STATS 211 - Reanalysis of "Does Daylight Saving Save Electricity?"

```r
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 4.0.5
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.5     v purrr   0.3.4
## v tibble  3.1.6     v dplyr   1.0.7
## v tidyr   1.1.4     v stringr 1.4.0
## v readr   2.1.0     v forcats 0.5.1
```

```
## Warning: package 'ggplot2' was built under R version 4.0.5
```

```
## Warning: package 'tibble' was built under R version 4.0.5
```

```
## Warning: package 'tidyr' was built under R version 4.0.5
```

```
## Warning: package 'readr' was built under R version 4.0.5
```

```
## Warning: package 'dplyr' was built under R version 4.0.5
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

Import Data

Original Paper: https://meta-analysis.cz/dst/dst.pdf

Website + Data: http://meta-analysis.cz/dst

```r
dst_df <- read.csv("dst.csv")
```

```r
dst_df %>% head()
```

```
##                      LABEL IDSTUDY IDAUTHOR COUNTRY COUNTRYA ESTIMATE
## 1             ADEME (1995)       1        1  France   France    -0.12
## 2             ADEME (2010)       2        1  France   France    -0.37
## 3 Ahuja & SenGupta (2012)       3        2   India    India    -0.29
## 4 Ahuja & SenGupta (2012)       3        2   India    India    -0.30
## 5     Ahuja et al. (2007)       4        2   India    India    -0.30
```

```
## 6           Binder (1976)       5        3     USA       USA    -1.00
##            SE TSTAT PRECISION PCC PCCSE  N  K DF REGRESSION SIMULATION RESIDENT
## 1          NA    NA        NA  NA    NA NA NA NA          0          0        0
## 2          NA    NA        NA  NA    NA NA NA NA          0          1        0
## 3 0.002959184   -98  337.9310  NA    NA NA NA NA          0          1        0
## 4 0.003061224   -98  326.6667  NA    NA NA NA NA          0          1        0
## 5          NA    NA        NA  NA    NA NA NA NA          0          1        0
## 6          NA    NA        NA  NA    NA NA NA NA          1          0        0
##   LIGHT USA PUBYEAR DSTYEAR PERIOD HOUR DAY MONTH DID LOG MAIN CITATIONS
## 1     1   0    1995    1995     NA    1   0     0   0   0    1         2
## 2     0   0    2010    2008      1    1   0     0   0   0    0         2
## 3     0   0    2012    2008      1    1   0     0   0   0    1         3
## 4     0   0    2012    2008      1    1   0     0   0   0    1         3
## 5     0   0    2007    2004      1    1   0     0   0   0    1         3
## 6     0   1    1975    1974      1    1   0     0   1   0    1         1
##   JOURNAL IMPACT WEIGHT LATITUDE DAYLIGHT EUROPE
## 1       0  0.000    1.0   46.000 15.75000      1
## 2       0  0.000    1.0   46.000 15.75000      1
## 3       1  0.027    0.5   20.000 13.33333      0
## 4       1  0.027    0.5   20.000 13.33333      0
## 5       1  0.027    1.0   20.000 13.33333      0
## 6       0  0.000    1.0   37.751 14.76667      0
##
## 1                              ADEME (1995): "Internal ADEME (French Environment and Energy Manageme
## 2                                                                                                  ADI
## 3
## 4
## 5
## 6 Binder, R. H. (1976): Testimony of Robert H. Binder, assistant secretary for policy, plans and inte
##    X
## 1 NA
## 2 NA
## 3 NA
## 4 NA
## 5 NA
## 6 NA
```

Add in some missing variables for later use

```
dst_df <- dst_df %>%
  mutate(OTHER_ANALYSIS = --!(REGRESSION | SIMULATION),
         COMMERCIAL = 1 - RESIDENT,
         UNREFEREED = 1 - JOURNAL,
         WITH_SE = --!is.na(SE),
         ALL = 1)
```

Recreate Initial BoxPlots

Figure 2: It looks like they are just plotting the variation of estimates reported in each study

```
dst_df %>%
  ggplot(aes(x=LABEL, y=ESTIMATE)) +
  geom_boxplot() +
  coord_flip() +
  geom_hline(yintercept=0, linetype='longdash')
```

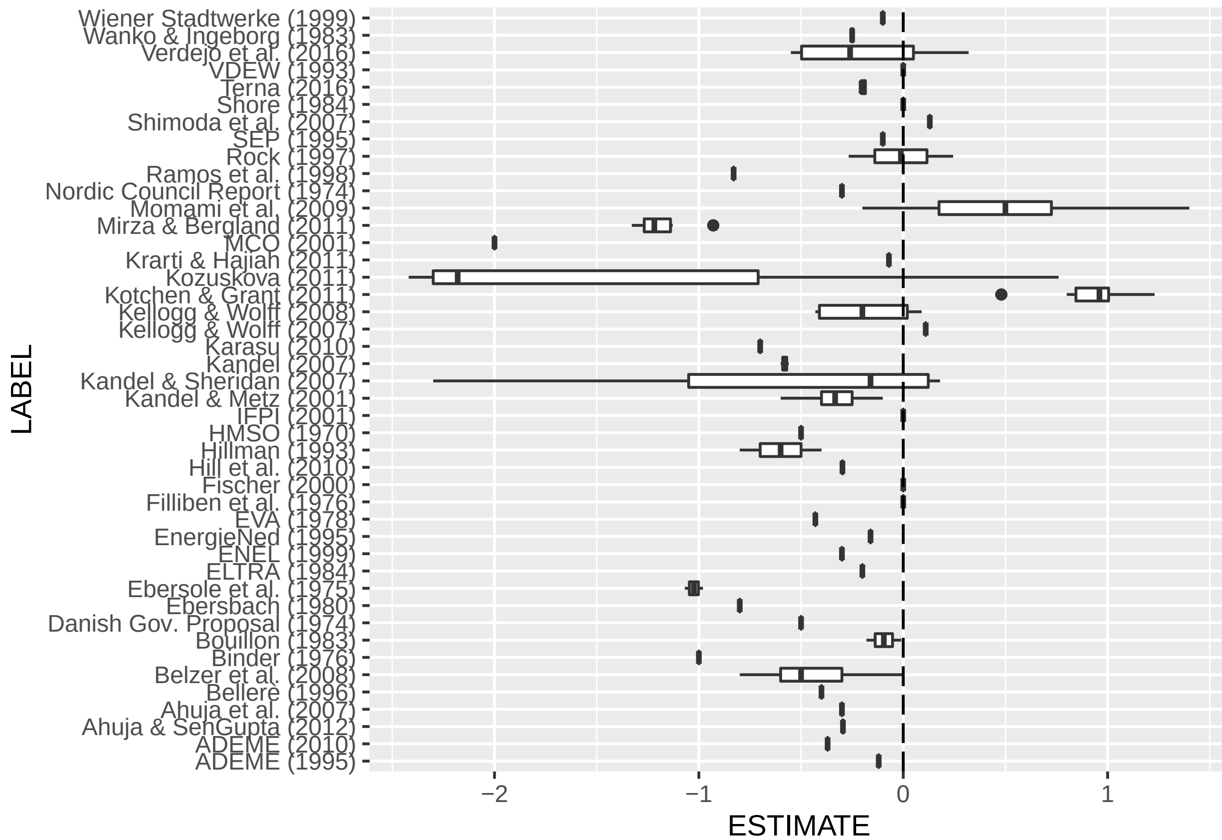
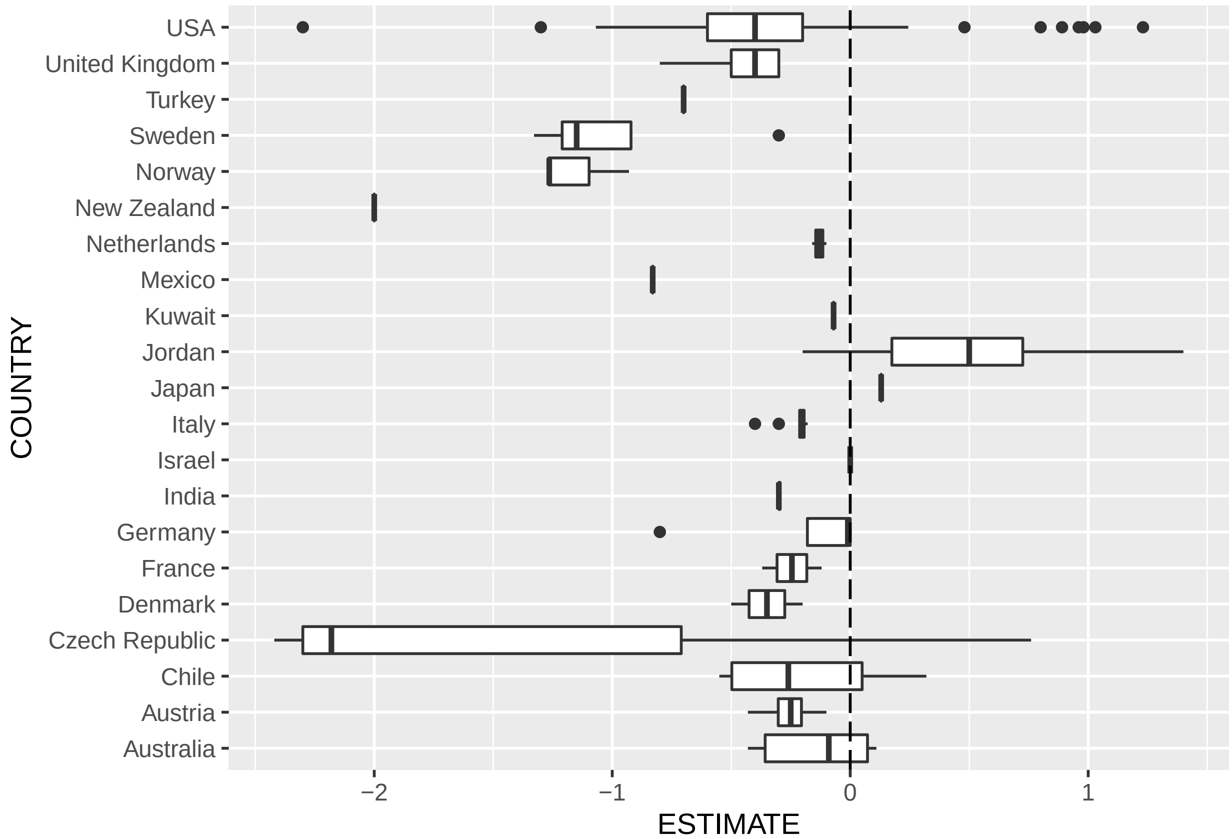


Figure 3

```
dst_df %>%
  ggplot(aes(x=COUNTRY, y=ESTIMATE)) +
  geom_boxplot() +
  coord_flip() +
  geom_hline(yintercept=0, linetype='longdash')
```



Main Estimate

Table 2: Looks like we're just doing mean of estimates with different slices

```
columns <- c(quo(HOUR), quo(DAY), quo(MAIN), quo(EUROPE), quo(USA), quo(REGRESSION), quo(SIMULATION), qu

report_df <- data.frame(
  subgroup = character(),
  n = integer(),
  mean = double(),
  ci_lower = double(),
  ci_upper = double(),
  w_mean = double(),
  w_ci_lower = double(),
  w_ci_upper = double(),
  SD = double(),
  w_SD = double()
)

for (col in columns) {
  report_df <- bind_rows(
    report_df,
    dst_df %>%
      filter(!!col == 1) %>%
      summarize(
        n = n(),
        mean = mean(ESTIMATE),
```

```r
        SD = sqrt(mean((ESTIMATE-mean)^2) / n),
        ci_lower = mean - qt(.975, df=n) * SD,
        ci_upper = mean + qt(.975, df=n) * SD,
        w_mean = weighted.mean(ESTIMATE, WEIGHT),
        w_SD = sqrt(weighted.mean((ESTIMATE-w_mean)^2, wt=WEIGHT) / n),
        w_ci_lower = mean - qt(.975, df=n) * w_SD,
        w_ci_upper = mean + qt(.975, df=n) * w_SD
      ) %>%
      mutate(subgroup = quo_name(col)) %>%
      select(
        subgroup,
        n,
        mean,
        ci_lower,
        ci_upper,
        w_mean,
        w_ci_lower,
        w_ci_upper,
        SD,
        w_SD
      )
  )
}

report_df
```

```
##          subgroup   n        mean   ci_lower    ci_upper     w_mean w_ci_lower
## 1            HOUR 139 -0.36145022 -0.4281210 -0.29477942 -0.3351577 -0.4282667
## 2             DAY  15 -0.68733333 -1.1988199 -0.17584676 -0.6541667 -1.1991455
## 3            MAIN  67 -0.25044982 -0.4087584 -0.09214127 -0.3384027 -0.4102046
## 4          EUROPE  43 -0.47386879 -0.6489290 -0.29880862 -0.3861882 -0.6509936
## 5             USA  94 -0.34122577 -0.4408314 -0.24162012 -0.3069850 -0.4410779
## 6      REGRESSION 117 -0.39525852 -0.4950156 -0.29550143 -0.4183492 -0.4951052
## 7      SIMULATION  21 -0.24077778 -0.4035501 -0.07800543 -0.2591357 -0.4037632
## 8  OTHER_ANALYSIS  24 -0.12000000 -0.3781965  0.13819649 -0.3203947 -0.3916486
## 9        RESIDENT  17  0.21923529 -0.1198702  0.55834074 -0.1167692 -0.1609676
## 10     COMMERCIAL 145 -0.39936952 -0.4798724 -0.31886662 -0.3819114 -0.4799234
## 11          LIGHT   7 -0.33714286 -0.5914492 -0.08283655 -0.3036364 -0.5932063
## 12            DID  94 -0.40722637 -0.5197100 -0.29474273 -0.4487523 -0.5200310
## 13        JOURNAL  41 -0.02557049 -0.2468989  0.19575797 -0.1206594 -0.2489217
## 14     UNREFEREED 121 -0.43911727 -0.5171435 -0.36109100 -0.4463816 -0.5171545
## 15        WITH_SE 101 -0.40195656 -0.5170946 -0.28681854 -0.4107411 -0.5171076
## 16            ALL 162 -0.33445420 -0.4190549 -0.24985349 -0.3427427 -0.4190647
##     w_ci_upper         SD       w_SD
## 1  -0.29463378 0.03372018 0.03379385
## 2  -0.17552118 0.23997123 0.24012398
## 3  -0.09069504 0.07931257 0.08003713
## 4  -0.29674397 0.08680560 0.08782938
## 5  -0.24137360 0.05016594 0.05029010
## 6  -0.29541188 0.05037098 0.05041620
## 7  -0.07779237 0.07827047 0.07837292
## 8   0.15164856 0.12510135 0.13161914
## 9   0.59943822 0.16072753 0.18020671
```

```
## 10 -0.31881564 0.04073086 0.04075666
## 11 -0.08107944 0.10754618 0.10828926
## 12 -0.29442171 0.05665188 0.05681356
## 13  0.19778069 0.10959345 0.11059502
## 14 -0.36108004 0.03941191 0.03941744
## 15 -0.28680548 0.05804114 0.05804773
## 16 -0.24984372 0.04284196 0.04284691
```

Notice how the weights are calculated by just dividing the number of observations reported per study instead of by sample size. This means this is NOT a meta analysis.

```r
dst_df %>%
  filter(!is.na(N)) %>%
  select(LABEL, N, WEIGHT)
```

```
##                       LABEL     N    WEIGHT
## 1      Belzer et al. (2008)  73920 0.0149254
## 2      Belzer et al. (2008)  41760 0.0149254
## 3      Belzer et al. (2008)   2112 0.0149254
## 4      Belzer et al. (2008)   2112 0.0149254
## 5      Belzer et al. (2008)   2112 0.0149254
## 6      Belzer et al. (2008)   2112 0.0149254
## 7      Belzer et al. (2008)   2112 0.0149254
## 8      Belzer et al. (2008)   2112 0.0149254
## 9      Belzer et al. (2008)   2112 0.0149254
## 10     Belzer et al. (2008)   2112 0.0149254
## 11     Belzer et al. (2008)   2112 0.0149254
## 12     Belzer et al. (2008)   2112 0.0149254
## 13     Belzer et al. (2008)   2112 0.0149254
## 14     Belzer et al. (2008)   2112 0.0149254
## 15     Belzer et al. (2008)   2112 0.0149254
## 16     Belzer et al. (2008)   2112 0.0149254
## 17     Belzer et al. (2008)   2112 0.0149254
## 18     Belzer et al. (2008)   2112 0.0149254
## 19     Belzer et al. (2008)   2112 0.0149254
## 20     Belzer et al. (2008)   2112 0.0149254
## 21     Belzer et al. (2008)   2112 0.0149254
## 22     Belzer et al. (2008)   2112 0.0149254
## 23     Belzer et al. (2008)   2112 0.0149254
## 24     Belzer et al. (2008)   2112 0.0149254
## 25     Belzer et al. (2008)   2112 0.0149254
## 26     Belzer et al. (2008)   2112 0.0149254
## 27     Belzer et al. (2008)   2112 0.0149254
## 28     Belzer et al. (2008)   2112 0.0149254
## 29     Belzer et al. (2008)   2112 0.0149254
## 30     Belzer et al. (2008)   2112 0.0149254
## 31     Belzer et al. (2008)   2112 0.0149254
## 32     Belzer et al. (2008)   2112 0.0149254
## 33     Belzer et al. (2008)   2112 0.0149254
## 34     Belzer et al. (2008)   2112 0.0149254
## 35     Belzer et al. (2008)   2112 0.0149254
## 36     Belzer et al. (2008)   2112 0.0149254
## 37     Belzer et al. (2008)   2112 0.0149254
```

```
## 38     Belzer et al. (2008)    1440 0.0149254
## 39     Belzer et al. (2008)    1440 0.0149254
## 40     Belzer et al. (2008)    1440 0.0149254
## 41     Belzer et al. (2008)    1440 0.0149254
## 42     Belzer et al. (2008)    1440 0.0149254
## 43     Belzer et al. (2008)    1440 0.0149254
## 44     Belzer et al. (2008)    1440 0.0149254
## 45     Belzer et al. (2008)    1440 0.0149254
## 46     Belzer et al. (2008)    1440 0.0149254
## 47     Belzer et al. (2008)    1440 0.0149254
## 48     Belzer et al. (2008)    1440 0.0149254
## 49     Belzer et al. (2008)    1440 0.0149254
## 50     Belzer et al. (2008)    1440 0.0149254
## 51     Belzer et al. (2008)    1440 0.0149254
## 52     Belzer et al. (2008)    1440 0.0149254
## 53     Belzer et al. (2008)    1440 0.0149254
## 54     Belzer et al. (2008)    1440 0.0149254
## 55     Belzer et al. (2008)    1440 0.0149254
## 56     Belzer et al. (2008)    1440 0.0149254
## 57     Belzer et al. (2008)    1440 0.0149254
## 58     Belzer et al. (2008)    1440 0.0149254
## 59     Belzer et al. (2008)    1440 0.0149254
## 60     Belzer et al. (2008)    1440 0.0149254
## 61     Belzer et al. (2008)    1440 0.0149254
## 62     Belzer et al. (2008)    1440 0.0149254
## 63     Belzer et al. (2008)    1440 0.0149254
## 64     Belzer et al. (2008)    1440 0.0149254
## 65     Belzer et al. (2008)    1440 0.0149254
## 66     Belzer et al. (2008)    1440 0.0149254
## 67     Belzer et al. (2008)    1440 0.0149254
## 68   Filliben et al. (1976)      56 1.0000000
## 69       Hill et al. (2010)   26829 0.5000000
## 70       Hill et al. (2010)   26829 0.5000000
## 71 Kandel & Sheridan (2007)     632 0.1666667
## 72 Kandel & Sheridan (2007)     632 0.1666667
## 73 Kandel & Sheridan (2007)     632 0.1666667
## 74 Kandel & Sheridan (2007)      62 0.1666667
## 75 Kandel & Sheridan (2007)      93 0.1666667
## 76 Kandel & Sheridan (2007)     186 0.1666667
## 77   Kellogg & Wolff (2007)   12960 1.0000000
## 78   Kellogg & Wolff (2008)   16224 0.2000000
## 79   Kellogg & Wolff (2008)   12960 0.2000000
## 80   Kellogg & Wolff (2008)   12960 0.2000000
## 81   Kotchen & Grant (2011) 3685287 0.1428571
## 82   Kotchen & Grant (2011) 3685287 0.1428571
## 83   Kotchen & Grant (2011) 3685287 0.1428571
## 84   Kotchen & Grant (2011) 3685287 0.1428571
## 85   Kotchen & Grant (2011)  580888 0.1428571
## 86   Kotchen & Grant (2011)  603253 0.1428571
## 87         Kozuskova (2011)     366 0.3333333
## 88         Kozuskova (2011)     366 0.3333333
## 89         Kozuskova (2011)     186 0.3333333
## 90  Mirza & Bergland (2011)   57696 0.1666667
## 91  Mirza & Bergland (2011)   57696 0.1666667
```

```
## 92  Mirza & Bergland (2011)   57696 0.1666667
## 93  Mirza & Bergland (2011)   57696 0.1666667
## 94  Mirza & Bergland (2011)   57696 0.1666667
## 95  Mirza & Bergland (2011)   57696 0.1666667
## 96              Rock (1997)     234 0.5000000
## 97              Rock (1997)     234 0.5000000
## 98    Shimoda et al. (2007) 1044000 1.0000000
```

A lot of estimates also missing SE.

```
dst_df %>% summarize(num_rows = n(), pct_has_se= sum(WITH_SE)/n())
```

```
##   num_rows pct_has_se
## 1      162  0.6234568
```

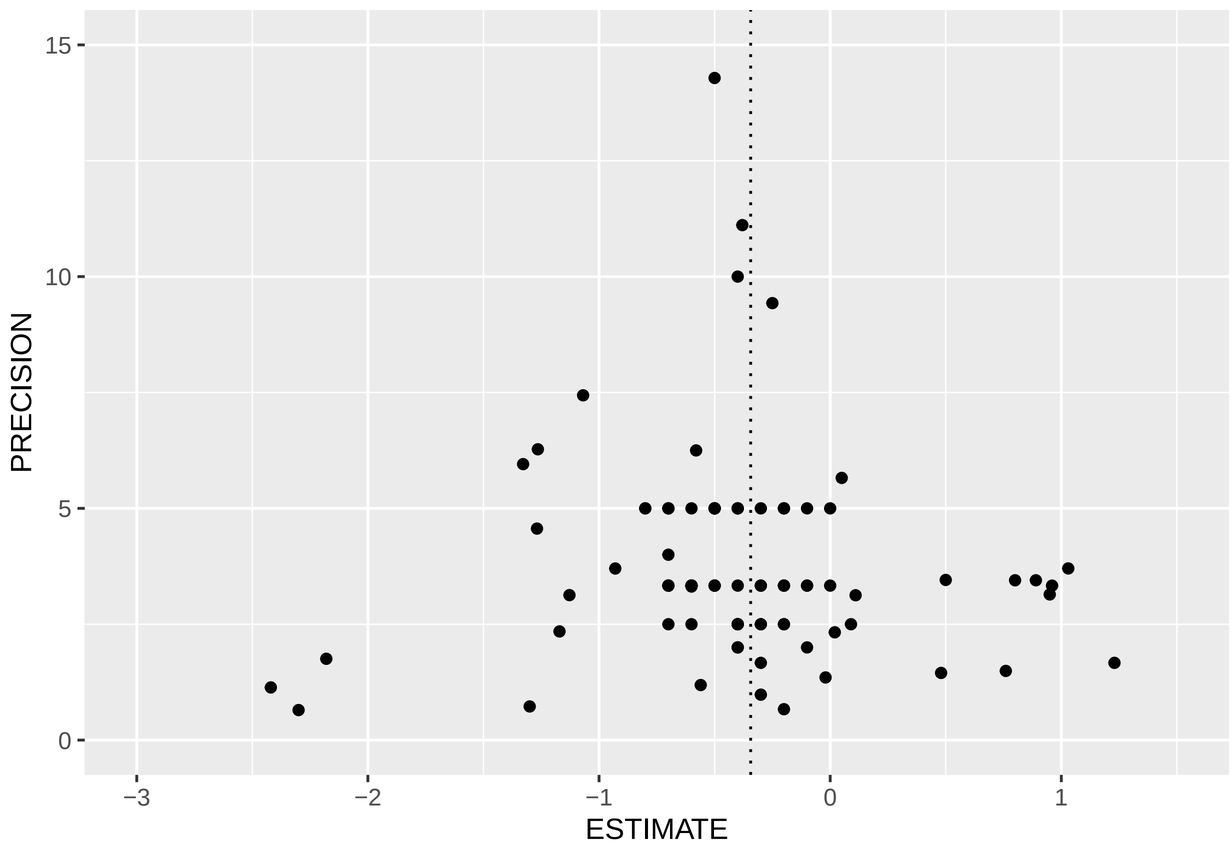
```
dst_df %>%
  group_by(IDSTUDY) %>%
  summarize(at_least_one_se = max(WITH_SE),
            all_se = min(WITH_SE)) %>%
  ungroup() %>%
  summarize(
    num_unique_studies = n(),
    at_least_one_se = sum(at_least_one_se),
    all_se = sum(all_se),
    pct_at_least_one_se = sum(at_least_one_se) / n(),
    pct_all_se = sum(all_se) / n()) %>%
  select(
    num_unique_studies, pct_at_least_one_se, pct_all_se
  )
```

```
## # A tibble: 1 x 3
##   num_unique_studies pct_at_least_one_se pct_all_se
##                <int>               <dbl>      <dbl>
## 1                 44               0.295      0.114
```

Publication Bias

Figure 5: Funnel plot recreation. Filters were applied in original.

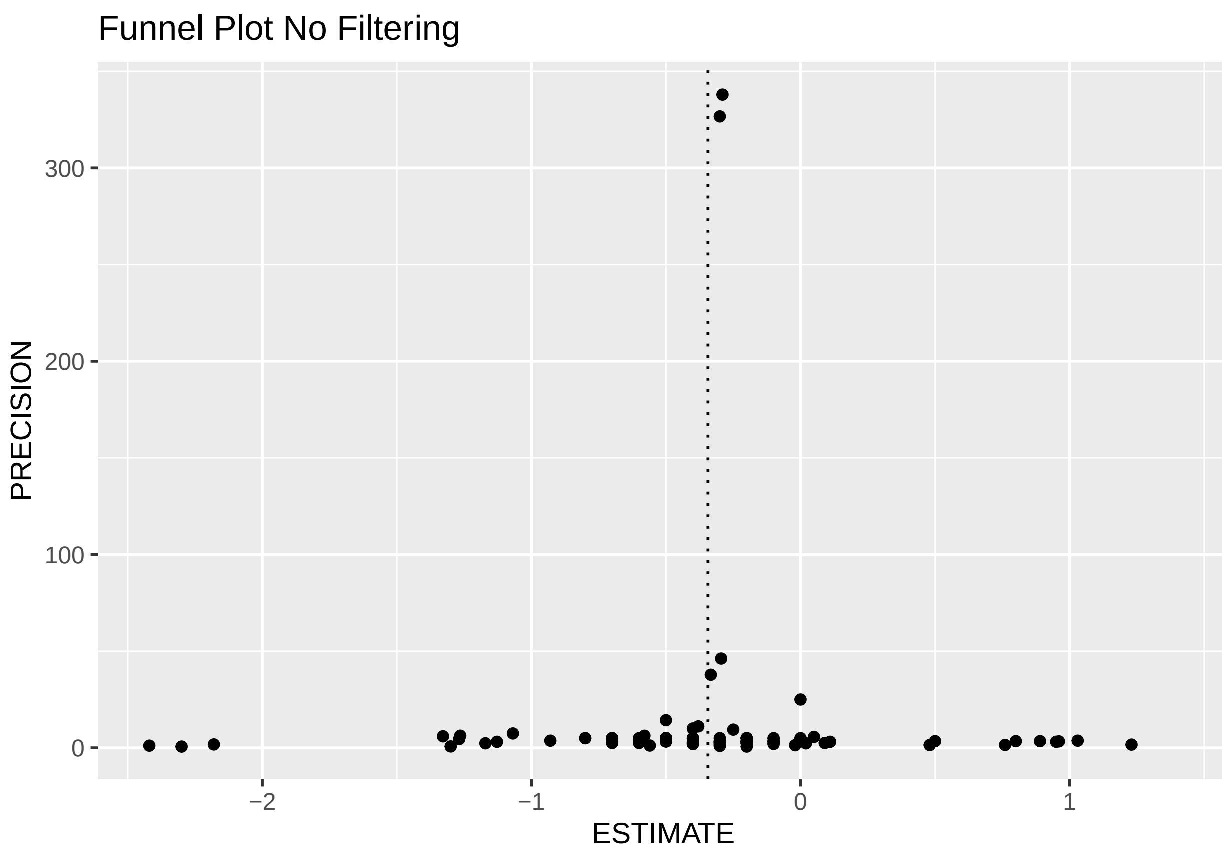
```
dst_df %>%
  filter(ESTIMATE > -5 & PRECISION < 15) %>%
  ggplot(aes(x=ESTIMATE, y=PRECISION)) +
  geom_point() +
  xlim(-3, 1.5) +
  ylim(0, 15) +
  geom_vline(xintercept=-0.344, linetype='dotted')
```



Remove filters

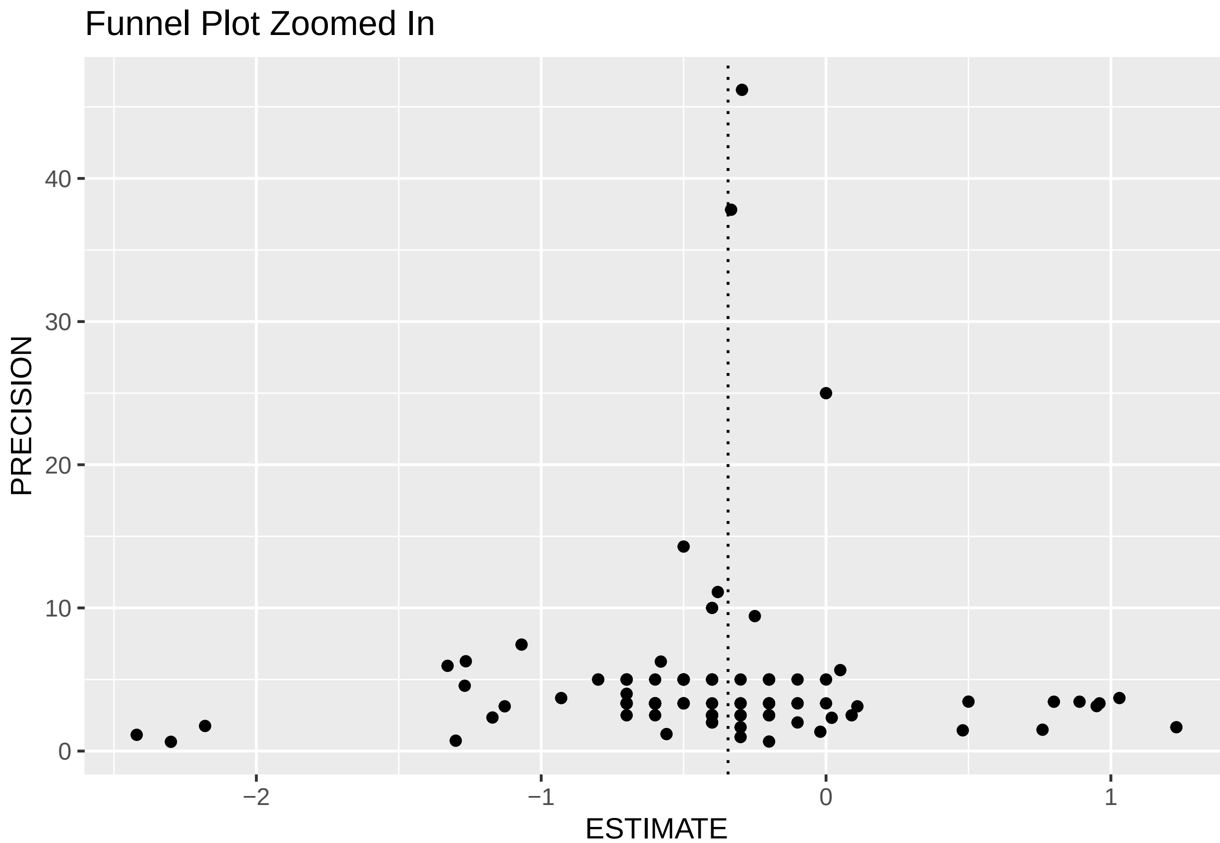
```
dst_df %>%
  ggplot(aes(x=ESTIMATE, y=PRECISION)) +
  geom_point() +
  geom_vline(xintercept=-0.344, linetype='dotted') +
  ggtitle('Funnel Plot No Filtering')
```

```
## Warning: Removed 61 rows containing missing values (geom_point).
```



Zoom in a little bit

```
dst_df %>%
  filter(PRECISION < 300) %>%
  ggplot(aes(x=ESTIMATE, y=PRECISION)) +
  geom_point() +
  geom_vline(xintercept=-0.344, linetype='dotted') +
  ggtitle('Funnel Plot Zoomed In')
```



A lot of precision numbers were missing, so the funnel plot doesn't seem representative. Let's calculate exactly how much was missing.

```
dst_df %>%
  summarize(
    num_rows = n(),
    missing_precision = sum(is.na(PRECISION)),
    pct_missing_precision = missing_precision / num_rows
  ) %>%
  select(num_rows, pct_missing_precision)
```

```
##   num_rows pct_missing_precision
## 1      162             0.3765432
```

Over 37% of rows are missing precision.

```
dst_df %>%
  group_by(IDSTUDY) %>%
  summarize(
    has_missing_precision = min(is.na(PRECISION))
  ) %>%
  ungroup() %>%
  summarize(
    studies_missing_all_precision = sum(has_missing_precision),
    pct_studies_missing_precision = sum(has_missing_precision)/n()
  )
```

```
## # A tibble: 1 x 2
##   studies_missing_all_precision pct_studies_missing_precision
##                           <int>                         <dbl>
## 1                            31                         0.705
```

Our funnel plot only represents ~30% of unique studies

Let's add in the missing study estimates at -10 just to see if they seem to skew in a certain way. It looks like a lot of estimates are closer to 0.

```
dst_df %>%
  mutate(
    PRECISION_FILLED = replace_na(PRECISION,-10),
    missing_precision = is.na(PRECISION)
  ) %>%
  filter(PRECISION_FILLED < 300) %>%
  ggplot(aes(x = ESTIMATE, y = PRECISION_FILLED)) +
  geom_point(aes(color = missing_precision)) +
  geom_vline(xintercept = -0.344, linetype = 'dotted') +
  ggtitle('Funnel Plot Adding Missing Studies')
```

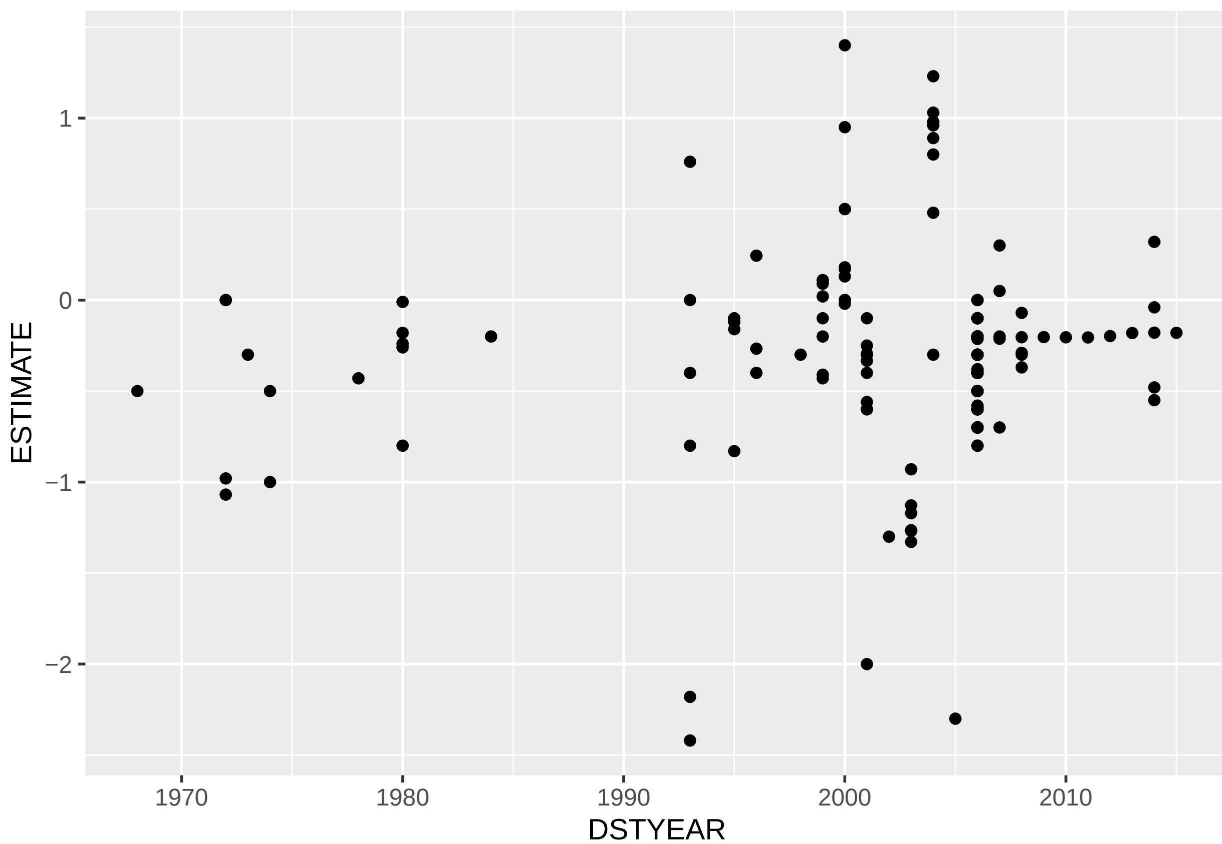
Table 3: It looks like a regular least squares gets similar estimates. The report did weighted least squares.

```
se_reg <- lm("ESTIMATE ~ SE", data=dst_df)
se_reg %>% summary()
```

```
##
```

```
## Call:
## lm(formula = "ESTIMATE ~ SE", data = dst_df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.79243 -0.21452 -0.05626  0.21070  1.74070
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.26027    0.09373  -2.777  0.00657 **
## SE          -0.41738    0.21790  -1.915  0.05832 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5785 on 99 degrees of freedom
##   (61 observations deleted due to missingness)
## Multiple R-squared:  0.03574,    Adjusted R-squared:  0.026
## F-statistic: 3.669 on 1 and 99 DF,  p-value: 0.05832
```

```
dst_df %>%
  # filter(ESTIMATE > -5 & PRECISION < 15) %>%
  ggplot(aes(x=DSTYEAR, y=ESTIMATE)) +
  geom_point()
```



Meta-Regression

```r
dst_df %>% str()
```

```
## 'data.frame':    162 obs. of  42 variables:
##  $ LABEL          : chr  "ADEME (1995)" "ADEME (2010)" "Ahuja & SenGupta (2012)" "Ahuja & SenGupta (20
##  $ IDSTUDY        : int  1 2 3 3 4 5 6 7 7 7 ...
##  $ IDAUTHOR       : int  1 1 2 2 2 3 4 5 5 5 ...
##  $ COUNTRY        : chr  "France" "France" "India" "India" ...
##  $ COUNTRYA       : chr  "France" "France" "India" "India" ...
##  $ ESTIMATE       : num  -0.12 -0.37 -0.29 -0.3 -0.3 -1 -0.4 -0.5 -0.38 -0.6 ...
##  $ SE             : num  NA NA 0.00296 0.00306 NA ...
##  $ TSTAT          : num  NA NA -98 -98 NA ...
##  $ PRECISION      : num  NA NA 338 327 NA ...
##  $ PCC            : num  NA NA NA NA NA ...
##  $ PCCSE          : logi  NA NA NA NA NA NA ...
##  $ N              : int  NA NA NA NA NA NA NA 73920 41760 2112 ...
##  $ K              : int  NA NA NA NA NA NA NA 450 450 450 ...
##  $ DF             : int  NA NA NA NA NA NA NA 73470 41310 1662 ...
##  $ REGRESSION     : int  0 0 0 0 0 1 0 1 1 1 ...
##  $ SIMULATION     : int  0 1 1 1 1 0 0 0 0 0 ...
##  $ RESIDENT       : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ LIGHT          : int  1 0 0 0 0 0 0 0 0 0 ...
##  $ USA            : int  0 0 0 0 0 1 0 1 1 1 ...
##  $ PUBYEAR        : int  1995 2010 2012 2012 2007 1975 1996 2008 2008 2008 ...
##  $ DSTYEAR        : int  1995 2008 2008 2008 2004 1974 1996 2006 2006 2006 ...
##  $ PERIOD         : int  NA 1 1 1 1 1 1 2 2 2 ...
##  $ HOUR           : int  1 1 1 1 1 1 0 1 1 1 ...
##  $ DAY            : int  0 0 0 0 0 0 1 0 0 0 ...
##  $ MONTH          : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ DID            : int  0 0 0 0 0 1 0 1 1 1 ...
##  $ LOG            : int  0 0 0 0 0 0 0 1 1 1 ...
##  $ MAIN           : int  1 0 1 1 1 1 1 1 1 0 ...
##  $ CITATIONS      : int  2 2 3 3 3 1 1 7 7 7 ...
##  $ JOURNAL        : int  0 0 1 1 1 0 0 0 0 0 ...
##  $ IMPACT         : num  0 0 0.027 0.027 0.027 0 0 0 0 0 ...
##  $ WEIGHT         : num  1 1 0.5 0.5 1 ...
##  $ LATITUDE       : num  46 46 20 20 20 ...
##  $ DAYLIGHT       : num  15.8 15.8 13.3 13.3 13.3 ...
##  $ EUROPE         : int  1 1 0 0 0 0 1 0 0 0 ...
##  $ REFERENCE      : chr  "ADEME (1995): \"Internal ADEME (French Environment and Energy Management Age
##  $ X              : logi  NA NA NA NA NA NA ...
##  $ OTHER_ANALYSIS: int  1 0 0 0 0 0 1 0 0 0 ...
##  $ COMMERCIAL     : num  1 1 1 1 1 1 1 1 1 1 ...
##  $ UNREFEREED     : num  1 1 0 0 0 1 1 1 1 1 ...
##  $ WITH_SE        : int  0 0 1 1 0 0 0 1 1 1 ...
##  $ ALL            : num  1 1 1 1 1 1 1 1 1 1 ...
```

Table 5: Running OLS on the relevant variables yields different estimates, but the same significant variables identified.

```r
meta_reg <- lm("ESTIMATE ~ PERIOD + MAIN + DAY + DAYLIGHT + USA + REGRESSION + SIMULATION + DID + RESID
meta_reg %>% summary()
```

```
##
```

```
## Call:
## lm(formula = "ESTIMATE ~ PERIOD + MAIN + DAY + DAYLIGHT + USA + REGRESSION + SIMULATION + DID + RESI
##     data = dst_df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.87613 -0.14628  0.00538  0.17558  1.99593
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -11.185175  10.566302  -1.059 0.291658
## PERIOD        0.006633   0.034134   0.194 0.846219
## MAIN          0.064556   0.106773   0.605 0.546442
## DAY          -0.464420   0.164772  -2.819 0.005539 **
## DAYLIGHT     -0.139467   0.042242  -3.302 0.001226 **
## USA           0.192468   0.140314   1.372 0.172399
## REGRESSION   -0.146962   0.160943  -0.913 0.362777
## SIMULATION   -0.564721   0.162787  -3.469 0.000699 ***
## DID          -0.431923   0.123523  -3.497 0.000635 ***
## RESIDENT      0.161634   0.179872   0.899 0.370441
## LIGHT         0.342888   0.235757   1.454 0.148120
## PUBYEAR       0.006544   0.005294   1.236 0.218529
## JOURNAL       0.206561   0.166689   1.239 0.217388
## IMPACT        0.597216   0.213450   2.798 0.005885 **
## CITATIONS     0.003311   0.004212   0.786 0.433189
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4287 on 137 degrees of freedom
##   (10 observations deleted due to missingness)
## Multiple R-squared:  0.4711, Adjusted R-squared:  0.4171
## F-statistic: 8.718 on 14 and 137 DF,  p-value: 2.614e-13
```

BMS

I could not recreate the bayesian model averaging. I tried copying their excel directly from clipboard but it doesn't work since the first column is not the estimate (which the bms package requires). I then copied from the estimate onwards and still was not able to run the package. Commenting out the code to show what I tried.

```
# library(BMS)
```

```
# datadaylight = read.table("clipboard-512", sep="\t", header=TRUE)
# datadaylight %>% head()
```

```
# bms_df <- dst_df %>%
#   select(ESTIMATE, HOUR, DAY, MAIN, EUROPE, USA, REGRESSION, SIMULATION, OTHER_ANALYSIS, RESIDENT, CO
```

```
# bms_df[complete.cases(bms_df),]
```

```
# daylight = bms(datadaylight, burn=1000000, iter=2000000, g="UIP", mprior="uniform", nmodel=5000, mcmc=
# daylight2 = bms(datadaylight, burn=1000000, iter=2000000, g="BRIC", mprior="random", nmodel=5000, mcm
# daylight3 = bms(datadaylight, burn=1000000,iter=2000000, g="hyper=BRIC", mprior="random", nmodel=5000
```