396.131
```

```
## [1] 1396.131
```

```
var(pqi08_data$pqi08, na.rm = T) # 669616.2
```

```
## [1] 669616.2
```

Due to the evidence of overdispersion, we can say the negative binomial is the right choice.

```
exp(coef(mod.nb_pqi08_2))
```

Interpretation

```
## (Intercept)
## 1.88564196
## expansion_status2
## 1.20515810
## expansion_status3
```

```
## 1.49873628
## expansion_status4
## 1.23478234
## percent_male
## 0.01118054
## percent_african_american
## 5.96804695
## percent_hispanic
## 1.64101007
## age65-74
## 0.76000999
## age75+
## 2.26975614
## percent_eligible_for_medicaid
## 12.53055363
## percent_eligible_for_medicaid_2
## 0.03860261
## expansion_status2:percent_eligible_for_medicaid
## 0.95401454
## expansion_status3:percent_eligible_for_medicaid
## 0.52538216
## expansion_status4:percent_eligible_for_medicaid
## 0.77061535
```

```
exp(confint(mod.nb_pqi08_2))
```

```
## Waiting for profiling to be done...
```

	2.5 %	97.5 %
## (Intercept)	1.586506347	2.24131789
## expansion_status2	1.141200851	1.27259059
## expansion_status3	1.401515732	1.60266133
## expansion_status4	1.172875047	1.29989803
## percent_male	0.007652943	0.01633597
## percent_african_american	5.524897110	6.44739466
## percent_hispanic	1.515213998	1.77759803
## age65-74	0.749967883	0.77018640
## age75+	2.239792476	2.30012017
## percent_eligible_for_medicaid	8.455761108	18.54796071
## percent_eligible_for_medicaid_2	0.019024975	0.07847272
## expansion_status2:percent_eligible_for_medicaid	0.754217239	1.20688146
## expansion_status3:percent_eligible_for_medicaid	0.392397869	0.70357760
## expansion_status4:percent_eligible_for_medicaid	0.623064550	0.95284009