

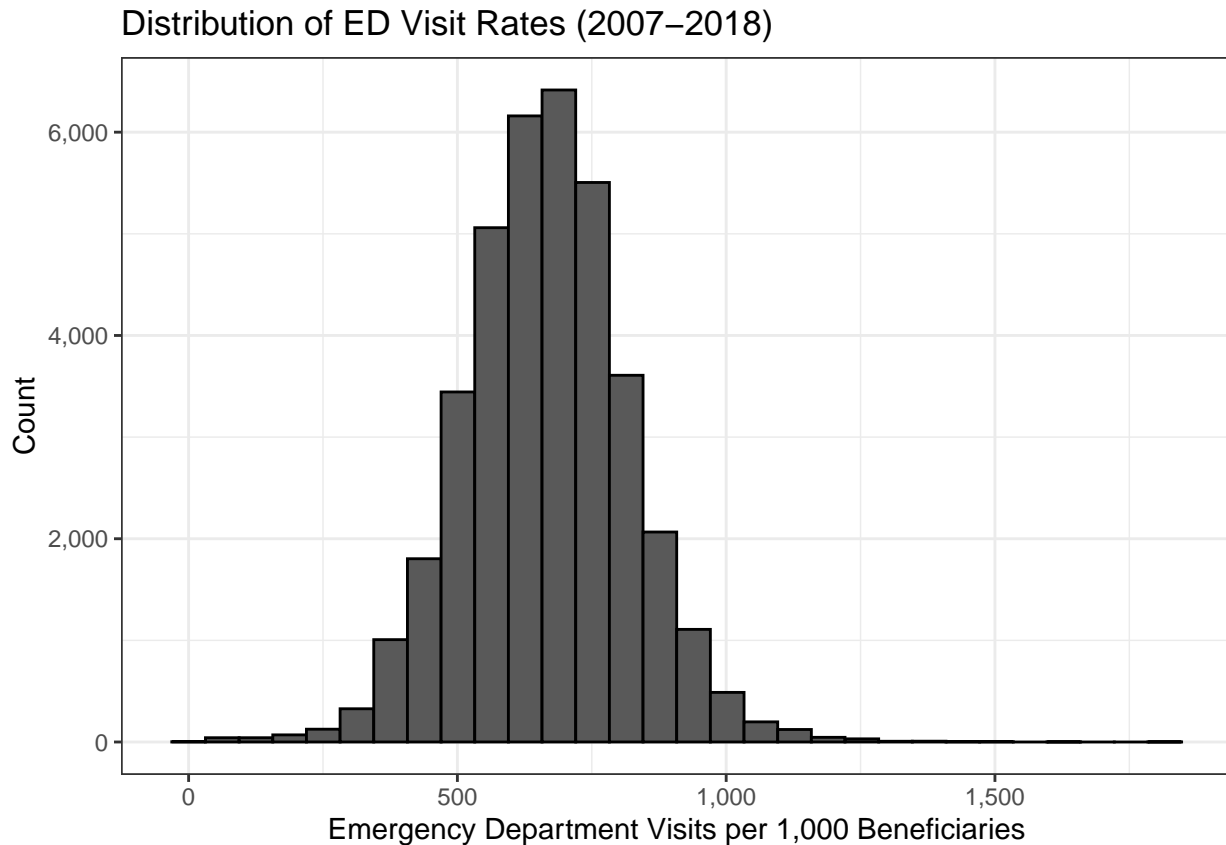
Section 4: Emergency Department Admissions Negative Binomial Model

We built a model to model the relationship between the emergency department visits per 1,000 beneficiaries 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).

Check distribution of outcome

```
medicare %>%
  ggplot(aes(x= emergency_department_visits_per_1000_beneficiaries)) +
  geom_histogram(color = "black") +
  scale_x_continuous(labels = scales::comma) +
  scale_y_continuous(labels = scales::comma) +
  xlab("Emergency Department Visits per 1,000 Beneficiaries") +
  ylab("Count") +
  ggtitle("Distribution of ED Visit Rates (2007-2018)") +
  theme_bw()

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
## Warning: Removed 41 rows containing non-finite values (stat_bin).
```



The outcome looks pretty normally distributed.

We next check for missingness in the outcome:

```
medicare %>%
  select(emergency_department_visits_per_1000_beneficiaries) %>%
  summarize_all(.funs = function(x)
    return(paste0(round(sum(is.na(x))/length(x), 3)*100, "%")) %>%
  t() %>% data.frame()
```

```
##
## emergency_department_visits_per_1000_beneficiaries 0.1%
```

Model Fitting

```
mod.nb_ed <- MASS::glm.nb(emergency_department_visits_per_1000_beneficiaries ~ expansion_status + offset(log(year)), data = medicare, init.theta = 17.76803829,
summary(mod.nb_ed)
```

Starting with our base model

```
##
## Call:
## MASS::glm.nb(formula = emergency_department_visits_per_1000_beneficiaries ~
##   expansion_status + offset(log(year)), data = medicare, init.theta = 17.76803829,
##   link = log)
##
## Deviance Residuals:
##   Min       1Q   Median       3Q      Max
## -8.3572  -0.6428   0.0015   0.5731   5.0732
```

```
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -1.173851   0.004759  -246.66  <2e-16 ***
## expansion_status2  0.074556   0.005301   14.06  <2e-16 ***
## expansion_status3  0.062586   0.005892   10.62  <2e-16 ***
## expansion_status4  0.071896   0.005058   14.21  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(17.768) family taken to be 1)
##
##      Null deviance: 38387  on 37646  degrees of freedom
## Residual deviance: 38173  on 37643  degrees of freedom
## (89 observations deleted due to missingness)
## AIC: 487645
##
## Number of Fisher Scoring iterations: 1
##
##
##              Theta: 17.768
##              Std. Err.: 0.132
##
## 2 x log-likelihood: -487634.657
# Model diagnostics table
model_diagnostics <- data.frame(
  Model = "ED Visits ~ Exp",
  AIC = AIC(mod.nb_ed),
  BIC = BIC(mod.nb_ed)
)
```

Improve Model Fit Using Other Covariates The covariates percent_male, percent_eligible_for_medicaid, average_age, do not satisfy the criteria for confounding the relationship between Medicaid Expansion status and Emergency Department Visit rates, so we will add each of these covariates one by one to determine if they improve the model fit using Likelihood Ratio tests.

Percent Male

```
mod.nb_ed_c1 <- MASS::glm.nb(emergency_department_visits_per_1000_beneficiaries ~ expansion_status + percent_male,
  summary(mod.nb_ed_c1)

##
## Call:
## MASS::glm.nb(formula = emergency_department_visits_per_1000_beneficiaries ~
##   expansion_status + percent_male + offset(log(year)), data = medicare,
##   init.theta = 18.03241502, link = log)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -8.3536  -0.6489  -0.0126   0.5673   4.7884
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -0.648440   0.023213  -27.935  <2e-16 ***
```

```
## expansion_status2 0.071795 0.005264 13.639 <2e-16 ***
## expansion_status3 0.056856 0.005850 9.719 <2e-16 ***
## expansion_status4 0.059482 0.005043 11.795 <2e-16 ***
## percent_male -1.129669 0.048963 -23.072 <2e-16 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(18.0324) family taken to be 1)
##
## Null deviance: 38884 on 37639 degrees of freedom
## Residual deviance: 38166 on 37635 degrees of freedom
## (96 observations deleted due to missingness)
## AIC: 487002
##
## Number of Fisher Scoring iterations: 1
##
##
## Theta: 18.032
## Std. Err.: 0.135
##
## 2 x log-likelihood: -486990.168
# Use Likelihood Ratio test to determine if confounder is necessary
anova(mod.nb_ed, mod.nb_ed_c1) # Yes

## Likelihood ratio tests of Negative Binomial Models
##
## Response: emergency_department_visits_per_1000_beneficiaries
##
## Model theta Resid. df
## 1 expansion_status + offset(log(year)) 17.76804 37643
## 2 expansion_status + percent_male + offset(log(year)) 18.03242 37635
## 2 x log-lik. Test df LR stat. Pr(Chi)
## 1 -487634.7
## 2 -486990.2 1 vs 2 8 644.4886 0

# # Model diagnostics table
model_diagnostics <- model_diagnostics %>%
  bind_rows(
    data.frame(
      Model = "ED Visits ~ Exp + Male",
      AIC = AIC(mod.nb_ed_c1),
      BIC = BIC(mod.nb_ed_c1)
    )
  )
)
```

By the LRT, we do add in the effects of percent_male.

Percent Eligible for Medicaid

```
mod.nb_ed_c2 <- MASS::glm.nb(emergency_department_visits_per_1000_beneficiaries ~ expansion_status + pe
summary(mod.nb_ed_c2)

##
## Call:
## MASS::glm.nb(formula = emergency_department_visits_per_1000_beneficiaries ~
## expansion_status + percent_male + percent_eligible_for_medicaid +
## offset(log(year)), data = medicare, init.theta = 25.64189931,
```

```

##      link = log)
##
## Deviance Residuals:
##      Min        1Q      Median        3Q        Max
## -11.3153    -0.5552     0.0145     0.5466     4.4794
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -0.781338   0.019806  -39.449 < 2e-16 ***
## expansion_status2    0.051741   0.004444   11.642 < 2e-16 ***
## expansion_status3    0.040550   0.004940    8.209 2.23e-16 ***
## expansion_status4    0.034023   0.004262    7.983 1.43e-15 ***
## percent_male    -1.484589   0.041908  -35.425 < 2e-16 ***
## percent_eligible_for_medicaid  1.434229   0.011642  123.192 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(25.6419) family taken to be 1)
##
##      Null deviance: 53305  on 37488  degrees of freedom
## Residual deviance: 38048  on 37483  degrees of freedom
##      (247 observations deleted due to missingness)
## AIC: 472391
##
## Number of Fisher Scoring iterations: 1
##
##
##              Theta: 25.642
##              Std. Err.: 0.195
##
##      2 x log-likelihood:  -472377.186
## Use Anova to determine if confounder is necessary
anova(mod.nb_ed_c1, mod.nb_ed_c2) # Yes

## Likelihood ratio tests of Negative Binomial Models
##
## Response: emergency_department_visits_per_1000_beneficiaries
##
##      1                                expansion_status + percent_male + offset(log(year))
##      2 expansion_status + percent_male + percent_eligible_for_medicaid + offset(log(year))
##      theta Resid. df      2 x log-lik.  Test      df LR stat. Pr(Chi)
##      1 18.03242    37635      -486990.2
##      2 25.64190    37483      -472377.2 1 vs 2    152 14612.98      0

## # Model diagnostics table
model_diagnostics <- model_diagnostics %>%
  bind_rows(
    data.frame(
      Model = "ED Visits ~ Exp + Male + % Elig",
      AIC = AIC(mod.nb_ed_c2),
      BIC = BIC(mod.nb_ed_c2)
    )
  )

```

By the LRT, we do add in the effects of percent_eligible_for_medicaid.

Average Age

```
mod.nb_ed_c3 <- MASS::glm.nb(emergency_department_visits_per_1000_beneficiaries ~ expansion_status + percent_male + percent_eligible_for_medicaid + average_age + offset(log(year)), data = medicare, init.theta = 29.66128298, link = log)
```

```
##
## Call:
## MASS::glm.nb(formula = emergency_department_visits_per_1000_beneficiaries ~
##   expansion_status + percent_male + percent_eligible_for_medicaid +
##   average_age + offset(log(year)), data = medicare, init.theta = 29.66128298,
##   link = log)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -11.0876   -0.6045   -0.0250    0.5399    5.6909
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      3.7969141   0.0619559  61.284 < 2e-16 ***
## expansion_status2  0.0347753   0.0041516   8.376 < 2e-16 ***
## expansion_status3  0.0213619   0.0046124   4.631 3.63e-06 ***
## expansion_status4  0.0170959   0.0039807   4.295 1.75e-05 ***
## percent_male     -2.5831239   0.0414482 -62.322 < 2e-16 ***
## percent_eligible_for_medicaid 0.6329201   0.0147607  42.879 < 2e-16 ***
## average_age      -0.0545008   0.0007074 -77.046 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(29.6613) family taken to be 1)
##
##      Null deviance: 61231  on 37488  degrees of freedom
## Residual deviance: 38123  on 37482  degrees of freedom
## (247 observations deleted due to missingness)
## AIC: 467206
##
## Number of Fisher Scoring iterations: 1
##
##              Theta: 29.661
##              Std. Err.: 0.228
##
## 2 x log-likelihood: -467190.146
```

```
# Use Likelihood ratio test to determine if covariate is necessary
anova(mod.nb_ed_c2, mod.nb_ed_c3) # Yes
```

```
## Likelihood ratio tests of Negative Binomial Models
##
## Response: emergency_department_visits_per_1000_beneficiaries
##
## 1          expansion_status + percent_male + percent_eligible_for_medicaid + offset(log(year))
## 2 expansion_status + percent_male + percent_eligible_for_medicaid + average_age + offset(log(year))
##      theta Resid. df    2 x log-lik.  Test      df LR stat. Pr(Chi)
## 1 25.64190    37483    -472377.2
## 2 29.66128    37482    -467190.1 1 vs 2      1 5187.04      0
```

```
## Model diagnostics table
model_diagnostics <- model_diagnostics %>%
  bind_rows(
    data.frame(
      Model = "ED Visits ~ Exp + Male + % Elig + Age",
      AIC = AIC(mod.nb_ed_c3),
      BIC = BIC(mod.nb_ed_c3)
    )
  )
)
```

By the LRT, we do add in the effects of average_age.

```
mod.nb_ed_1 <- MASS::glm.nb(emergency_department_visits_per_1000_beneficiaries ~ expansion_status + percent_eligible_for_medicaid + average_age)
summary(mod.nb_ed_1)
```

Check for quadratic effects of percent_eligible_for_medicaid

```
##
## Call:
## MASS::glm.nb(formula = emergency_department_visits_per_1000_beneficiaries ~
##   expansion_status + percent_male + percent_eligible_for_medicaid +
##   average_age + percent_eligible_for_medicaid_2 + offset(log(year)),
##   data = medicare %>% mutate(percent_eligible_for_medicaid_2 = percent_eligible_for_medicaid^2),
##   init.theta = 31.38808918, link = log)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -9.9333  -0.6014  -0.0234   0.5509   5.8131
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)      2.7416814   0.0651791  42.064 < 2e-16 ***
## expansion_status2  0.0383764   0.0040422   9.494 < 2e-16 ***
## expansion_status3  0.0206499   0.0044900   4.599 4.24e-06 ***
## expansion_status4  0.0240737   0.0038784   6.207 5.40e-10 ***
## percent_male      -2.2597600   0.0411092 -54.970 < 2e-16 ***
## percent_eligible_for_medicaid  2.6954272   0.0469964  57.354 < 2e-16 ***
## average_age       -0.0452695   0.0007228 -62.627 < 2e-16 ***
## percent_eligible_for_medicaid_2 -3.7076762   0.0794541 -46.664 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(31.3881) family taken to be 1)
##
## Null deviance: 64603  on 37488  degrees of freedom
## Residual deviance: 38150  on 37481  degrees of freedom
## (247 observations deleted due to missingness)
## AIC: 465202
##
## Number of Fisher Scoring iterations: 1
##
##
## Theta: 31.388
```

```
##          Std. Err.:  0.242
##
## 2 x log-likelihood:  -465183.777
anova(mod.nb_ed_c3, mod.nb_ed_1) # Yes

## Likelihood ratio tests of Negative Binomial Models
##
## Response: emergency_department_visits_per_1000_beneficiaries
##
## 1                                expansion_status + percent_male + percent_eligible_for_medicaid +
## 2 expansion_status + percent_male + percent_eligible_for_medicaid + average_age + percent_eligible_f
##      theta Resid. df    2 x log-lik.  Test    df LR stat. Pr(Chi)
## 1 29.66128    37482    -467190.1
## 2 31.38809    37481    -465183.8 1 vs 2      1 2006.369      0

model_diagnostics <- model_diagnostics %>%
  bind_rows(
    data.frame(
      Model = "ED Visits ~ Exp + Male + % Elig + Age + % Elig^2",
      AIC = AIC(mod.nb_ed_1),
      BIC = BIC(mod.nb_ed_1)
    )
  )
)
```

By the LRT, we do add in the quadratic effects of 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_ed_2 <- MASS::glm.nb(emergency_department_visits_per_1000_beneficiaries ~ expansion_status + per
summary(mod.nb_ed_2)
```

```
##
## Call:
## MASS::glm.nb(formula = emergency_department_visits_per_1000_beneficiaries ~
##   expansion_status + percent_male + percent_eligible_for_medicaid +
##   average_age + percent_eligible_for_medicaid_2 + percent_eligible_for_medicaid *
##   expansion_status + offset(log(year)), data = medicare %>%
##   mutate(percent_eligible_for_medicaid_2 = percent_eligible_for_medicaid^2),
##   init.theta = 31.51175527, link = log)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -9.6219  -0.6080  -0.0241   0.5535   5.7001
##
## Coefficients:
##                                Estimate Std. Error z value
## (Intercept)                   2.7444137  0.0653659  41.985
## expansion_status2              0.0136760  0.0110109   1.242
## expansion_status3              0.0583827  0.0123470   4.728
## expansion_status4             -0.0254955  0.0104870  -2.431
## percent_male                  -2.2289477  0.0412870 -53.987
## percent_eligible_for_medicaid  2.5747088  0.0621452  41.431
## average_age                   -0.0451236  0.0007259 -62.158
```



```
## percent_eligible_for_medicaid_2          -3.7661832  0.0797081 -47.250
## expansion_status2:percent_eligible_for_medicaid  0.1231502  0.0496178   2.482
## expansion_status3:percent_eligible_for_medicaid -0.1663355  0.0551050  -3.019
## expansion_status4:percent_eligible_for_medicaid  0.2381265  0.0474591   5.018
##                                           Pr(>|z|)
## (Intercept)                                < 2e-16 ***
## expansion_status2                          0.21422
## expansion_status3                          2.26e-06 ***
## expansion_status4                          0.01505 *
## percent_male                              < 2e-16 ***
## percent_eligible_for_medicaid             < 2e-16 ***
## average_age                               < 2e-16 ***
## percent_eligible_for_medicaid_2           < 2e-16 ***
## expansion_status2:percent_eligible_for_medicaid  0.01307 *
## expansion_status3:percent_eligible_for_medicaid  0.00254 **
## expansion_status4:percent_eligible_for_medicaid  5.23e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for Negative Binomial(31.5118) family taken to be 1)
##
##      Null deviance: 64843  on 37488  degrees of freedom
## Residual deviance: 38149  on 37478  degrees of freedom
## (247 observations deleted due to missingness)
## AIC: 465065
##
## Number of Fisher Scoring iterations: 1
##
##
##              Theta:  31.512
##             Std. Err.:  0.243
##
## 2 x log-likelihood:  -465040.689
```

```
anova(mod.nb_ed_1, mod.nb_ed_2)
```

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

```
##
```

```
## Response: emergency_department_visits_per_1000_beneficiaries
```

```
##
```

```
## 1                                     expansion_status + percent_male + percent_eligible_for_medicaid
## 2 expansion_status + percent_male + percent_eligible_for_medicaid + average_age + percent_eligible_for_medicaid^2
##      theta Resid. df    2 x log-lik.   Test      df LR stat. Pr(Chi)
## 1 31.38809    37481      -465183.8
## 2 31.51176    37478      -465040.7 1 vs 2      3 143.0884      0
```

```
model_diagnostics <- model_diagnostics %>%
```

```
  bind_rows(
```

```
    data.frame(
```

```
      Model = "ED Visits ~ Exp + Male + % Elig + Age + % Elig^2 + % Elig*Exp",
```

```
      AIC = AIC(mod.nb_ed_2),
```

```
      BIC = BIC(mod.nb_ed_2)
```

```
    )
```

```
)
```

By the LRT, we do add the interaction between Medicaid expansion status and percent_eligible_for_medicaid.

```
model_diagnostics
```

Model Diagnostics

```
##                                Model      AIC
## 1                        ED Visits ~ Exp 487644.7
## 2                        ED Visits ~ Exp + Male 487002.2
## 3                        ED Visits ~ Exp + Male + % Elig 472391.2
## 4                        ED Visits ~ Exp + Male + % Elig + Age 467206.1
## 5                        ED Visits ~ Exp + Male + % Elig + Age + % Elig^2 465201.8
## 6 ED Visits ~ Exp + Male + % Elig + Age + % Elig^2 + % Elig*Exp 465064.7
##      BIC
## 1 487687.3
## 2 487053.4
## 3 472450.9
## 4 467274.4
## 5 465278.6
## 6 465167.1
```

We have the smallest AIC and BIC for the full model.

```
mod.pois_ed <- glm(emergency_department_visits_per_1000_beneficiaries ~
  expansion_status + percent_male + percent_eligible_for_medicaid +
  average_age + percent_eligible_for_medicaid_2 + percent_eligible_for_medicaid *
  expansion_status + offset(log(year)), data = medicare %>% mutate(percent_eligible_for_medicaid_2 =
    percent_eligible_for_medicaid^2))

# Evaluate overdispersion
deviance(mod.pois_ed)/mod.pois_ed$df.residual # 19.99481
```

Check Poisson model

```
## [1] 19.99481

pearson.stat_ed <- sum((na.omit(medicare %>% select(percent_male, percent_eligible_for_medicaid, average_age))$pearson_resid)^2)
pearson.stat_ed/mod.pois_ed$df.residual # 19.4187

## [1] 19.4187

# Evidence of overdispersion
mean(medicare$emergency_department_visits_per_1000_beneficiaries, na.rm = T) # 664.9697

## [1] 664.9697

var(medicare$emergency_department_visits_per_1000_beneficiaries, na.rm = T) # 23235.82

## [1] 23235.82
```

Since the ratio of the Pearson χ^2 statistic and deviance to their respective degrees of freedom are greater than 1, the negative binomial model seems appropriate.

Interpretation Output which is interpreted in the paper.

```
exp(coef(mod.nb_ed_2))

##                                (Intercept)
##                                15.55549171
##                                expansion_status2
```

```
##                1.01376991
##                expansion_status3
##                1.06012059
##                expansion_status4
##                0.97482680
##                percent_male
##                0.10764164
##                percent_eligible_for_medicaid
##                13.12749383
##                average_age
##                0.95587935
##                percent_eligible_for_medicaid_2
##                0.02314022
## expansion_status2:percent_eligible_for_medicaid
##                1.13105433
## expansion_status3:percent_eligible_for_medicaid
##                0.84676212
## expansion_status4:percent_eligible_for_medicaid
##                1.26886969
```

```
exp(confint(mod.nb_ed_2))
```

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

	2.5 %	97.5 %
## (Intercept)	13.66168322	17.71150246
## expansion_status2	0.99224216	1.03576400
## expansion_status3	1.03453004	1.08634513
## expansion_status4	0.95512236	0.99493825
## percent_male	0.09909400	0.11692989
## percent_eligible_for_medicaid	11.61599511	14.83545949
## average_age	0.95450185	0.95725899
## percent_eligible_for_medicaid_2	0.01978178	0.02707666
## expansion_status2:percent_eligible_for_medicaid	1.02676325	1.24571669
## expansion_status3:percent_eligible_for_medicaid	0.75918263	0.94433981
## expansion_status4:percent_eligible_for_medicaid	1.15682391	1.39147343