technische universität dortmund

ADVANCED BAYESIAN DATA ANALYSIS

BAYESIAN DATA ANALYSIS ON SUICIDES IN INDIA

TU DORTMUND | DATA SCIENCE

Group Members

Aakash Goyal (229975)

Jaykumar Savani (230443)

Supervisors

Prof. Dr. Paul Bürkner

Prof. Dr. Katja Ickstadt

Table of contents

Research Objective

How the number of suicides vary by gender, age group, cause and each state in India?

Specific Objectives:

- Investigate the relationship between different age groups and the number of suicides in India
- Analyze the differences in number of suicides between males and females in India
- Identify the suicide methods that pose the highest risk of suicide in India

Data Description

- **Total observations :** 2,37,520
- **Total columns** : 7
- **Data source:** https://data.world/rajanand/suicides-in-india

Column Name	Data Type	Description
State	Categorical	29 Indian states and 7 Union territories
Year	Numeric	From 2001 to 2012
Type_Code	Categorical	5 Category of Suicide
Type	Categorical	69 different reasons of suicide
Gender	Categorical	Male or Female
Age_group	Categorical	6 Groups: 0-14, 15-29, 30-44, 45-60, 60+, 0-100+(Total)
Total	Numeric	Total number of suicides

Table 1: Dataset description

Data Processing

- Selected one year i.e. 2012 (most recent)
- Remove age group 0-100+ as it contains total sum of all suicides for the other age groups
- Remove unwanted special characters present in Type column
- Map down 69 categories of suicide causes into 9 different categories

Type	Cause
"By_Consuming_Insecticides","Consuming_Insecticides", "By_Consuming_Other_Poison", "By_Over_Alcoholism", "By_Overdose_of_sleeping_pills", "Drug_Abuse_Addiction"	Drugs
"Ideological_Causes_Hero_Worshipping", "Love_Affairs", "Fall_in_Social_Reputation", "Physical_Abuse_Rape_Incest_Etc_"	Social
"Property_Dispute", "Family_Problems", "Cancellation_Non_Settlement_of_Marriage", "Married", "Never_Married", "Divorcee","Seperated","Dowry_Dispute", "Divorce", "Widowed_Widower", "Death_of_Dear_Person","Suspected_Illicit_Relation"	Family
"Insanity_Mental_Illness", "Illness_Aids_STD", "Not_having_Children_Barrenness_Impotency", "Other_Prolonged_Illness", "By_Self_Infliction_of_injury", "Illegitimate_Pregnancy", "Cancer", "Paralysis", "Insanity_Mental_Illness"	Health
"No_Education", "Student", "Hr__Secondary_Intermediate_Pre_Universit", "Failure_in_Examination", "Post_Graduate_and_Above", "Middle", "Primary", "Diploma", "Matriculate_Secondary", "Unemployed", "Graduate"	Educational
"Self_employed_Business_activity", "ServicePrivate", "Poverty", "Professional_Career_Problem", "Retired_Person", "House_Wife", "Bankruptcy_or_Sudden_change_in_Economic", "Unemployment", "Service_Government","Public_Sector_Undertaking","Self_employed_Business_activity"	Financial & Career
"By_Drowning", "By_touching_electric_wires", "Professional_Activity", "Farming_Agriculture_Activity","By_coming_under_running_vehicles_trains","By_Machine"	Accidental
"By_Hanging", "By_Jumping_from_Building", "By_Fire_Arms", "By_Fire_Self_Immolation","By_Jumping_from_Other_sites", "By_Jumping_off_Moving_Vehicles_Trains"	Intentional
Others_Please_Specify", "By_Other_means_please_specify","Other_Causes_Please_Specity", "Causes_Not_known", "Others_Please_Specify"	Others

Table 2: Mapping of all Type to Cause Categories

Cleaned dataset

	State	Year	Gender	Age_group	Cause	Total
1	ANDHRA PRADESH	2012	Female	0-14	Accidental	24
2	ANDHRA PRADESH	2012	Female	0-14	Drugs	70
3	ANDHRA PRADESH	2012	Female	0-14	Educational	65
4	ANDHRA PRADESH	2012	Female	0-14	Family	19
5	ANDHRA PRADESH	2012	Female	0-14	Financial & Career	1
6	ANDHRA PRADESH	2012	Female	0-14	Health	24
7	ANDHRA PRADESH	2012	Female	0-14	Intentional	45
8	ANDHRA PRADESH	2012	Female	0-14	Others	158
9	ANDHRA PRADESH	2012	Female	0-14	Social	23
10	ANDHRA PRADESH	2012	Female	15-29	Accidental	332
11	ANDHRA PRADESH	2012	Female	15-29	Drugs	912
12	ANDHRA PRADESH	2012	Female	15-29	Educational	368
13	ANDHRA PRADESH	2012	Female	15-29	Family	667
14	ANDHRA PRADESH	2012	Female	15-29	Financial & Career	1192
15	ANDHRA PRADESH	2012	Female	15-29	Health	672
16	ANDHRA PRADESH	2012	Female	15-29	Intentional	818
17	ANDHRA PRADESH	2012	Female	15-29	Others	1038
18	ANDHRA PRADESH	2012	Female	15-29	Social	139
19	ANDHRA PRADESH	2012	Female	30-44	Accidental	335
20	ANDHRA PRADESH	2012	Female	30-44	Drugs	712

Fig 1.1: Processed/Cleaned dataset

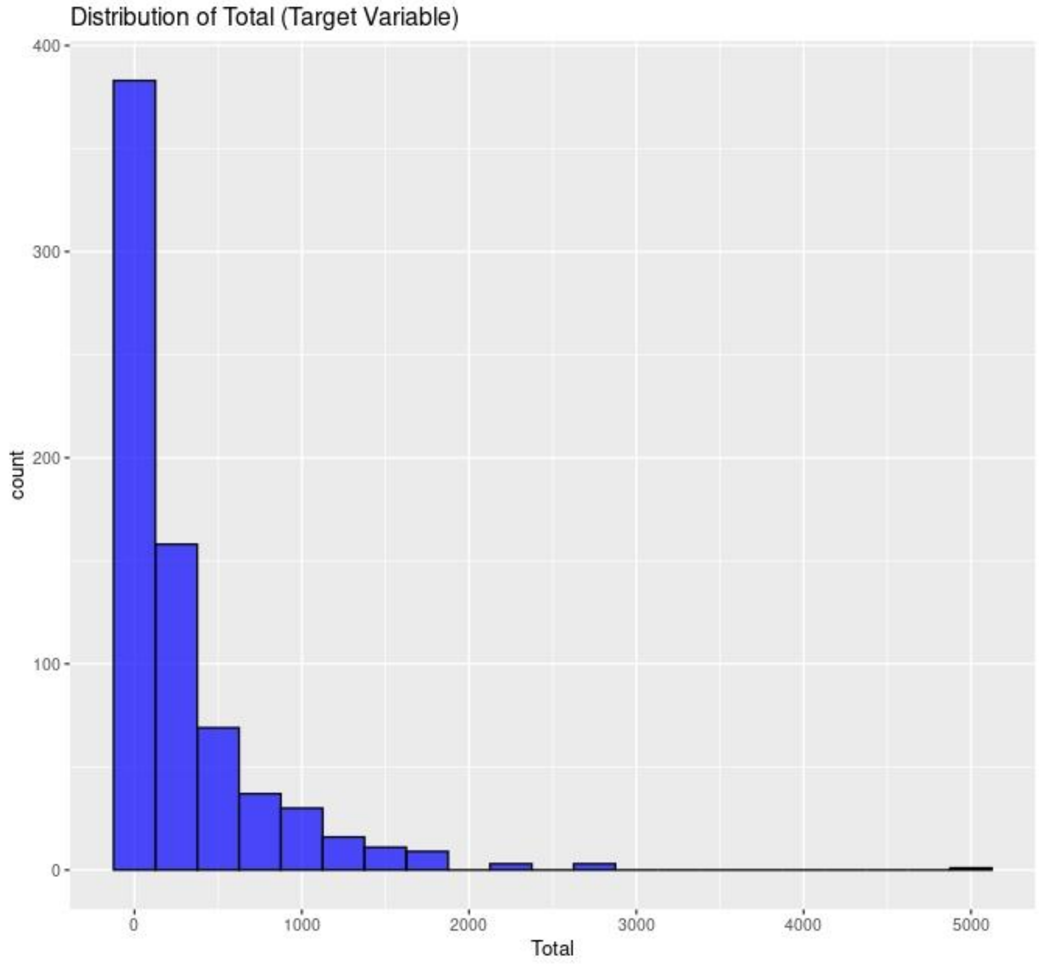


Fig 1.2:Distribution of Total

Models

Which model to choose ?

- Poisson Model
- Negative Binomial
- Zero Inflated Negative Binomial

*Source: brms: An R Package for Bayesian Multilevel Models using Stan By Paul-Christian Bürkner

Poisson | Neg Binomial | Zero inflated NB

technische universität dortmund

Family: Poisson
Links: mu = log ()
Response variable: Total
Group level Term: State, random intercept is for State variable with 8 unique levels.
sd(Intercept): represents standard deviation at group level which measure variability across different states.

Significant covariates:

- All covariates are significant.

ESS: captures how many independent draws contain the same amount of information as the dependent sample obtained by the MCMC algorithm.

- High Bulk_ESS*
- High Tail_ESS

*Bulk_ESS > num of chains * 100
- Source: Runtime warnings and convergence problems by Stan Development Team (link)

```
> summary(poisson_model)
 Family: poisson
  Links: mu = log
Formula: Total ~ Cause + Gender + Age_group + (1 | State)
   Data: subset_df1 (Number of observations: 720)
  Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup draws = 4000

Group-Level Effects:
~State (Number of levels: 8)
              Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sd(Intercept)     0.88      0.29     0.51     1.62 1.00     1015     1573

Population-Level Effects:
                      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
Intercept                 2.34      0.30     1.71     2.93 1.01      828     1037
CauseDrugs                0.46      0.01     0.45     0.48 1.00     2023     2595
CauseEducational         -0.65      0.01    -0.68    -0.63 1.00     2565     2804
CauseFamily               0.11      0.01     0.09     0.13 1.00     2122     2541
CauseFinancial&Career     0.47      0.01     0.45     0.49 1.00     2104     2565
CauseHealth              -0.16      0.01    -0.18    -0.14 1.00     2240     2541
CauseIntentional          0.68      0.01     0.66     0.70 1.00     1836     2426
CauseOthers               0.96      0.01     0.95     0.98 1.00     1840     2257
CauseSocial              -1.87      0.02    -1.91    -1.83 1.00     2833     2674
GenderMale                0.64      0.00     0.63     0.65 1.00     2190     1805
Age_group15M29            3.05      0.02     3.02     3.09 1.00     1824     1865
Age_group30M44            3.04      0.02     3.00     3.07 1.00     1930     2095
Age_group45M59            2.53      0.02     2.50     2.57 1.00     1938     1993
Age_group60P              1.61      0.02     1.57     1.65 1.00     2019     2084

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
>
```

Fig 2(a) : Poisson Model Summary

Poisson | Neg Binomial | Zero inflated NB

Convergence Diagnostics:

- R-hat compares between and within chain estimates of model parameters.
- Rhat for the model parameters are less than 1.01.
- Chains have mixed well and he chains have converged to a common distribution.

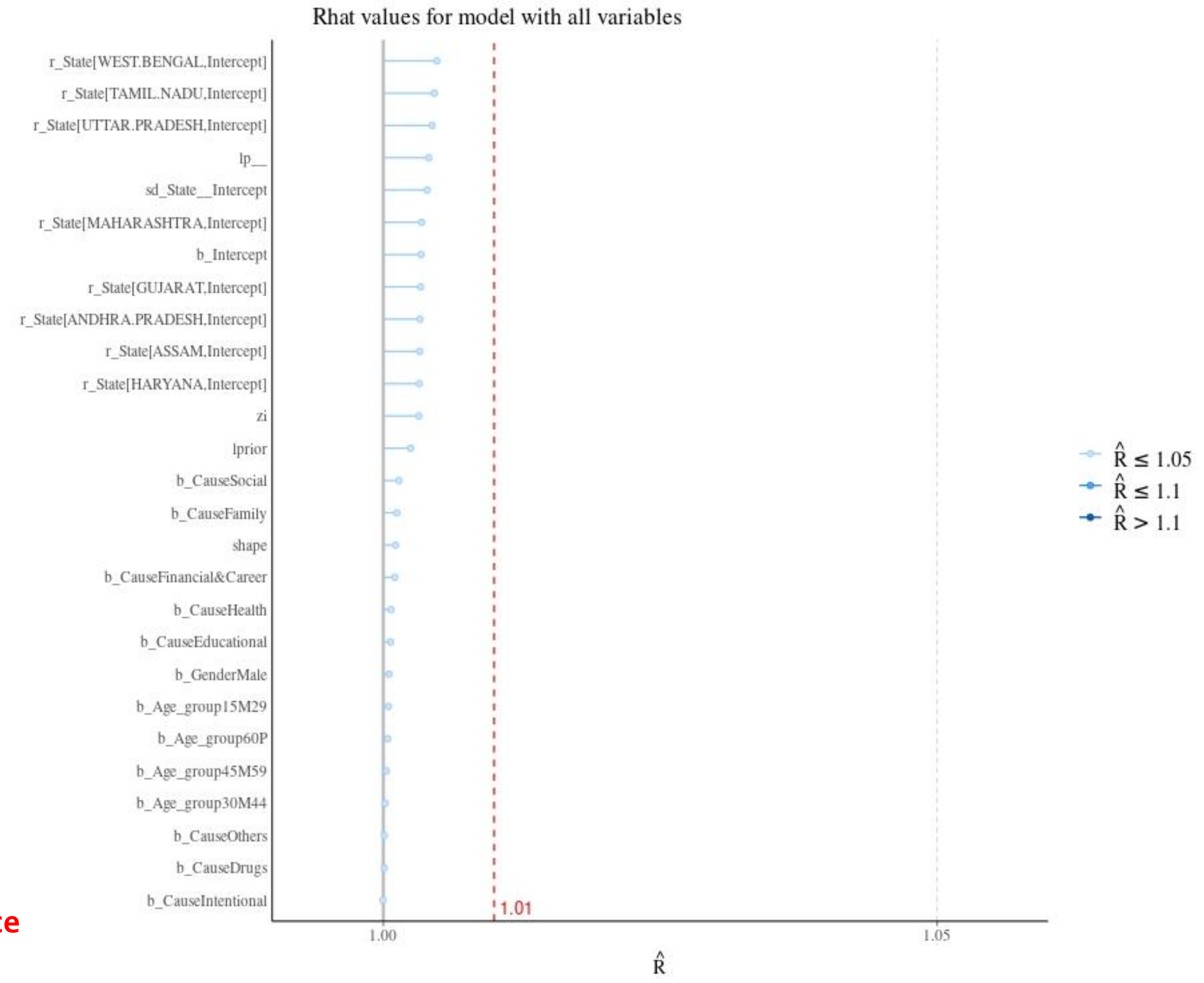


Fig2(b): R-hat values

* R-hat < 1.01
- Source: Runtime warnings and convergence problems by Stan Development (link)

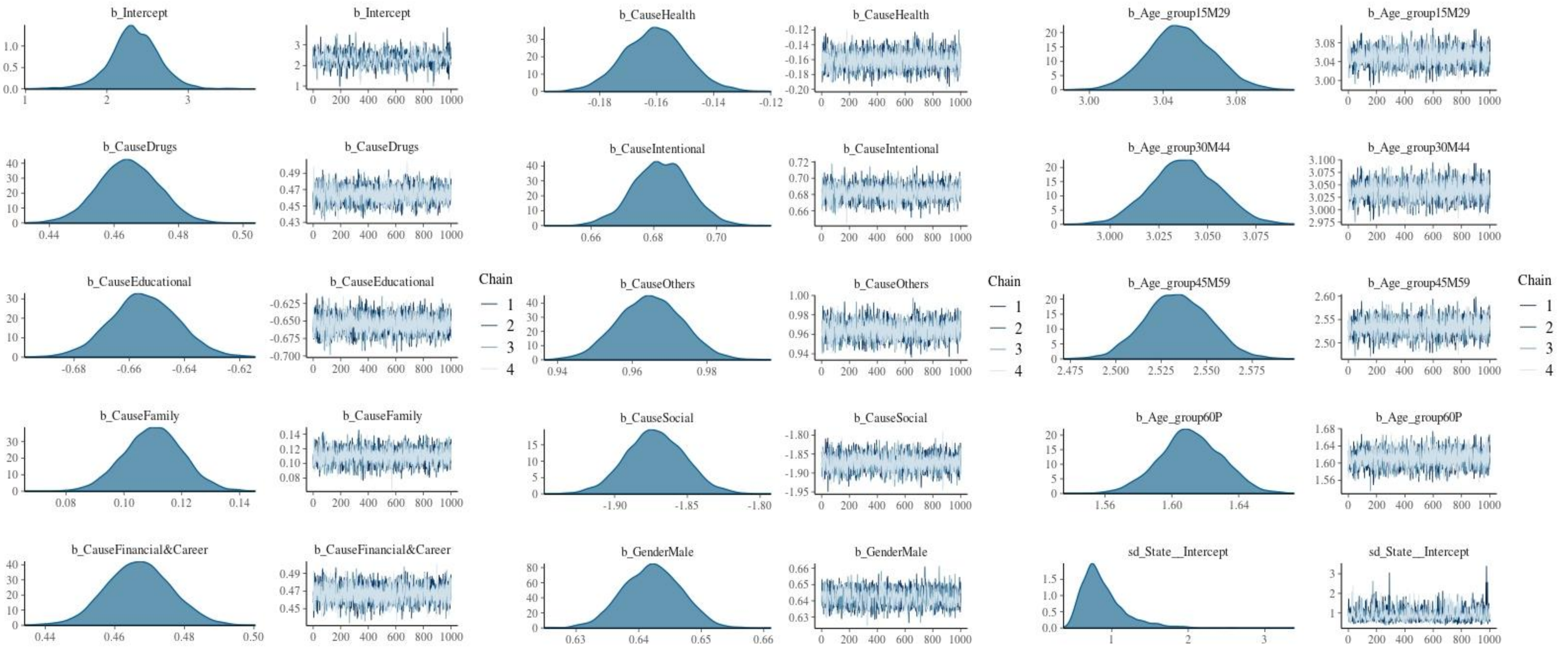
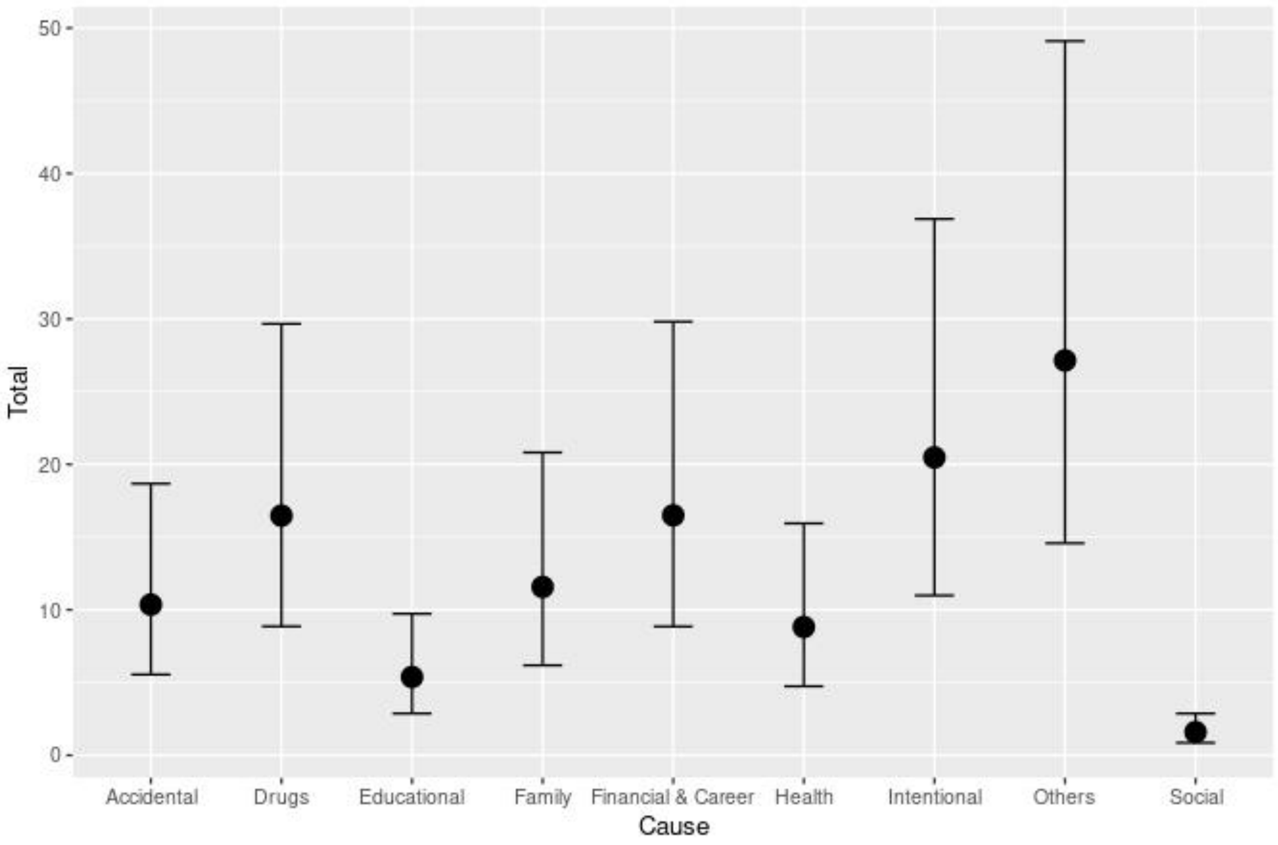


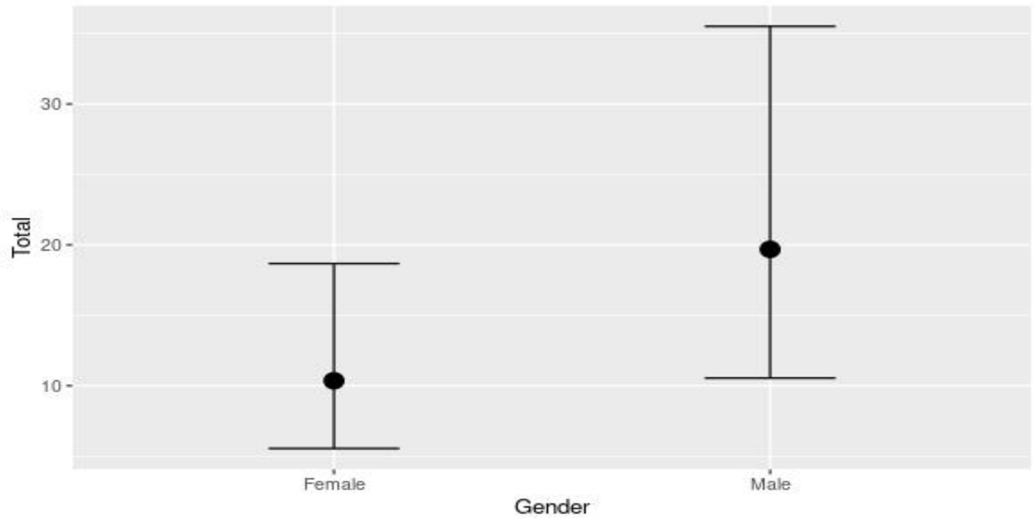
Fig 3: Trace and Density plots of all relevant parameters of the poisson model

- Trace plots shows that model has explored all the possible values it could look at so it converged well.
- Density plot shows the estimate found by model in Summary

Poisson | Neg Binomial | Zero inflated NB

technische universität dortmund





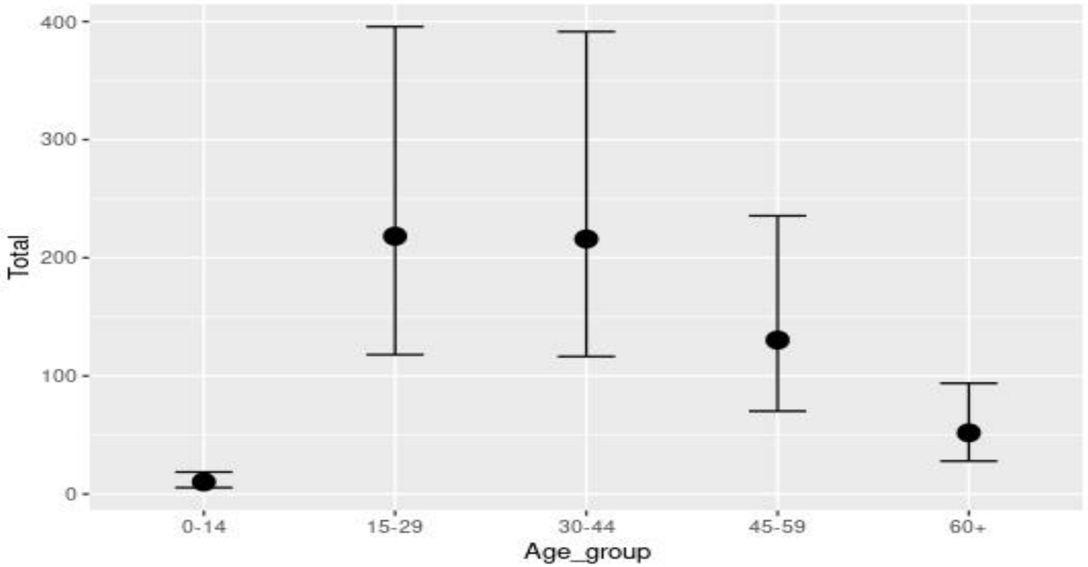


Fig 4: Conditional effects plots of all population-level predictors (fitted)

Poisson | Neg Binomial | Zero inflated NB

technische universität
dortmund

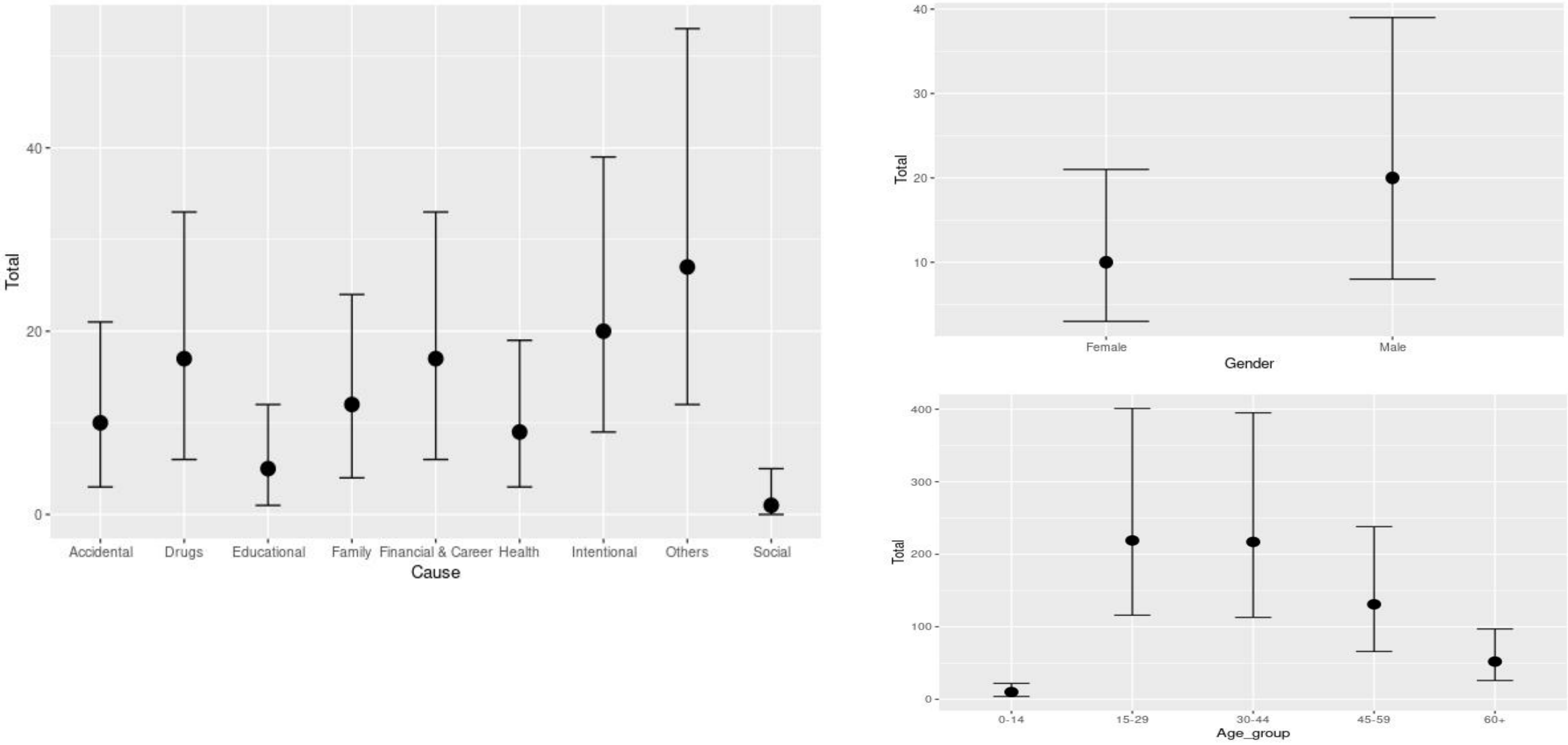
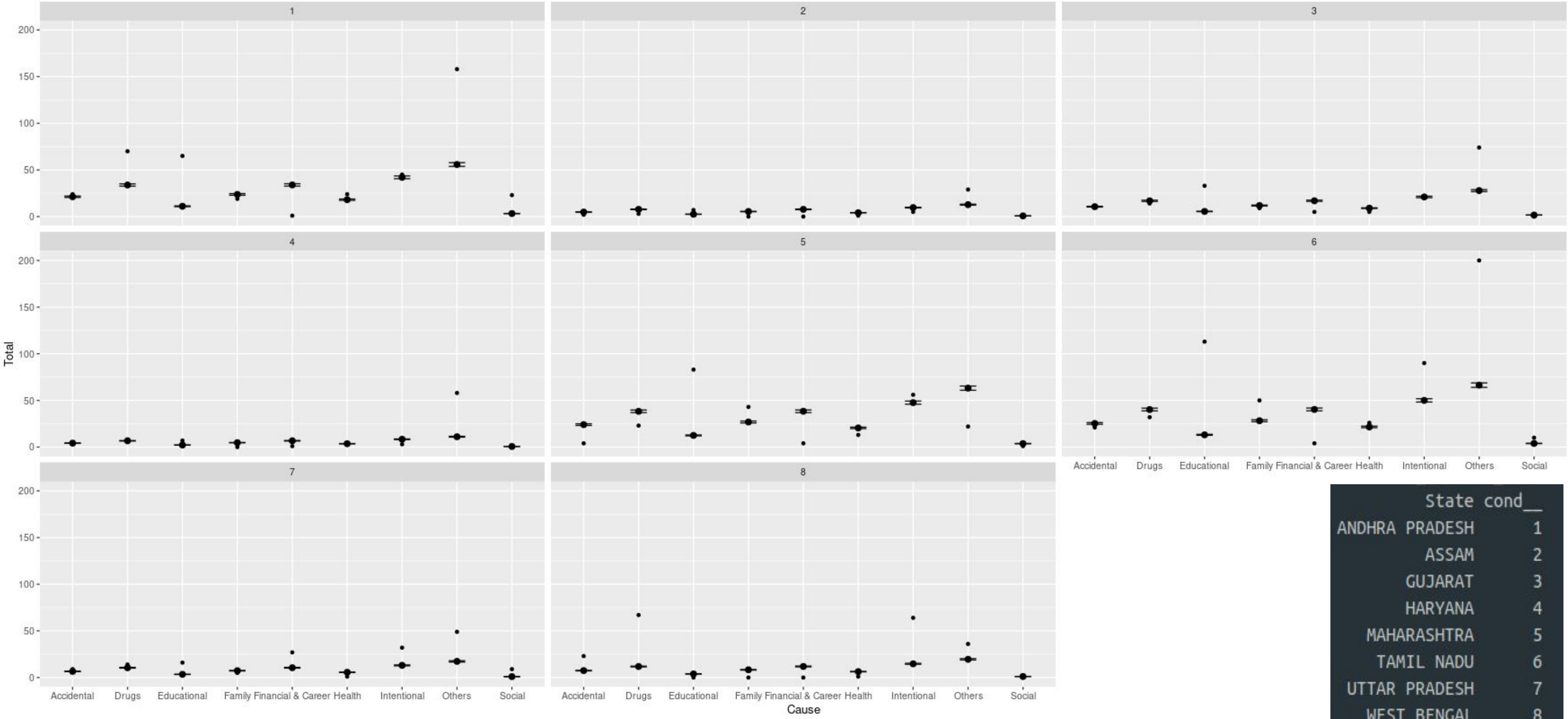


Fig 4(b): Conditional effects plots of all population-level predictors (Predict)

Poisson | Neg Binomial | Zero inflated NB



State	cond__
ANDHRA PRADESH	1
ASSAM	2
GUJARAT	3
HARYANA	4
MAHARASHTRA	5
TAMIL NADU	6
UTTAR PRADESH	7
WEST BENGAL	8

Fig 5(a): Conditional effects plots of all population-level predictors across the states (fitted)

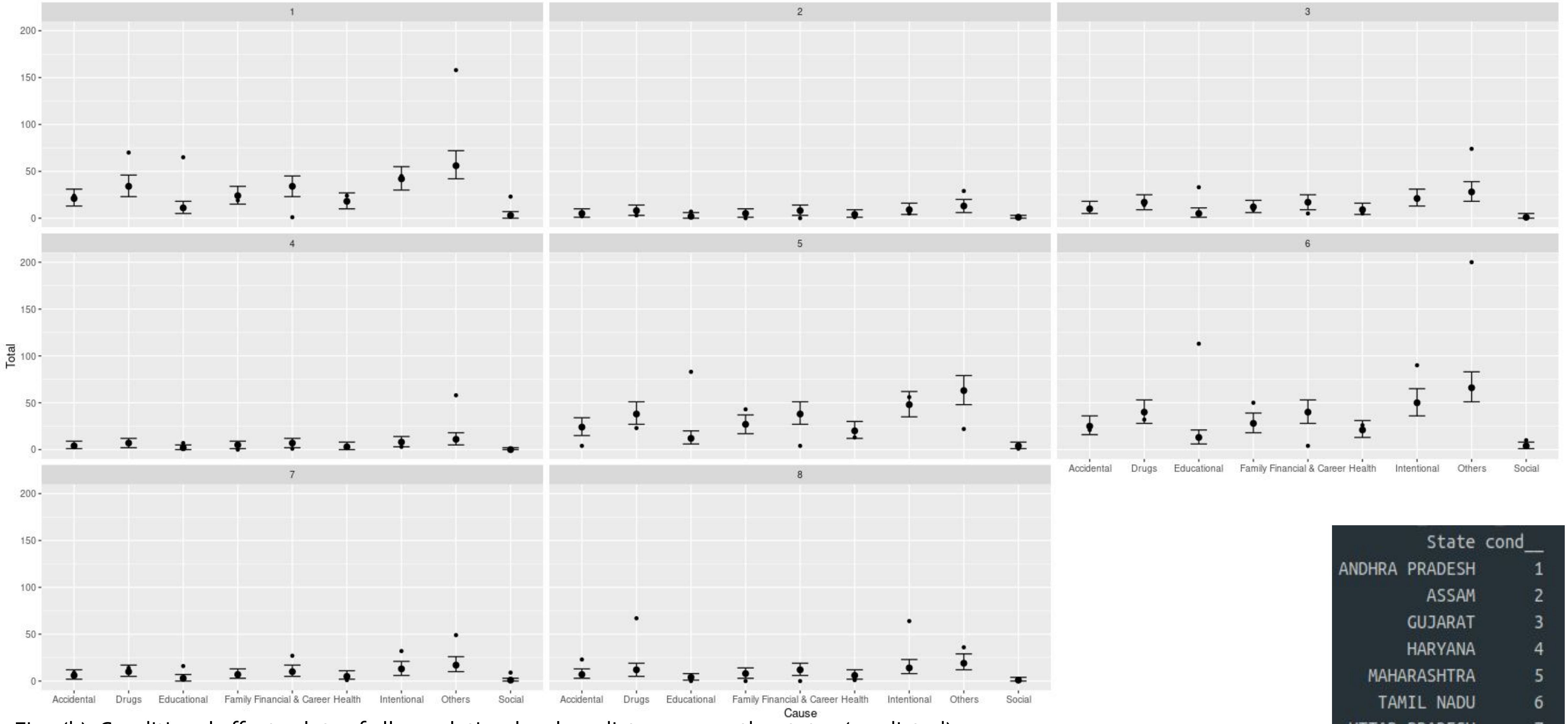


Fig 5(b): Conditional effects plots of all population-level predictors across the states (predicted)

State	cond__
ANDHRA PRADESH	1
ASSAM	2
GUJARAT	3
HARYANA	4
MAHARASHTRA	5
TAMIL NADU	6
UTTAR PRADESH	7
WEST BENGAL	8

Poisson | **Neg Binomial** | Zero inflated NB

technische universität dortmund

Significant covariates:

- Drugs
- Educational
- Intentional
- Social
- Others

Convergence diagnostics:

- Values of Rhat close to 1.00 suggest that the chains have converged to a common distribution.

ESS:

- Relatively high Bulk_ESS*.
- High Tail_ESS

Shape:

- Under Dispersion as Estimate(0.98) < 1
- observed counts have less variability in counts than would be expected under a Poisson distribution, although the deviation from 1 is relatively small.

***Bulk_ESS > num of chains * 100**

- **Source: Runtime warnings and convergence problems by Stan Development Team (link)**

```
> summary(negBinomial_model)
 Family: negbinomial
  Links: mu = log; shape = identity
Formula: Total ~ Cause + Gender + Age_group + (1 | State)
   Data: subset_df1 (Number of observations: 720)
  Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup draws = 4000

Group-Level Effects:
~State (Number of levels: 8)
              Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sd(Intercept)     0.98      0.32     0.56     1.77 1.00      917     1655

Population-Level Effects:
                      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
Intercept                 2.52      0.39     1.74     3.30 1.00     1009     1434
CauseDrugs                0.57      0.17     0.25     0.90 1.00     1958     2919
CauseEducational         -0.49      0.17    -0.83    -0.15 1.00     1811     2808
CauseFamily              -0.18      0.17    -0.51     0.14 1.00     2106     2897
CauseFinancial&Career     0.31      0.17    -0.01     0.63 1.00     2093     2724
CauseHealth              -0.21      0.17    -0.53     0.12 1.00     2027     2794
CauseIntentional          0.85      0.16     0.53     1.18 1.00     1936     2580
CauseOthers               1.31      0.17     0.98     1.65 1.00     2120     2597
CauseSocial              -1.75      0.18    -2.09    -1.40 1.00     2058     2789
GenderMale                0.62      0.08     0.46     0.77 1.00     4048     3184
Age_group15M29            3.06      0.13     2.81     3.31 1.00     2559     2772
Age_group30M44            2.82      0.13     2.57     3.08 1.00     2611     3043
Age_group45M59            2.18      0.13     1.92     2.44 1.00     2556     2992
Age_group60P              1.06      0.13     0.79     1.31 1.00     2654     2870

Family Specific Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
shape     0.98      0.05     0.88     1.08 1.00     4474     2757

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
```

Fig 6: Negative Binomial Model Summary

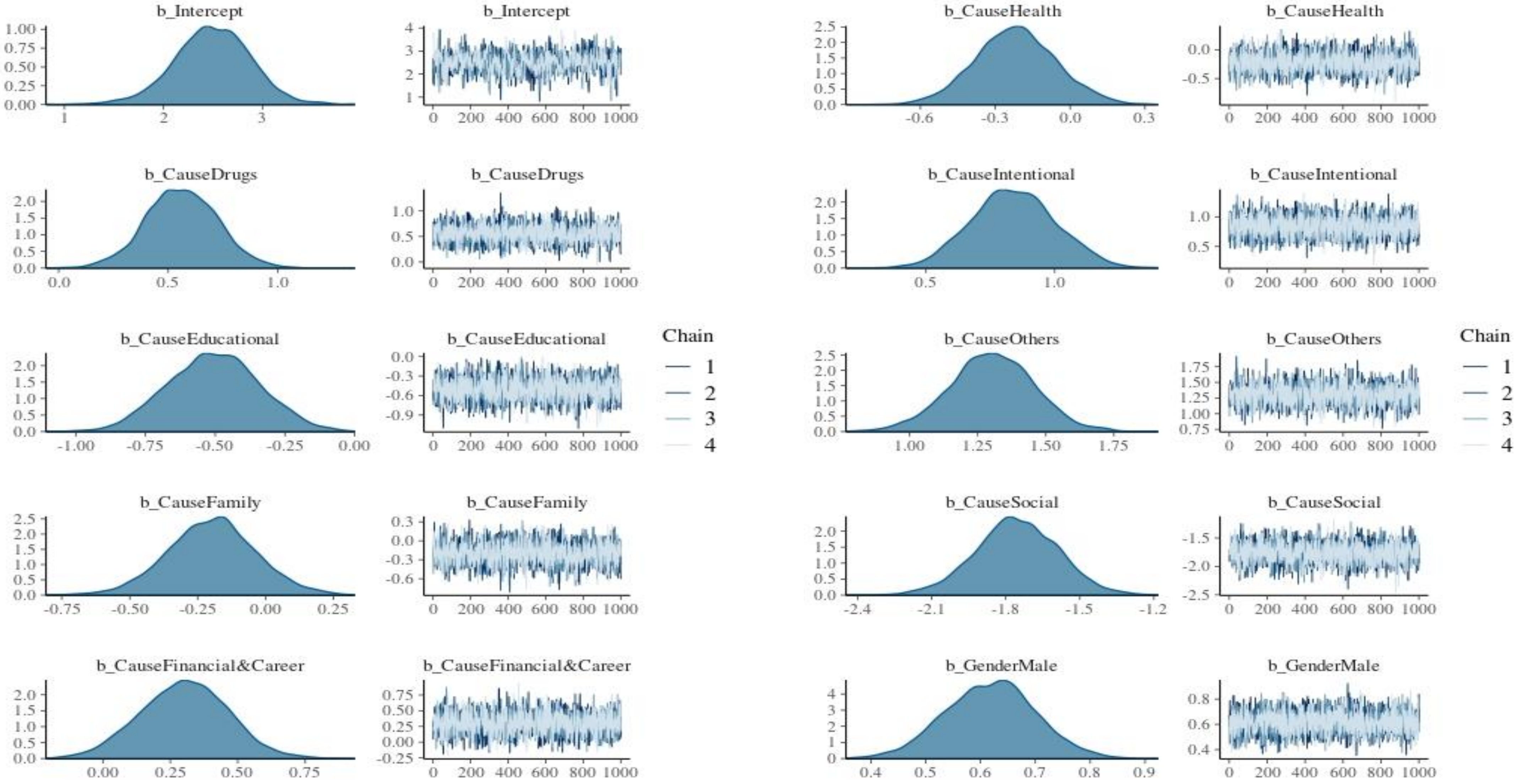
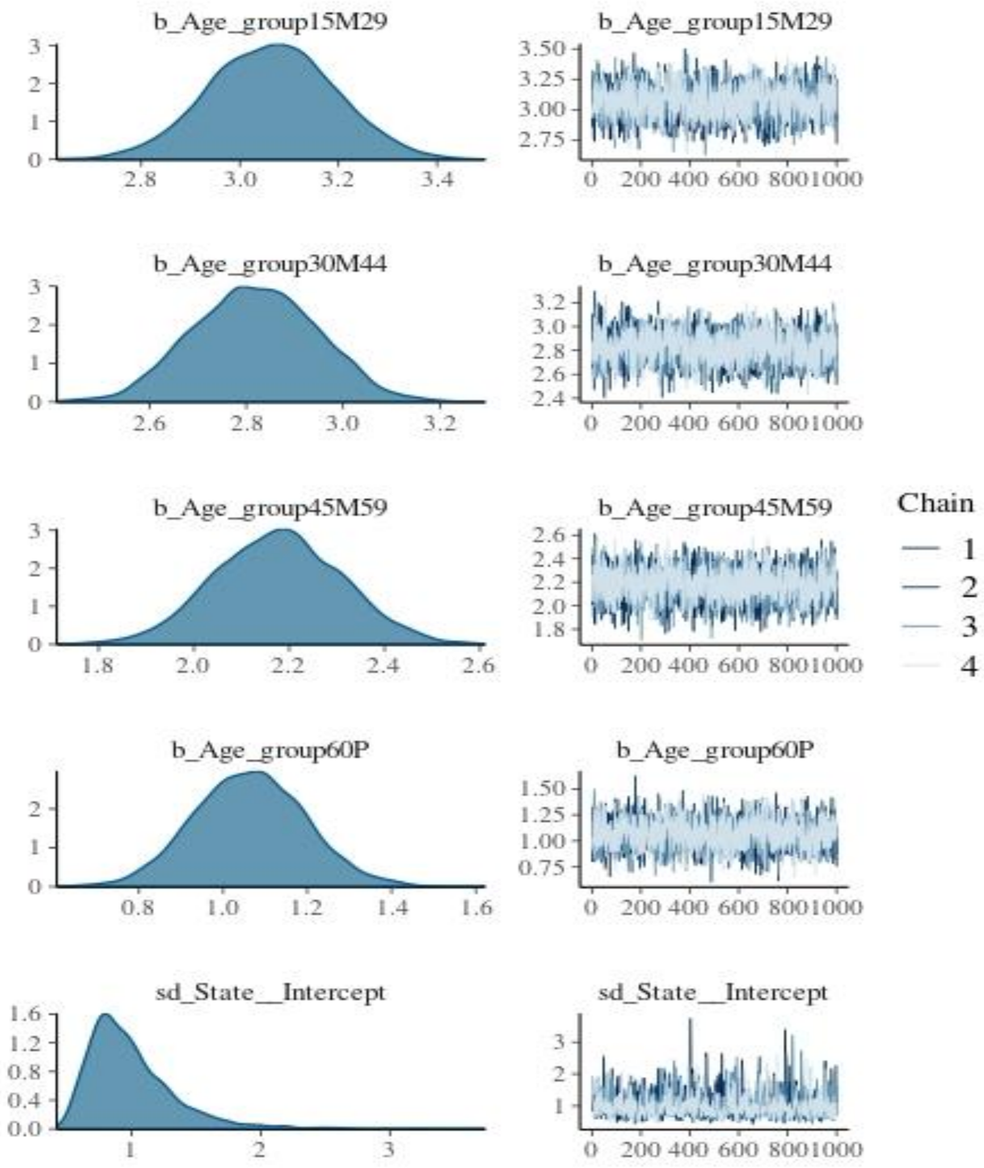
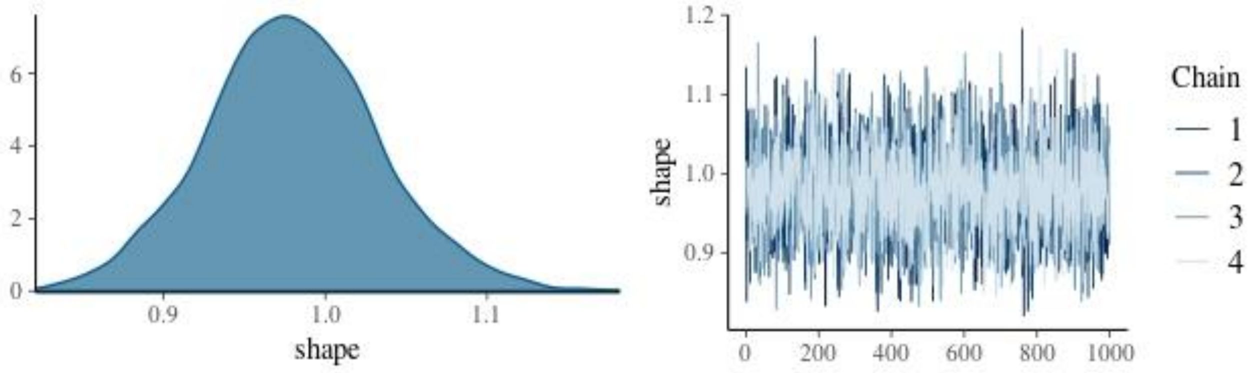


Fig 7: Trace and Density plots of all relevant parameters of the Negative Binomial model

Poisson | **Neg Binomial** | Zero inflated NB

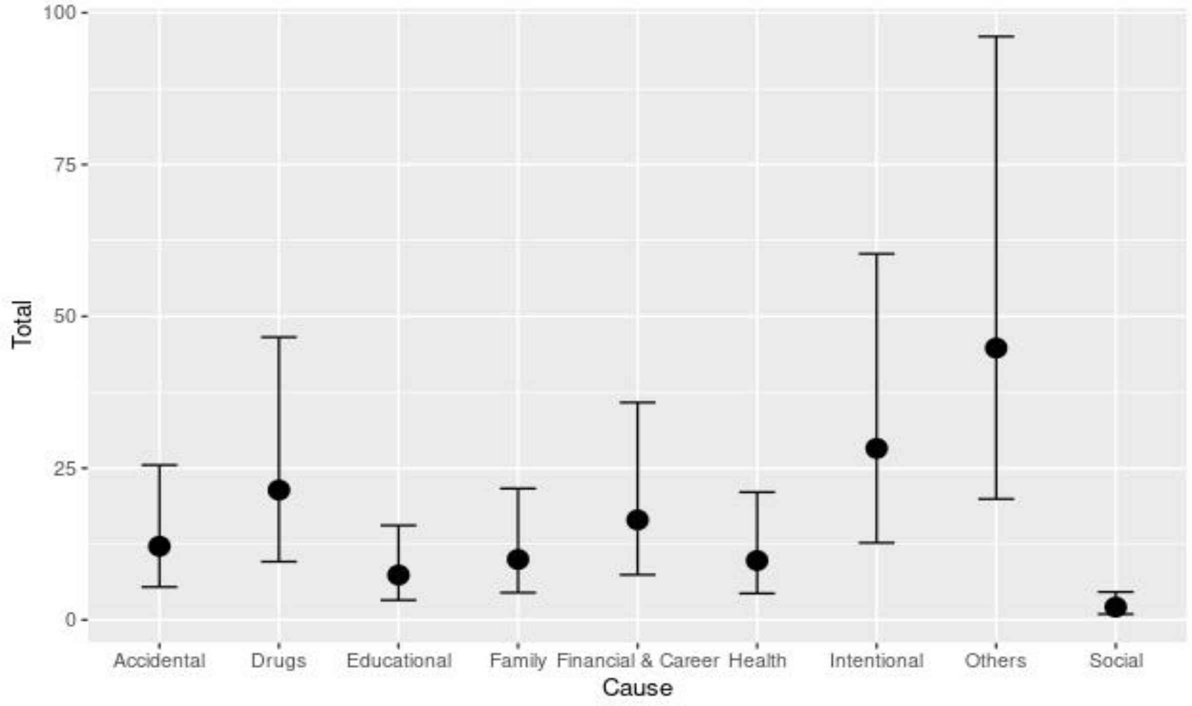


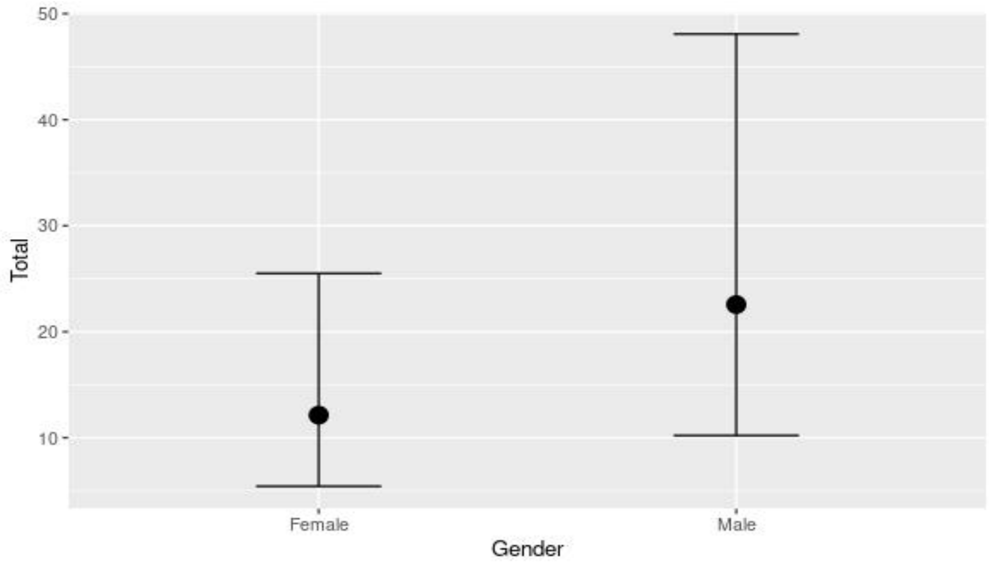
Extra Parameter for over dispersion Shape



- Trace plots shows that model has explored all the possible values it could look at so it converged well.
- Density plot shows the estimate found by model in Summary

Fig 7: Trace and Density plots of all relevant parameters of the Negative Binomial model





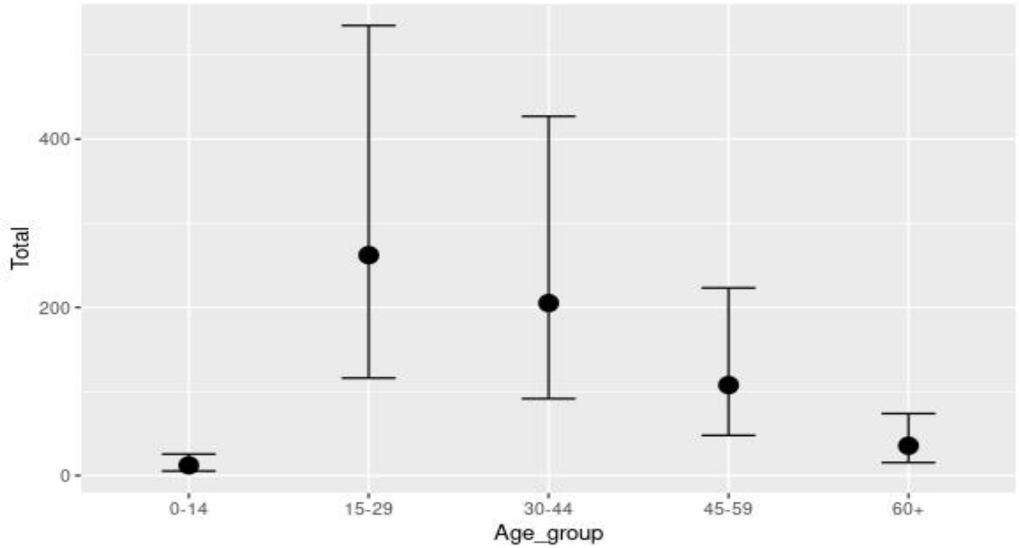


Fig 8(a): Conditional effects plots of all population-level predictors (fitted) technische universität dortmund

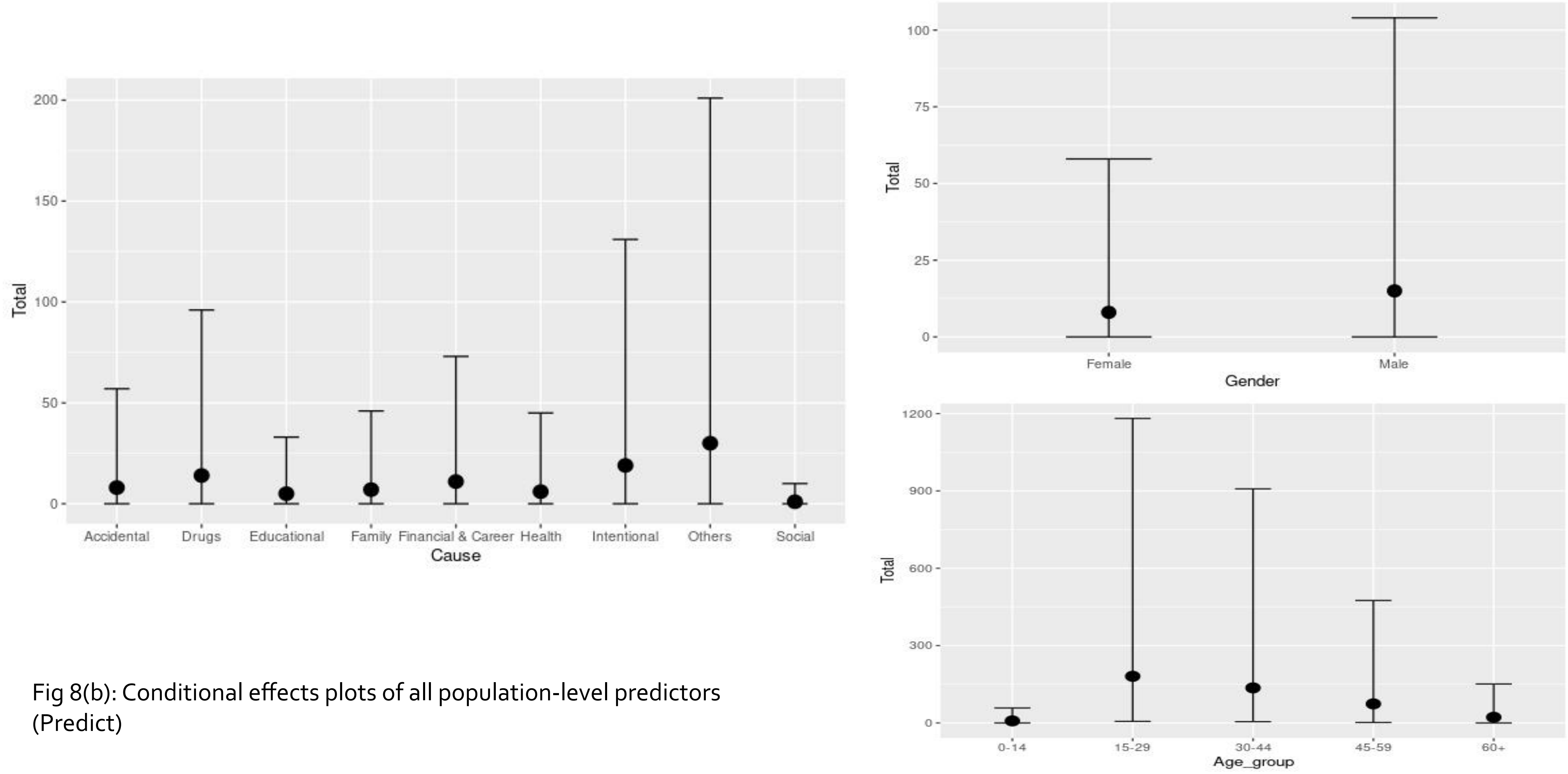


Fig 8(b): Conditional effects plots of all population-level predictors (Predict)

Poisson | **Neg Binomial** | Zero inflated NB

technische universität dortmund

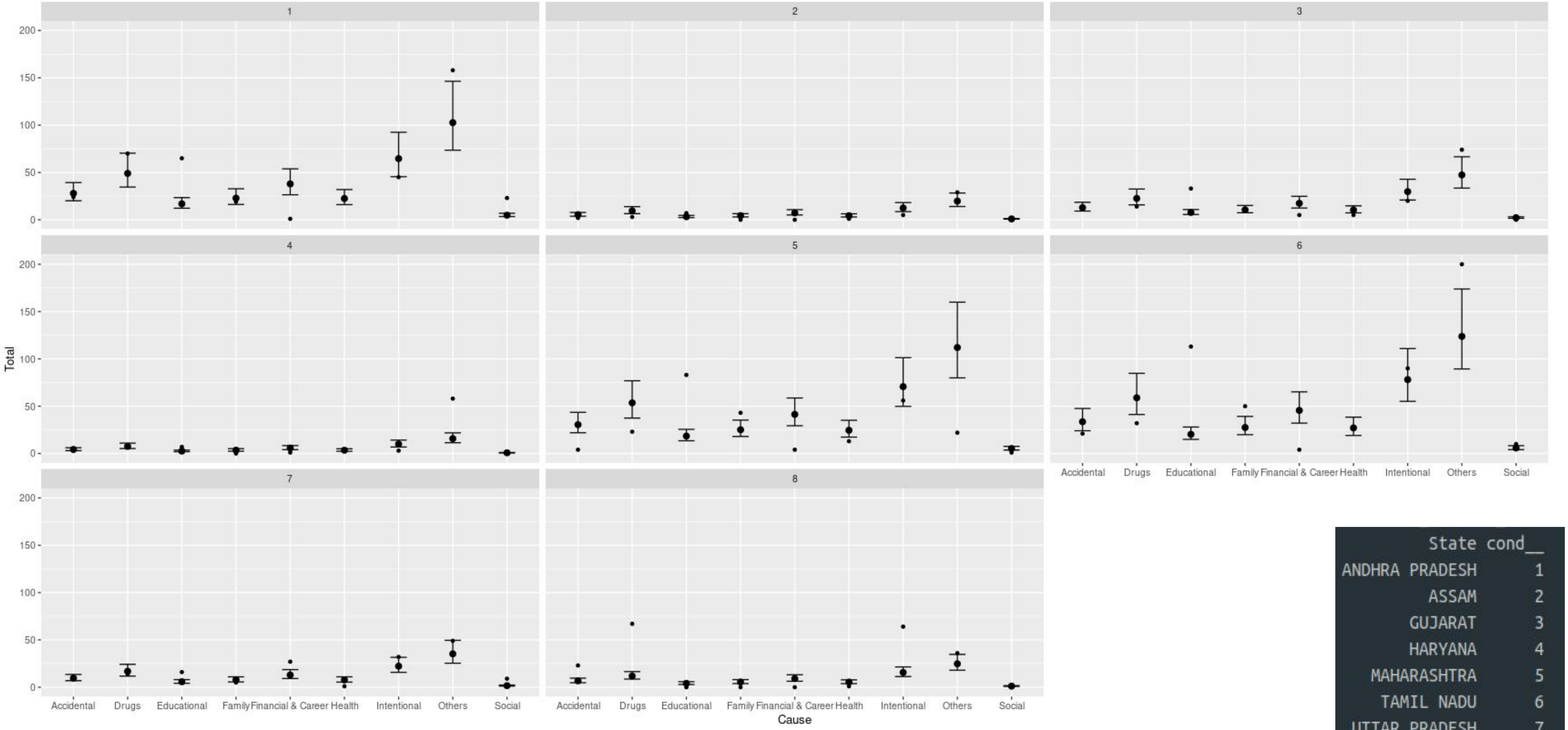
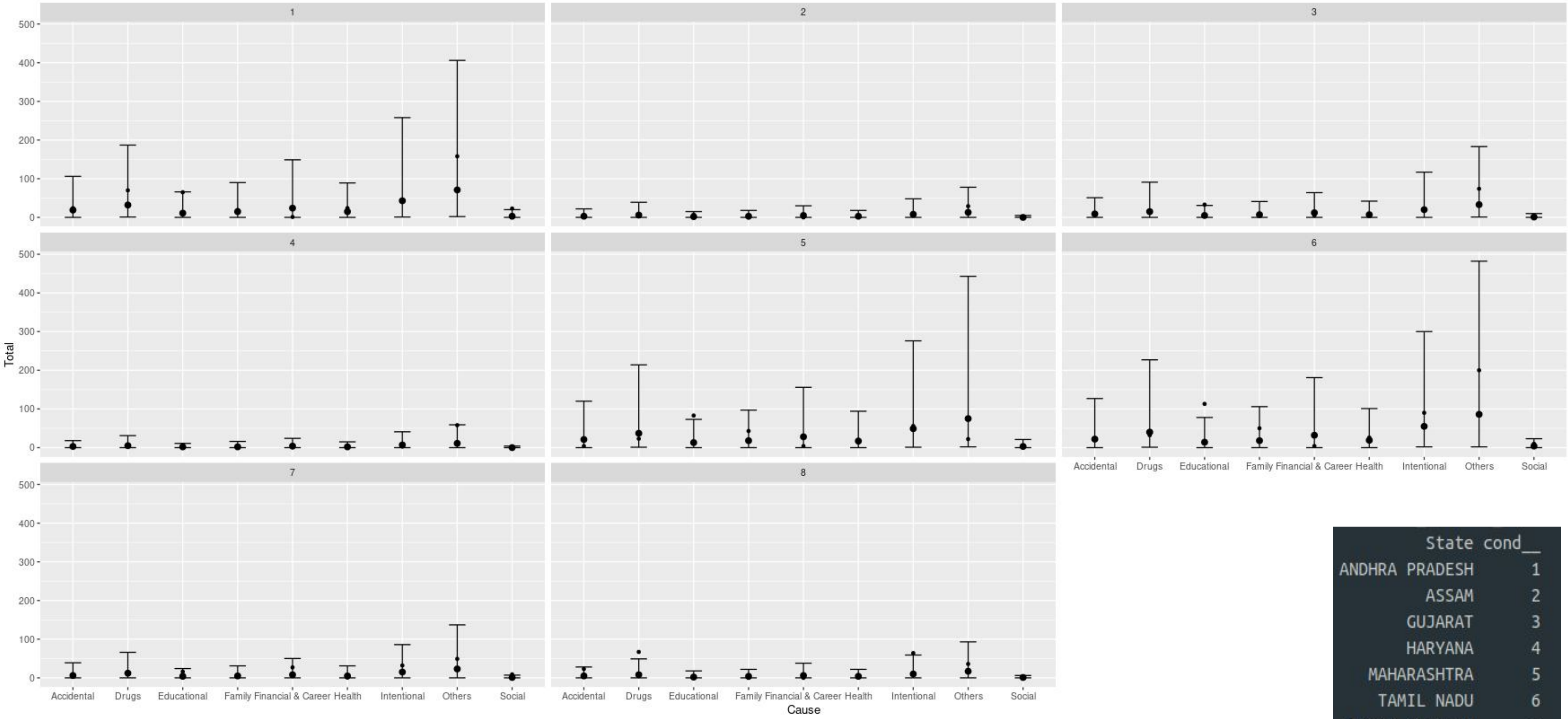


Fig 9(a) : Conditional effects plots of all population-level predictors across the states (fitted)

State	cond__
ANDHRA PRADESH	1
ASSAM	2
GUJARAT	3
HARYANA	4
MAHARASHTRA	5
TAMIL NADU	6
UTTAR PRADESH	7
WEST BENGAL	8

Poisson | **Neg Binomial** | Zero inflated NB



State	cond__
ANDHRA PRADESH	1
ASSAM	2
GUJARAT	3
HARYANA	4
MAHARASHTRA	5
TAMIL NADU	6
UTTAR PRADESH	7
WEST BENGAL	8

Fig 9(a) : Conditional effects plots of all population-level predictors across the states (predicted)

Poisson | Neg Binomial | **Zero inflated NB**

technische universität dortmund

Significant covariates:
- Drugs
- Financial & Career
- Intentional
- Social
- Others

Convergence diagnostics:
- Values of Rhat close to 1.00 suggest that the chains have converged to a common distribution.

ESS:
- High Bulk_ESS*.
- High Tail_ESS

Shape (Dispersion parameter):
- Over Dispersion as Estimate(1.54) > 1
- observed counts have more variability in the counts than would be expected under a Poisson distribution.

Zi (zero-inflation parameter):
- probability of excess zeros in the data
- small probability of observing excess zeros

***Bulk_ESS > num of chains * 100**
- **Source: Runtime warnings and convergence problems by Stan Development Team (link)**

```
> summary(zi_nbm)
 Family: zero_inflated_negbinomial
  Links: mu = log; shape = identity; zi = identity
Formula: Total ~ Cause + Gender + Age_group + (1 | State)
   Data: subset_df1 (Number of observations: 720)
  Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup draws = 4000

Group-Level Effects:
~State (Number of levels: 8)
              Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sd(Intercept)     0.94      0.29     0.55     1.67 1.00     1136     1891

Population-Level Effects:
                      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
Intercept                 2.46      0.35     1.75     3.18 1.01      805     1329
CauseDrugs                0.45      0.13     0.20     0.71 1.00     1841     2042
CauseEducational         -0.26      0.14    -0.55     0.02 1.00     2050     2484
CauseFamily               0.03      0.14    -0.24     0.29 1.00     1943     2608
CauseFinancial&Career     0.56      0.14     0.30     0.84 1.00     1845     2394
CauseHealth              -0.15      0.13    -0.41     0.10 1.00     1909     2263
CauseIntentional          0.74      0.13     0.48     1.01 1.00     1904     2351
CauseOthers               1.29      0.13     1.04     1.55 1.00     1944     2436
CauseSocial              -1.52      0.15    -1.82    -1.24 1.00     2205     2607
GenderMale                0.66      0.07     0.53     0.79 1.00     3790     3013
Age_group15M29            3.07      0.11     2.85     3.28 1.00     1896     2577
Age_group30M44            2.83      0.11     2.61     3.04 1.00     1900     2786
Age_group45M59            2.20      0.11     1.98     2.41 1.00     1757     2580
Age_group60P              1.13      0.11     0.90     1.34 1.00     1938     2710

Family Specific Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
shape     1.54      0.09     1.37     1.73 1.00     4290     2898
zi        0.06      0.01     0.05     0.08 1.00     4405     2492

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
```

Fig 10 :Zero Inflated Negative Binomial Model Summary

Poisson | Neg Binomial | **Zero inflated NB**

technische universität dortmund

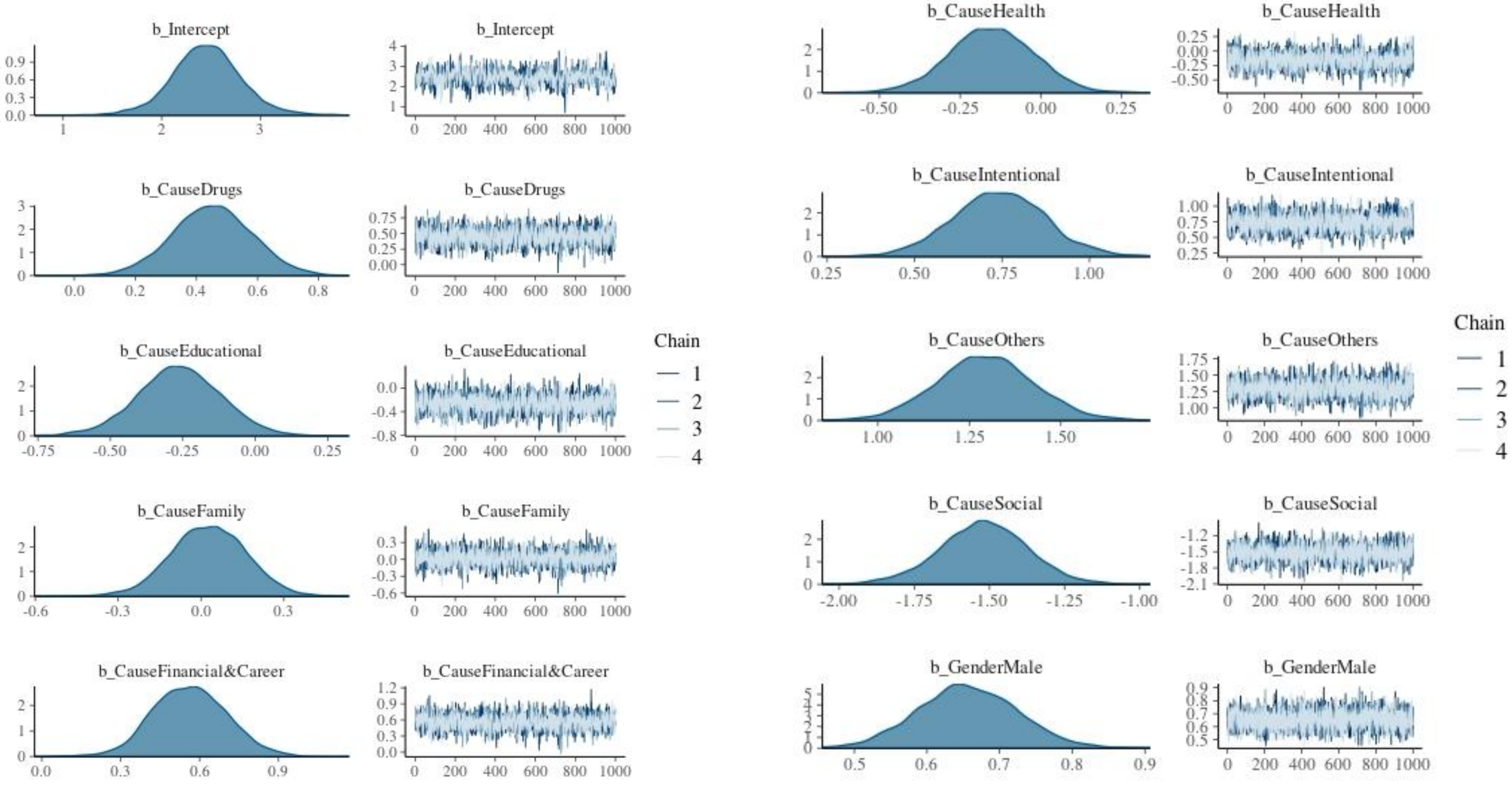
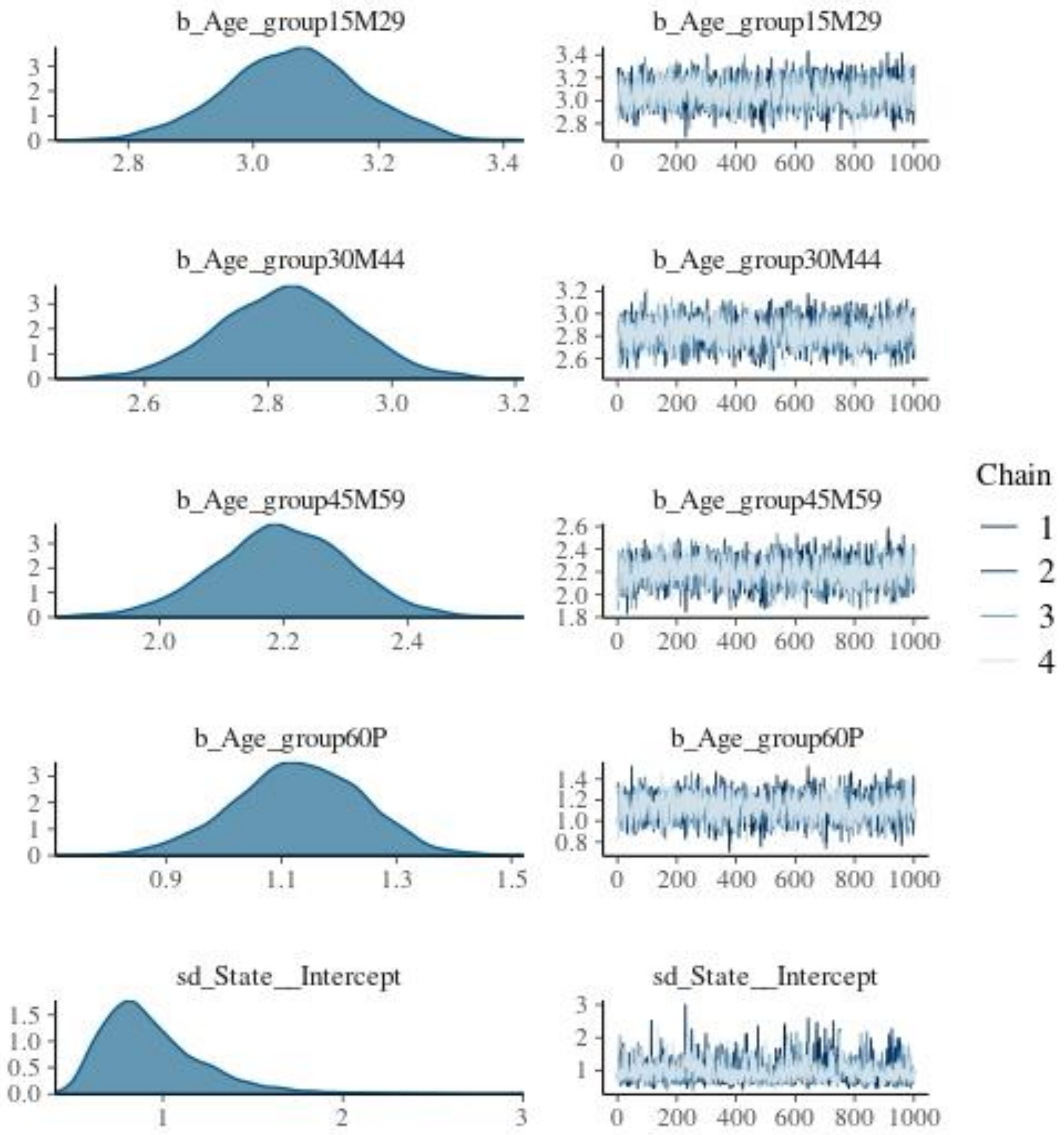


Fig 11: Trace and Density plots of all relevant parameters of Zero inflated Negative Binomial model



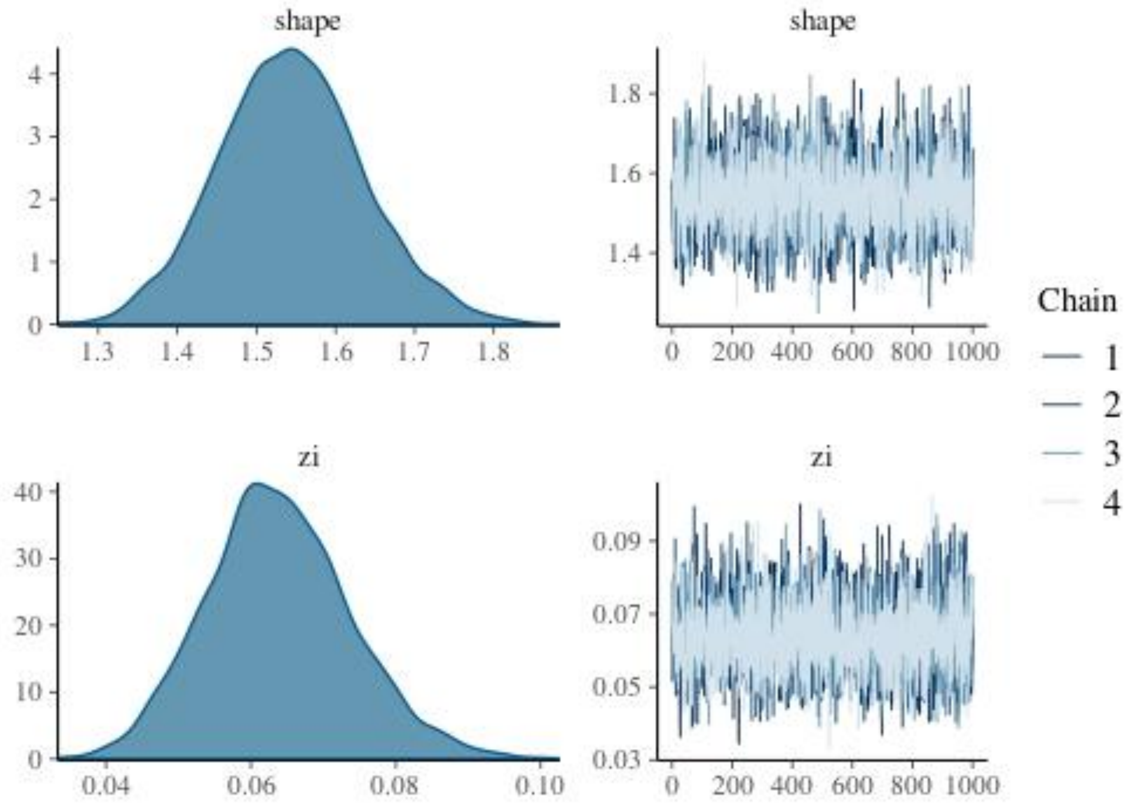


Fig 11: Trace and Density plots of all relevant parameters of Zero inflated Negative Binomial model

Poisson | Neg Binomial | **Zero inflated NB**

technische universität dortmund

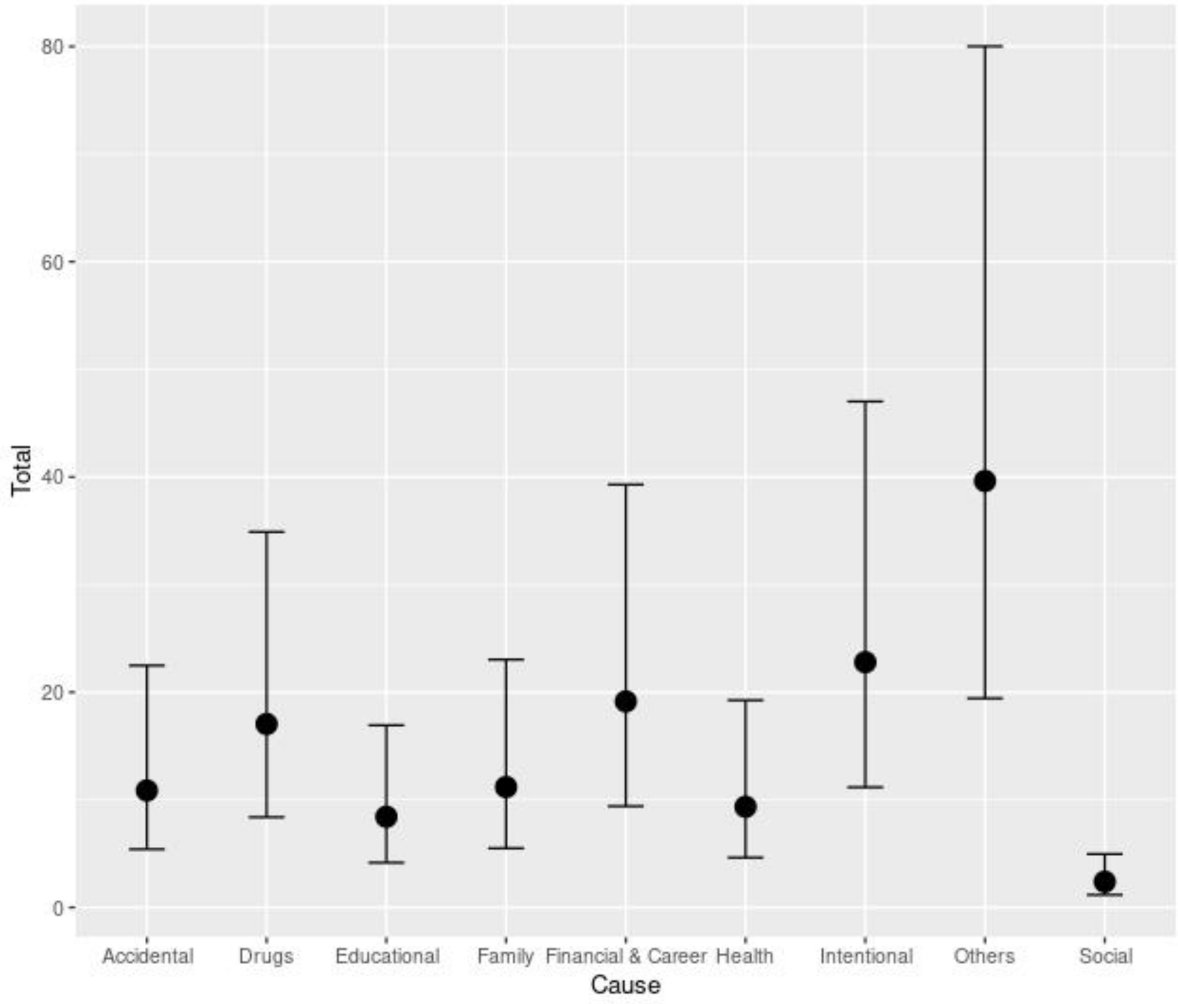
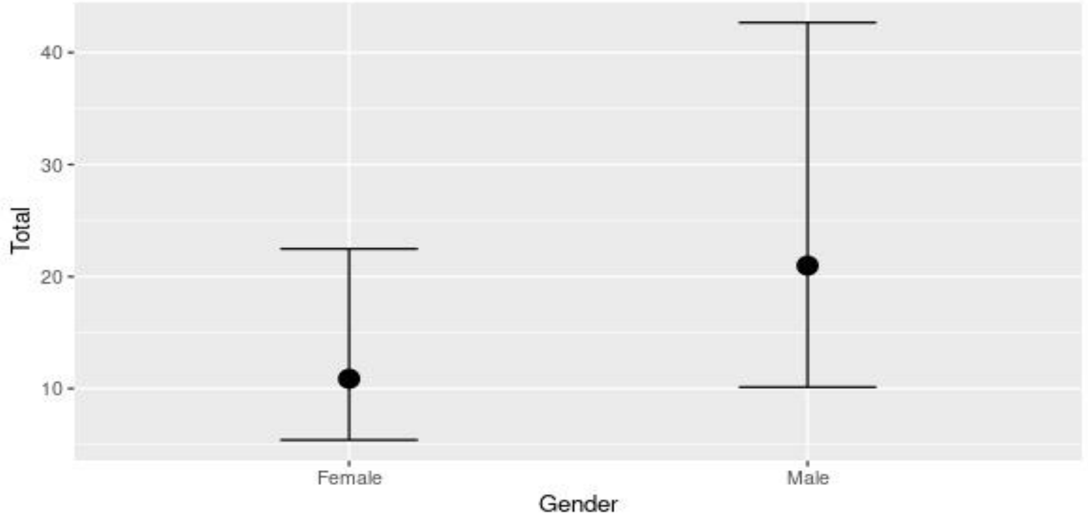
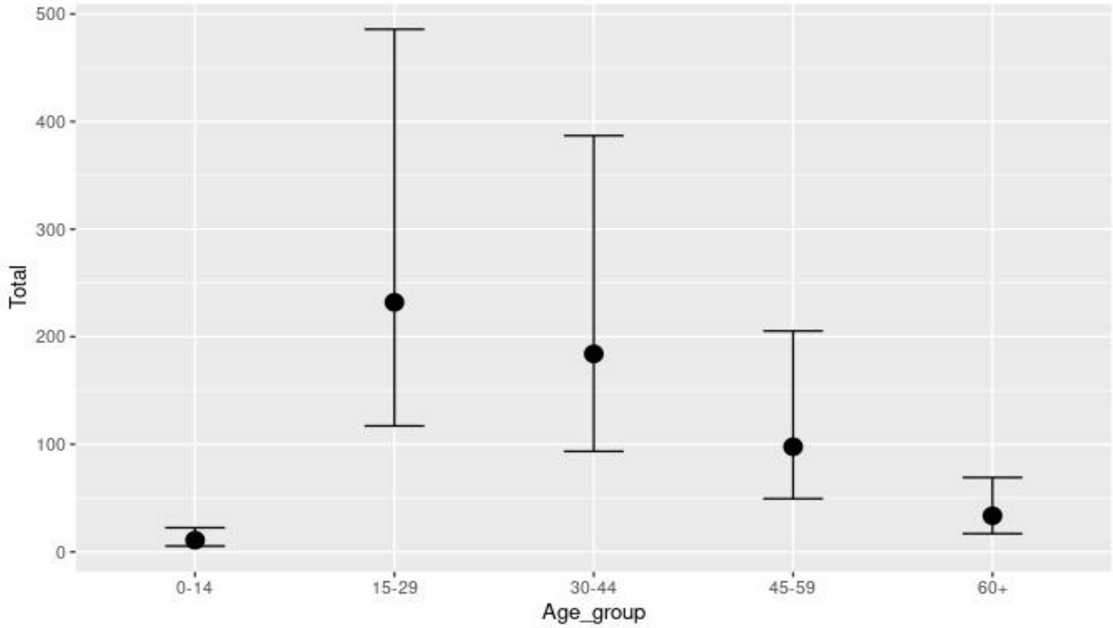
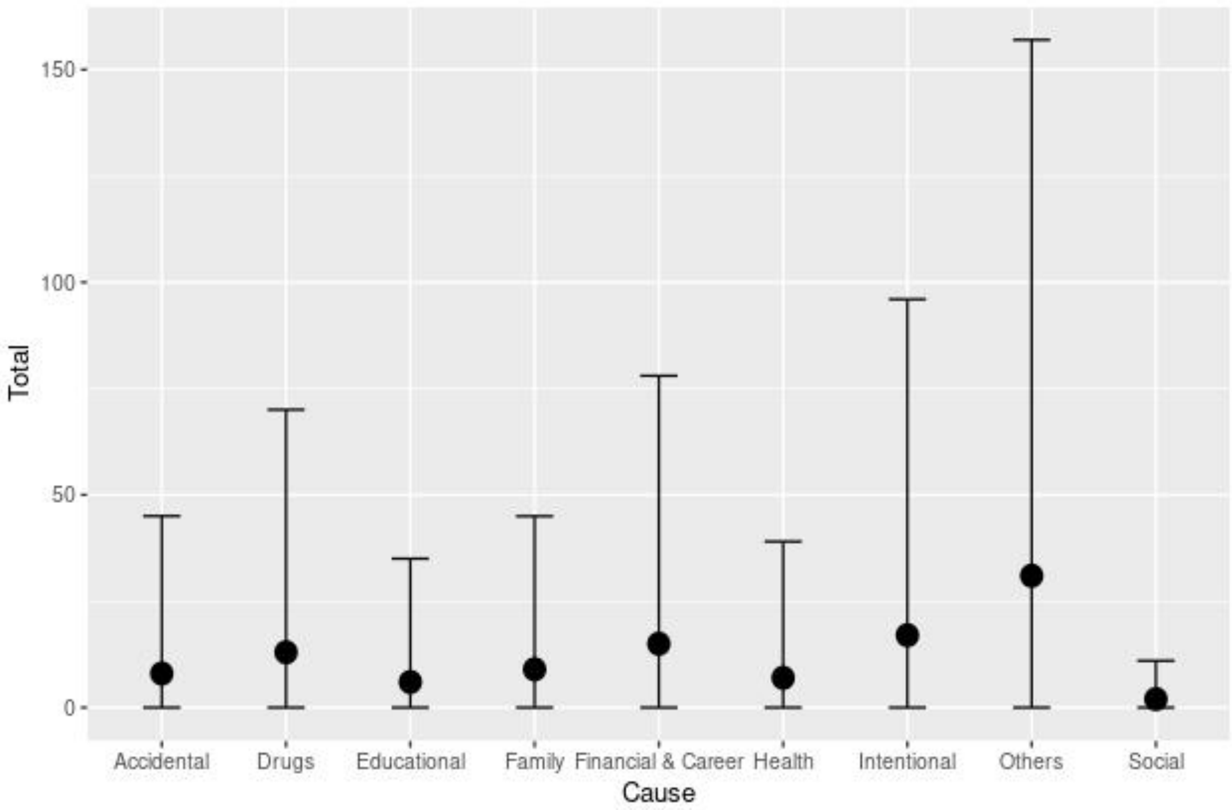


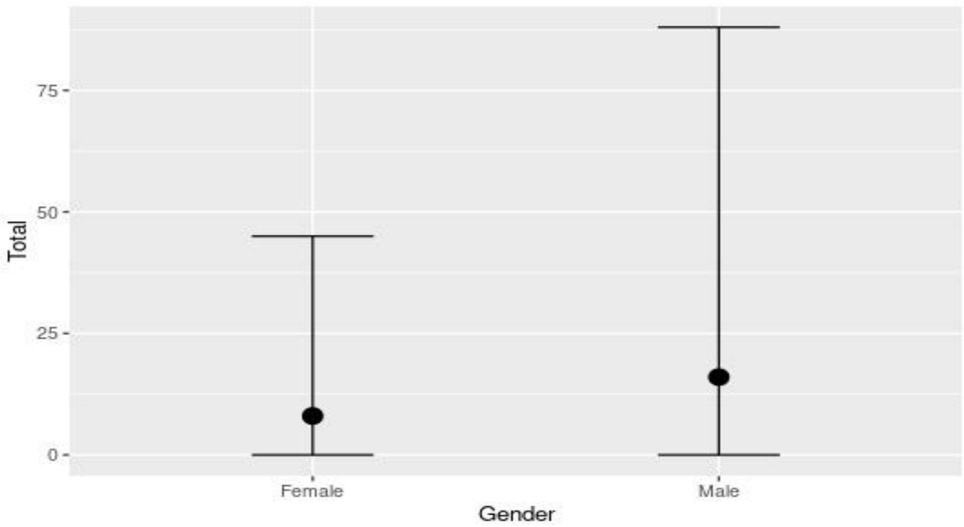
Fig 12(a): Conditional effects plots of all population-level predictors (fitted)





Poisson | Neg Binomial | **Zero inflated NB**





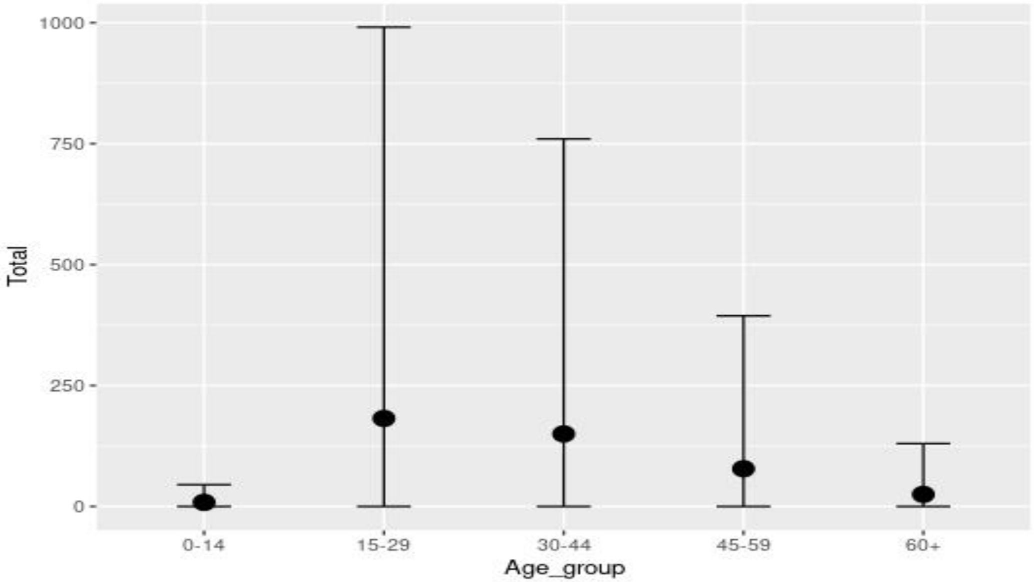


Fig 12(b): Conditional effects plots of all population-level predictors (predicted)

Poisson | Neg Binomial | **Zero inflated NB**

technische universität dortmund

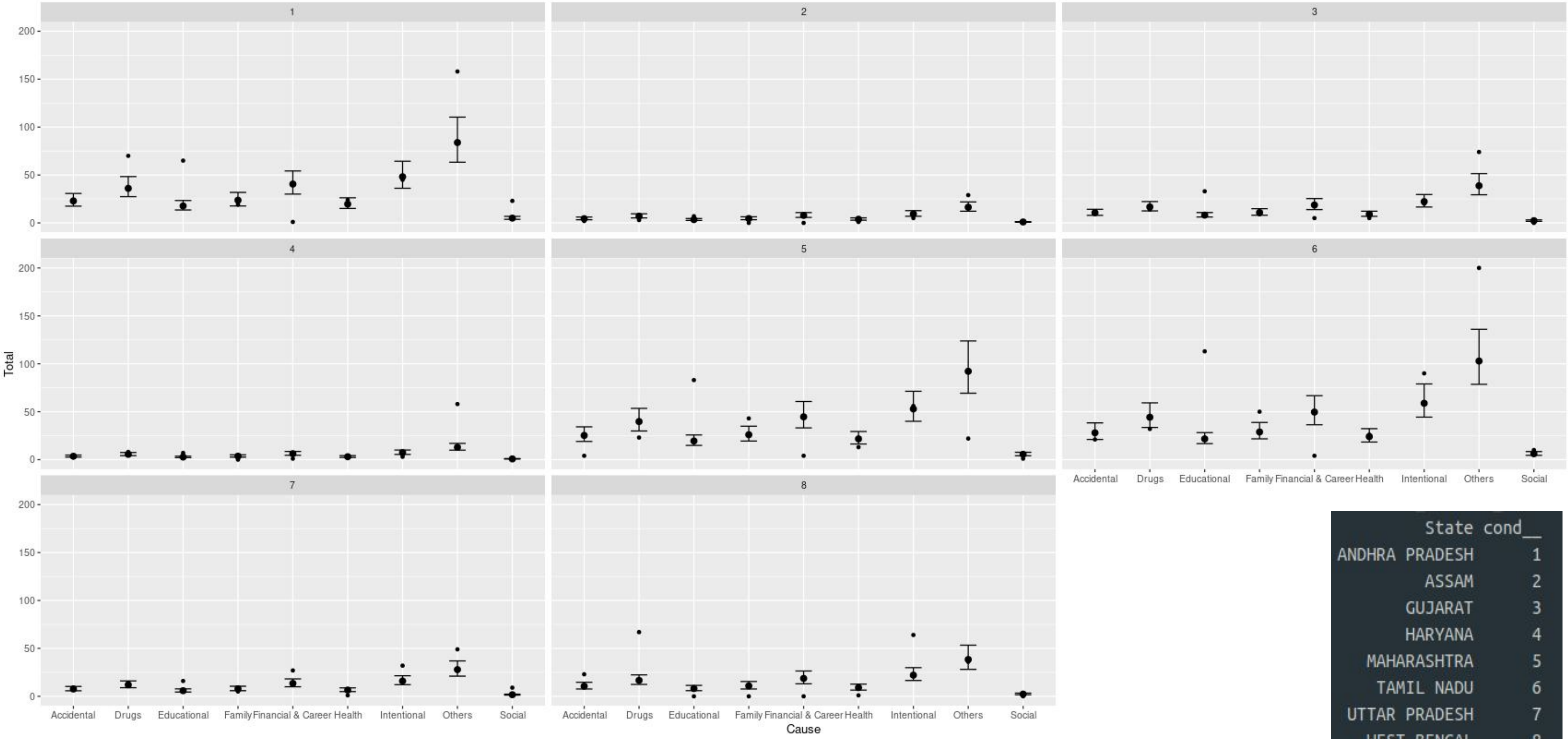
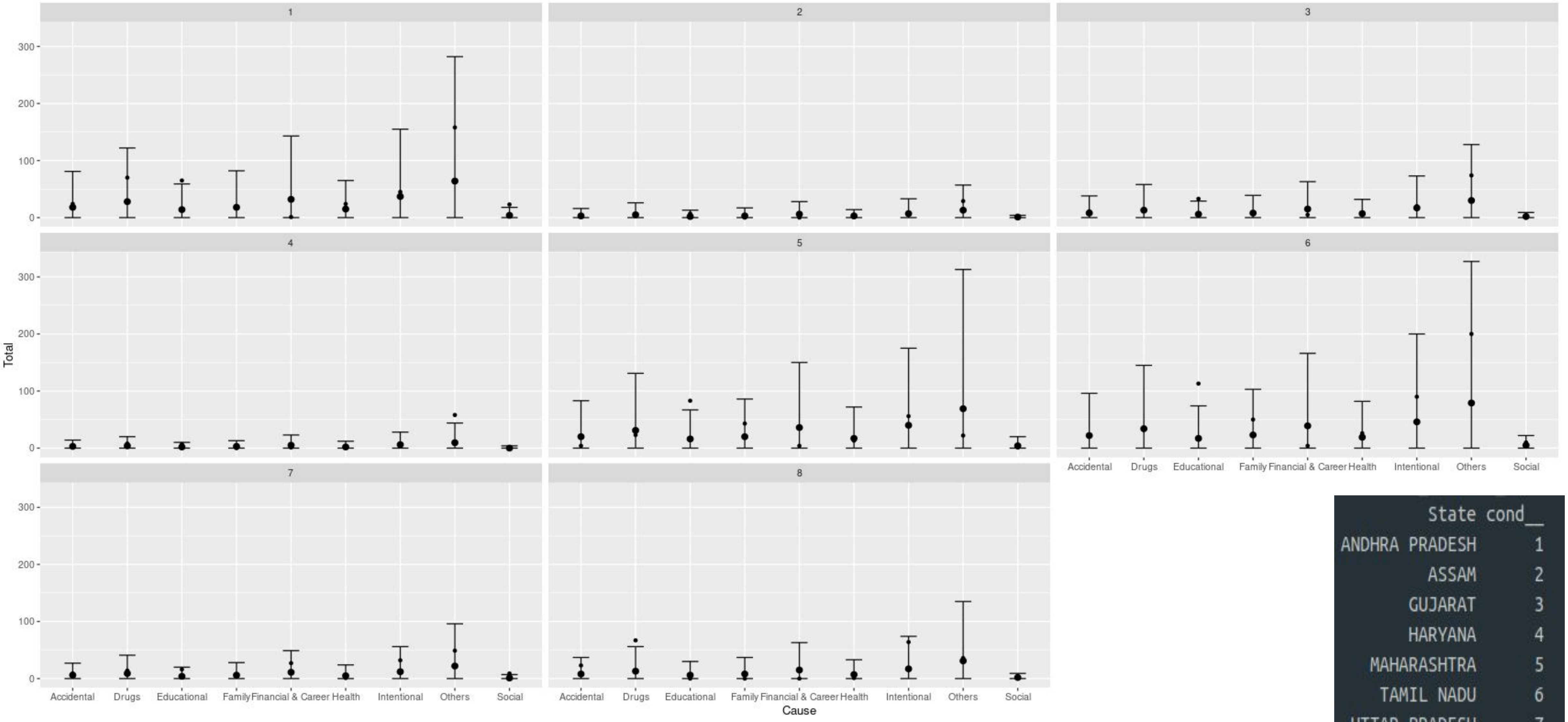


Fig 13(a): Conditional effects plots of all population-level predictors across the states (fitted)

State	cond__
ANDHRA PRADESH	1
ASSAM	2
GUJARAT	3
HARYANA	4
MAHARASHTRA	5
TAMIL NADU	6
UTTAR PRADESH	7
WEST BENGAL	8

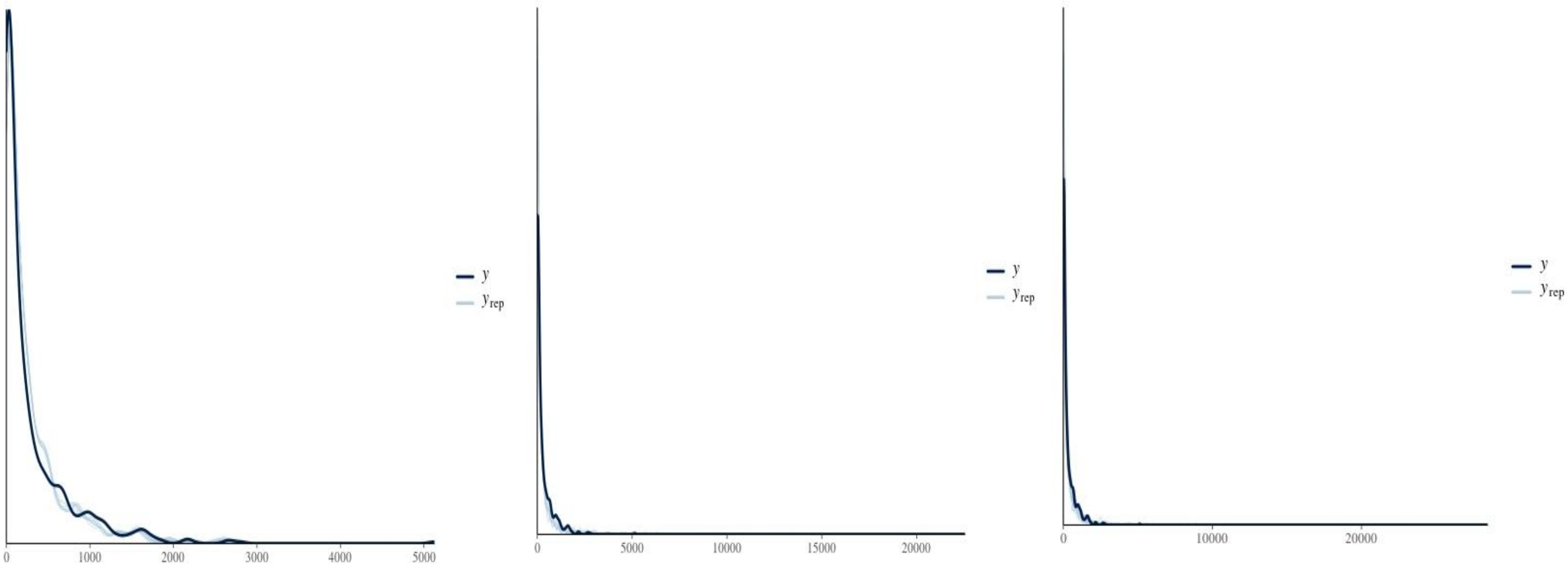
Poisson | Neg Binomial | **Zero inflated NB**



State	cond__
ANDHRA PRADESH	1
ASSAM	2
GUJARAT	3
HARYANA	4
MAHARASHTRA	5
TAMIL NADU	6
UTTAR PRADESH	7
WEST BENGAL	8

Fig 13(b): Conditional effects plots of all population-level predictors across the states (predicted)

Posterior Check



Poisson

Negative Binomial

Zero inflated negative binomial

Fig 14 : Posterior Predictive Check Graph for all models

Model Comparison

Models	elpd_diff	se_diff
Zero-inflated Negative Binomial Model	0.0	0.0
Negative Binomial Model	-55.2	12.9
Poisson Model	-34573.2	2687.0

Table 3: Model comparison using "loo()" function

- Higher the ELPD score the more well model fits the data.
- Zero-inflated negative binomial model performs best as it has highest ELPD (Expected log predictive density) score.
- It is taken as a base model for comparison.

Prior Sensitivity Check

```
> prior_summary(zi_nbm)
               prior     class                 coef group resp dpar nlpar lb ub       source
              (flat)         b                                                       default
              (flat)         b        Age_group15M29                            (vectorized)
              (flat)         b        Age_group30M44                            (vectorized)
              (flat)         b        Age_group45M59                            (vectorized)
              (flat)         b          Age_group60P                            (vectorized)
              (flat)         b            CauseDrugs                            (vectorized)
              (flat)         b       CauseEducational                           (vectorized)
              (flat)         b           CauseFamily                            (vectorized)
              (flat)         b CauseFinancial&Career                            (vectorized)
              (flat)         b           CauseHealth                            (vectorized)
              (flat)         b      CauseIntentional                            (vectorized)
              (flat)         b           CauseOthers                            (vectorized)
              (flat)         b           CauseSocial                            (vectorized)
              (flat)         b            GenderMale                            (vectorized)
 student_t(3, 4.6, 2.5) Intercept                                                    default
   student_t(3, 0, 2.5)        sd                                          0         default
   student_t(3, 0, 2.5)        sd                      State               0    (vectorized)
   student_t(3, 0, 2.5)        sd            Intercept State               0    (vectorized)
      gamma(0.01, 0.01)     shape                                          0         default
            beta(1, 1)        zi                                           0  1      default
```

Fig 15: Default Prior Settings

Prior Sensitivity Check

(a)

```
priors_set1<- c(
  set_prior("normal(0, 5)", class = "b"),
  set_prior("normal(0, 10)", class = "Intercept"),
  set_prior("cauchy(0, 2)", class = "sd"),
  set_prior("beta(1, 1)", class = "zi"),
  set_prior("gamma(2, 0.5)", class = "shape")
)
```

(b)

```
priors_set2 <- c(
  set_prior("normal(0, 2)", class = "b"),
  set_prior("normal(0, 5)", class = "Intercept"),
  set_prior("cauchy(0, 1)", class = "sd"),
  set_prior("beta(2, 2)", class = "zi"),
  set_prior("gamma(2, 0.1)", class = "shape")
)
```

Fig 16):Different Prior Settings (a) Prior Setting-1, (b) Prior Setting-2

Prior Sensitivity Check

Fig 17: Prior Summary of Two models fitted with different Prior settings

```
> prior_summary(PSA_zi_nbm1)
        prior     class                 coef group resp dpar nlpar lb ub       source
  normal(0, 5)        b                                                          user
  normal(0, 5)        b      Age_group15M29                                (vectorized)
  normal(0, 5)        b      Age_group30M44                                (vectorized)
  normal(0, 5)        b      Age_group45M59                                (vectorized)
  normal(0, 5)        b        Age_group60P                                (vectorized)
  normal(0, 5)        b          CauseDrugs                                (vectorized)
  normal(0, 5)        b     CauseEducational                               (vectorized)
  normal(0, 5)        b         CauseFamily                                (vectorized)
  normal(0, 5)        b CauseFinancial&Career                              (vectorized)
  normal(0, 5)        b         CauseHealth                                (vectorized)
  normal(0, 5)        b     CauseIntentional                               (vectorized)
  normal(0, 5)        b         CauseOthers                                (vectorized)
  normal(0, 5)        b         CauseSocial                                (vectorized)
  normal(0, 5)        b          GenderMale                                (vectorized)
 normal(0, 10) Intercept                                                         user
  cauchy(0, 2)       sd                                          0               user
  cauchy(0, 2)       sd                       State              0         (vectorized)
  cauchy(0, 2)       sd             Intercept State              0         (vectorized)
 gamma(2, 0.5)    shape                                          0               user
    beta(1, 1)       zi                                          0  1            user
```

Fig 17(a): Prior setting-1

```
> prior_summary(PSA_zi_nbm2)
        prior     class                 coef group resp dpar nlpar lb ub       source
  normal(0, 2)        b                                                          user
  normal(0, 2)        b      Age_group15M29                                (vectorized)
  normal(0, 2)        b      Age_group30M44                                (vectorized)
  normal(0, 2)        b      Age_group45M59                                (vectorized)
  normal(0, 2)        b        Age_group60P                                (vectorized)
  normal(0, 2)        b          CauseDrugs                                (vectorized)
  normal(0, 2)        b     CauseEducational                               (vectorized)
  normal(0, 2)        b         CauseFamily                                (vectorized)
  normal(0, 2)        b CauseFinancial&Career                              (vectorized)
  normal(0, 2)        b         CauseHealth                                (vectorized)
  normal(0, 2)        b     CauseIntentional                               (vectorized)
  normal(0, 2)        b         CauseOthers                                (vectorized)
  normal(0, 2)        b         CauseSocial                                (vectorized)
  normal(0, 2)        b          GenderMale                                (vectorized)
  normal(0, 5) Intercept                                                         user
  cauchy(0, 1)       sd                                          0               user
  cauchy(0, 1)       sd                       State              0         (vectorized)
  cauchy(0, 1)       sd             Intercept State              0         (vectorized)
 gamma(2, 0.1)    shape                                          0               user
    beta(2, 2)       zi                                          0  1            user
```

Fig 17(b): Prior setting-2

Prior Sensitivity Check

```
> summary(PSA_zi_nbm1)
 Family: zero_inflated_negbinomial
  Links: mu = log; shape = identity; zi = identity
Formula: Total ~ Cause + Gender + Age_group + (1 | State)
   Data: subset_df1 (Number of observations: 720)
  Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup draws = 4000

Group-Level Effects:
~State (Number of levels: 8)
              Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sd(Intercept)     0.93      0.31     0.52     1.69 1.00      801     1224

Population-Level Effects:
                      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
Intercept                 2.48      0.38     1.76     3.24 1.00      782      877
CauseDrugs                0.45      0.13     0.20     0.70 1.00     1631     1888
CauseEducational         -0.26      0.14    -0.55     0.01 1.00     1922     2532
CauseFamily               0.03      0.13    -0.24     0.29 1.00     1929     2635
CauseFinancial&Career     0.56      0.14     0.30     0.83 1.00     1853     2224
CauseHealth              -0.15      0.13    -0.41     0.10 1.00     1814     2223
CauseIntentional          0.74      0.13     0.48     0.99 1.00     1761     1755
CauseOthers               1.29      0.13     1.04     1.54 1.00     1570     2427
CauseSocial              -1.52      0.15    -1.81    -1.23 1.00     1826     2128
GenderMale                0.66      0.07     0.52     0.79 1.00     3454     3294
Age_group15M29            3.06      0.11     2.85     3.27 1.00     2237     2679
Age_group30M44            2.83      0.11     2.61     3.05 1.00     2166     2491
Age_group45M59            2.19      0.11     1.98     2.41 1.00     2086     2703
Age_group60P              1.13      0.11     0.90     1.34 1.00     2175     2564

Family Specific Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
shape     1.55      0.09     1.37     1.73 1.00     4221     2654
zi        0.06      0.01     0.05     0.08 1.00     3484     2834

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
```

Fig 18 (a): Model summary of prior settings-1

```
> summary(PSA_zi_nbm2)
 Family: zero_inflated_negbinomial
  Links: mu = log; shape = identity; zi = identity
Formula: Total ~ Cause + Gender + Age_group + (1 | State)
   Data: subset_df1 (Number of observations: 720)
  Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup draws = 4000

Group-Level Effects:
~State (Number of levels: 8)
              Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sd(Intercept)     0.89      0.27     0.53     1.54 1.00      972     1614

Population-Level Effects:
                      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
Intercept                 2.47      0.34     1.79     3.13 1.00     1065     1487
CauseDrugs                0.45      0.13     0.19     0.70 1.00     1911     2632
CauseEducational         -0.27      0.15    -0.55     0.02 1.00     1816     2761
CauseFamily               0.03      0.13    -0.24     0.29 1.00     1997     2827
CauseFinancial&Career     0.57      0.14     0.30     0.84 1.00     1716     2511
CauseHealth              -0.15      0.13    -0.41     0.11 1.00     1766     2462
CauseIntentional          0.74      0.13     0.48     0.99 1.00     2010     2742
CauseOthers               1.29      0.13     1.03     1.55 1.00     1853     2658
CauseSocial              -1.51      0.15    -1.79    -1.23 1.00     1983     2761
GenderMale                0.66      0.07     0.53     0.78 1.00     3230     2613
Age_group15M29            3.05      0.11     2.84     3.25 1.00     2142     2502
Age_group30M44            2.81      0.11     2.59     3.02 1.00     1934     2564
Age_group45M59            2.18      0.11     1.97     2.40 1.00     1950     2338
Age_group60P              1.12      0.11     0.89     1.33 1.00     1954     2191

Family Specific Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
shape     1.56      0.09     1.39     1.74 1.00     3928     2524
zi        0.07      0.01     0.05     0.09 1.00     3822     2700

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
```

Fig 18 (b): Model summary of prior settings-2

Prior Sensitivity Check

technische universität dortmund

Model Comparison

Zero-inflated Negative Binomial Model	**elpd_diff**	**se_diff**
with default settings	0.0	0.0
with Prior settings 2	-0.2	0.4
with Prior settings 1	-0.5	0.2

Table 4: Model comparison with different priors

- Zero inflated Negative binomial model is not sensitive to priors as the difference in the elpd score is quite small.

Results

- The number of suicides tend to be lower among children and older adults compared to middle-aged adults.
- Suicide is a more prevalent issue among males compareda to females in India.
- The main cause of suicides in India are intentional and financial & career-related problems.

Summary

- Cleaned Dataset contains 720 rows and 4 covariates
- Different Models like Poisson Model, Negative Binomial and Zero-inflated negative binomial models are fitted and variation of number of suicides based on gender, age groups, cause and each state in India is discussed
- Models are compared and Zero-inflated negative binomial model is the best fit
- Prior Sensitivity Analysis is performed on the best model.
- The model is not sensitive to priors as elpd_diff is quite small.

Contact info:

Aakash Goyal (229975)
aakash.goyal@tu-dortmund.de

Jaykumar Savani (230443)
jaykumar.savani@tu-dortmund.de