NHANES Analysis

14/04/23

List version and package information

```
## [1] "R version 4.2.2 (2022-10-31)"
```

```
## [1] "Package versions:"
```

	x
loo	2.5.1
doParallel	1.0.17
iterators	1.0.14
foreach	1.5.2
rstiefel	1.0.1
knitr	1.41
patchwork	1.1.2
tidybayes	3.0.2
cmdstanr	0.5.3
forcats	0.5.2
stringr	1.5.0
dplyr	1.0.10
purrr	1.0.1
readr	2.1.3
tidyr	1.2.1
tibble	3.1.8
ggplot2	3.3.6
tidyverse	1.3.2

```
## [1] "Fraction of treated units is 0.073297"
```

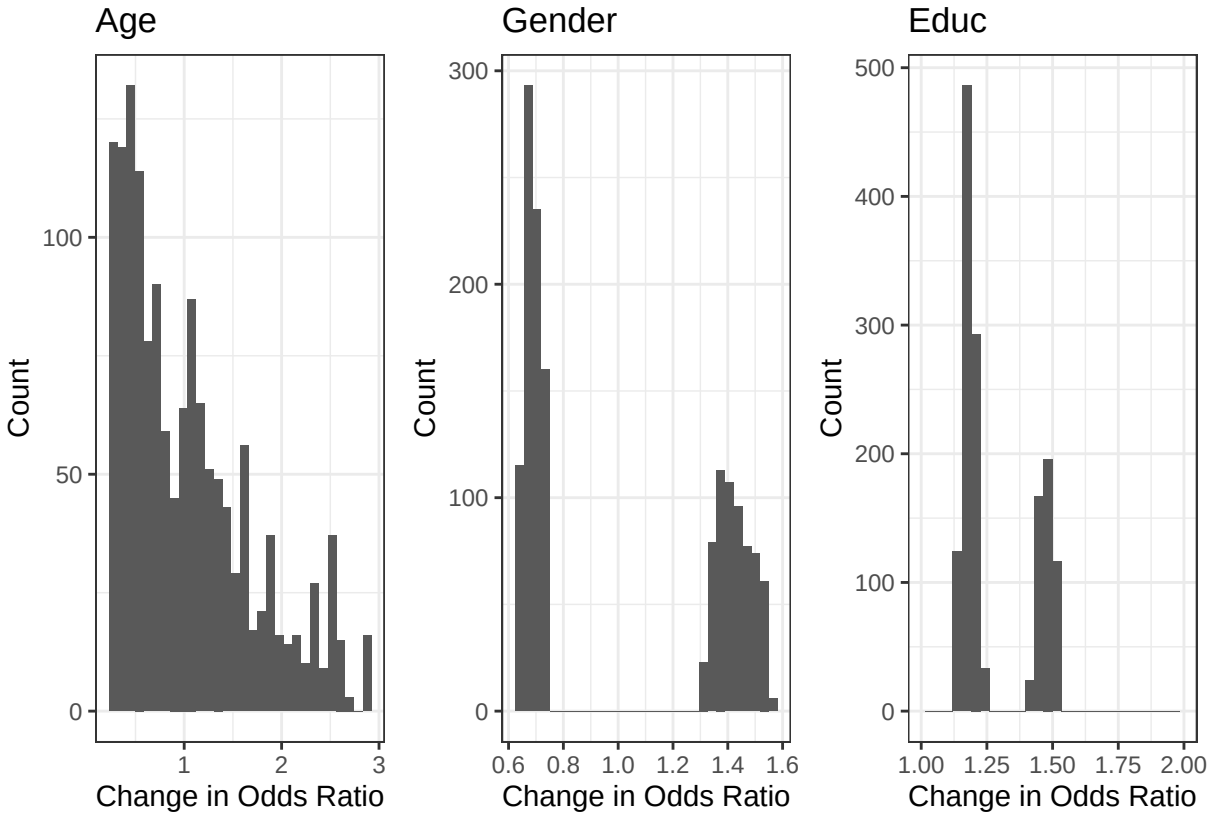
```
## [1] "Summary information about treatment and covariates"
```

light_drinking	age	gender	educ
Min. :0.00000	Min. :20.00	Min. :0.000	Min. :0.0000
1st Qu.:0.00000	1st Qu.:36.00	1st Qu.:0.000	1st Qu.:0.0000
Median :0.00000	Median :52.00	Median :1.000	Median :1.0000
Mean :0.07922	Mean :50.69	Mean :0.558	Mean :0.6505
3rd Qu.:0.00000	3rd Qu.:64.00	3rd Qu.:1.000	3rd Qu.:1.0000
Max. :1.00000	Max. :80.00	Max. :1.000	Max. :1.0000

Calibration

Compute propensity scores given observed covariates

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

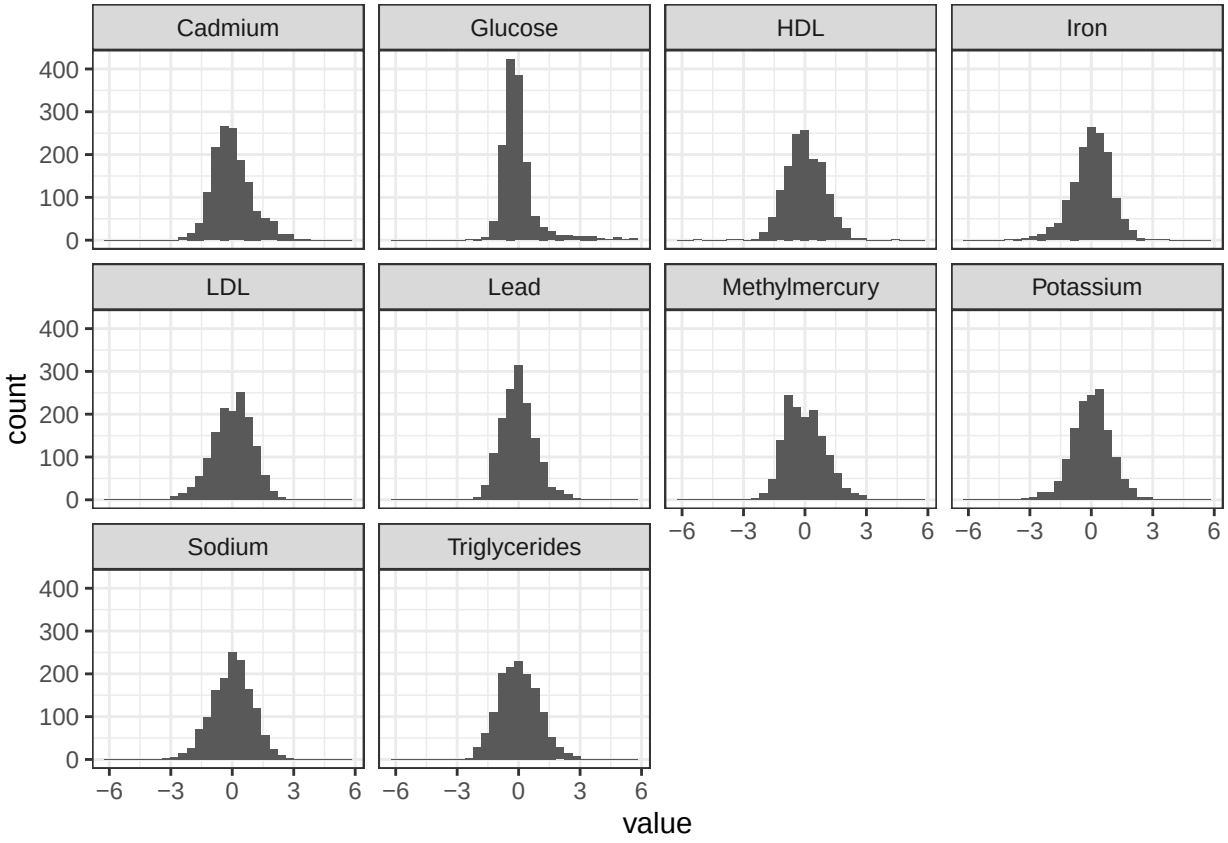


Compute benchmark Lambda

Compute benchmark lambda values by calibrating in R2 space and converting black (see paper Sections 5.1 and 5.2)

```
## Benchmark Lambdas:
## Lambda_age = 3.608499
## Lambda_gender = 2.098668
## Lambda_educ =1.632084
```

Histograms of outcomes



Run stan model and process results

Note: chunk option is set to FALSE by default (don't rerun inference every time we knit). Focus on rank 5 model for analysis

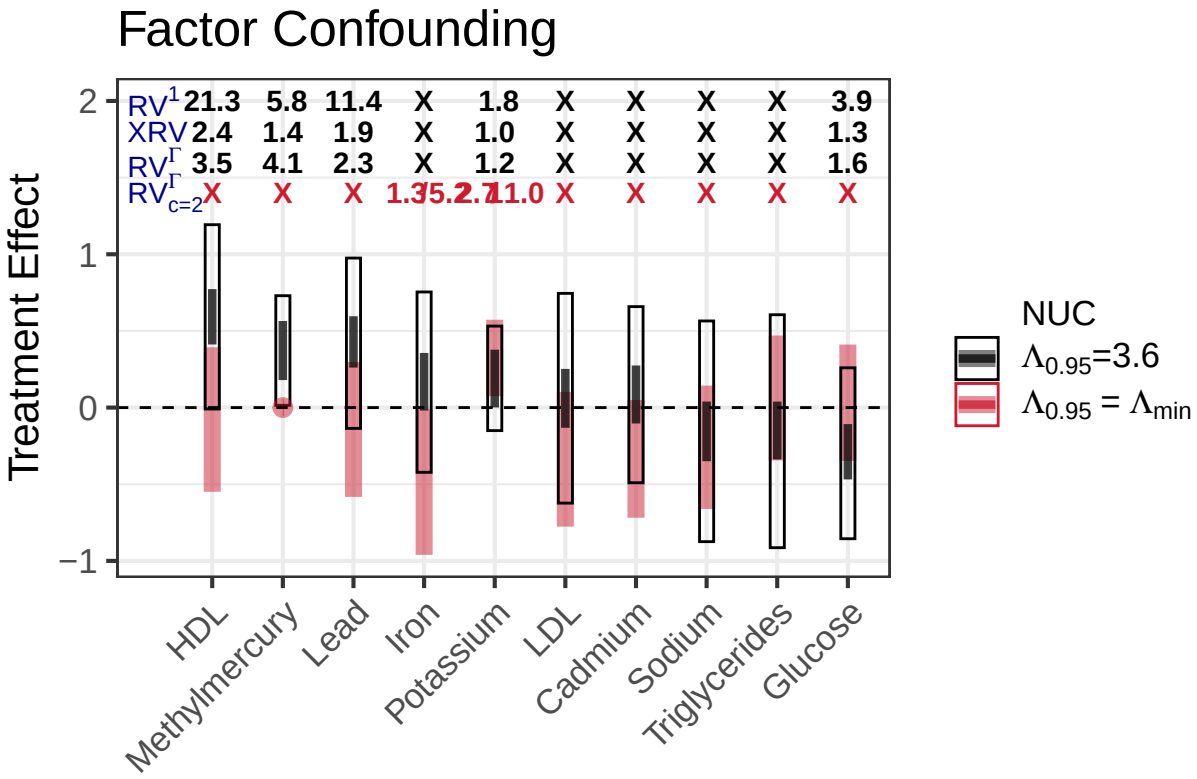
```
## Warning: `gather_()` was deprecated in tidyr 1.2.0.
## i Please use `gather()` instead.
## i The deprecated feature was likely used in the tidybayes package.
##   Please report the issue at <https://github.com/mjskay/tidybayes/issues/new>.
```

```
## `summarise()` has grouped output by 'J'. You can override using the `.groups`
## argument.
```

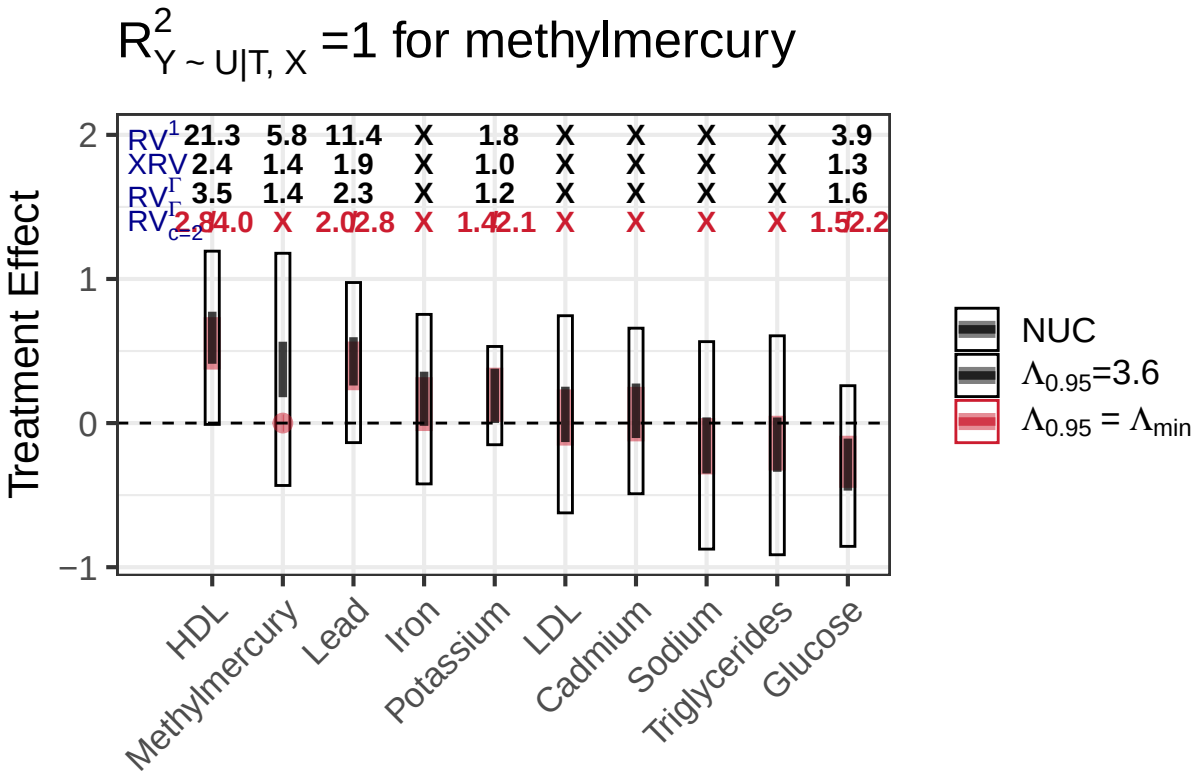
Posterior mean and intervals for causal effects under NUC

feature	q025	q975	J	K	sig
HDL	0.412	0.773	1	1	TRUE
Lead	0.261	0.596	1	7	TRUE
Methylmercury	0.179	0.565	1	2	TRUE
Potassium	0.001	0.378	1	4	TRUE
Iron	-0.019	0.356	1	6	FALSE
Cadmium	-0.104	0.274	1	10	FALSE
LDL	-0.132	0.252	1	8	FALSE
Triglycerides	-0.338	0.039	1	9	FALSE
Sodium	-0.350	0.039	1	5	FALSE
Glucose	-0.469	-0.108	1	3	TRUE

Figure 2 in paper



R2Y = 1 for Mercury (Appendix Figure 4)



Partial coefficients of determination (Appendix, Table 1)

	R2_Y~Age|-Age	R2_Y~Gender|-Gender	R2_Y~Educ|-Educ	R2_Y~U|X
HDL	0.013	0.12	0.01	0.501
Methylmercury	0.019	NS	0.018	0.066
Glucose	0.092	0.014	0.005	0.192
Potassium	0.005	0.045	NS	0.05
Sodium	0.008	0.003	NS	0.59
Iron	0.007	0.071	0.006	0.285
Lead	0.215	0.047	0.014	0.412
LDL	NS	NS	NS	0.46
Triglycerides	0.034	0.009	0.003	0.86
Cadmium	0.041	0.028	0.013	0.289

Model checking: Table 2 from Appendix

```
## Warning: Some Pareto k diagnostic values are slightly high. See help('pareto-k-diagnostic') for deta
```

	elpd_diff	se_diff
full_rank	0.00	0.00
rank6	-0.62	17.92
rank5	-4.11	18.17
rank4	-21.24	19.22
rank3	-45.93	20.98
rank2	-101.93	24.02
rank1	-223.64	29.68

Gamma heat map: Figure 3 (Appendix)

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.1622  0.5447  0.6124  0.6079  0.6786  0.9241
```