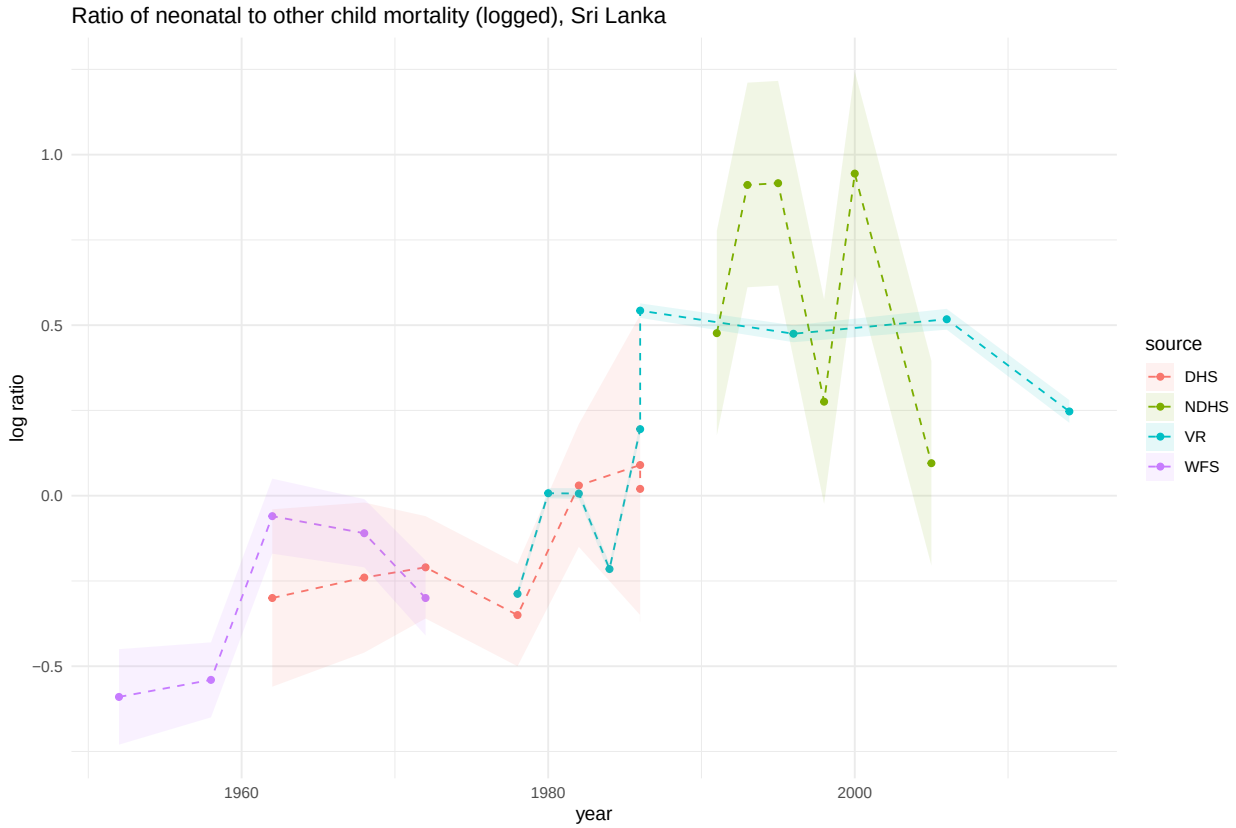
Week 10: Temporal data

Quynh Vu

2023-03-27

Child mortality in Sri Lanka

In this lab you will be fitting a couple of different models to the data about child mortality in Sri Lanka, which was used in the lecture. Here's the data and the plot from the lecture:



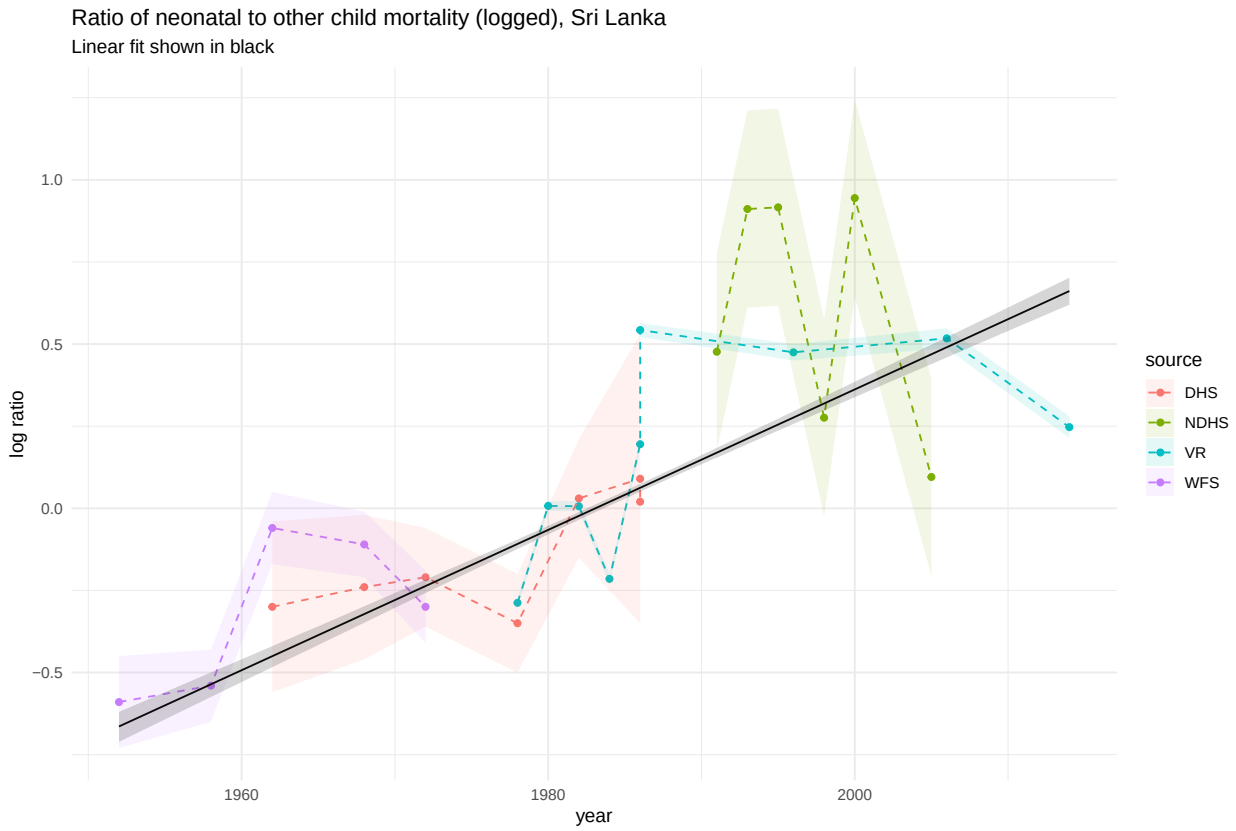
Fitting a linear model: Let's firstly fit a linear model in time to these data. Here's the code to do this:

Extract the results:

```
## # A tibble: 6 x 9
##       t .variable .value .lower .upper .width .point .interval  year
##   <int> <chr>      <dbl>  <dbl>  <dbl>  <dbl> <chr>  <chr>     <int>
## 1     1 mu        -0.664 -0.711 -0.620   0.95 median qi         1952
## 2     2 mu        -0.643 -0.688 -0.600   0.95 median qi         1953
## 3     3 mu        -0.621 -0.666 -0.580   0.95 median qi         1954
```

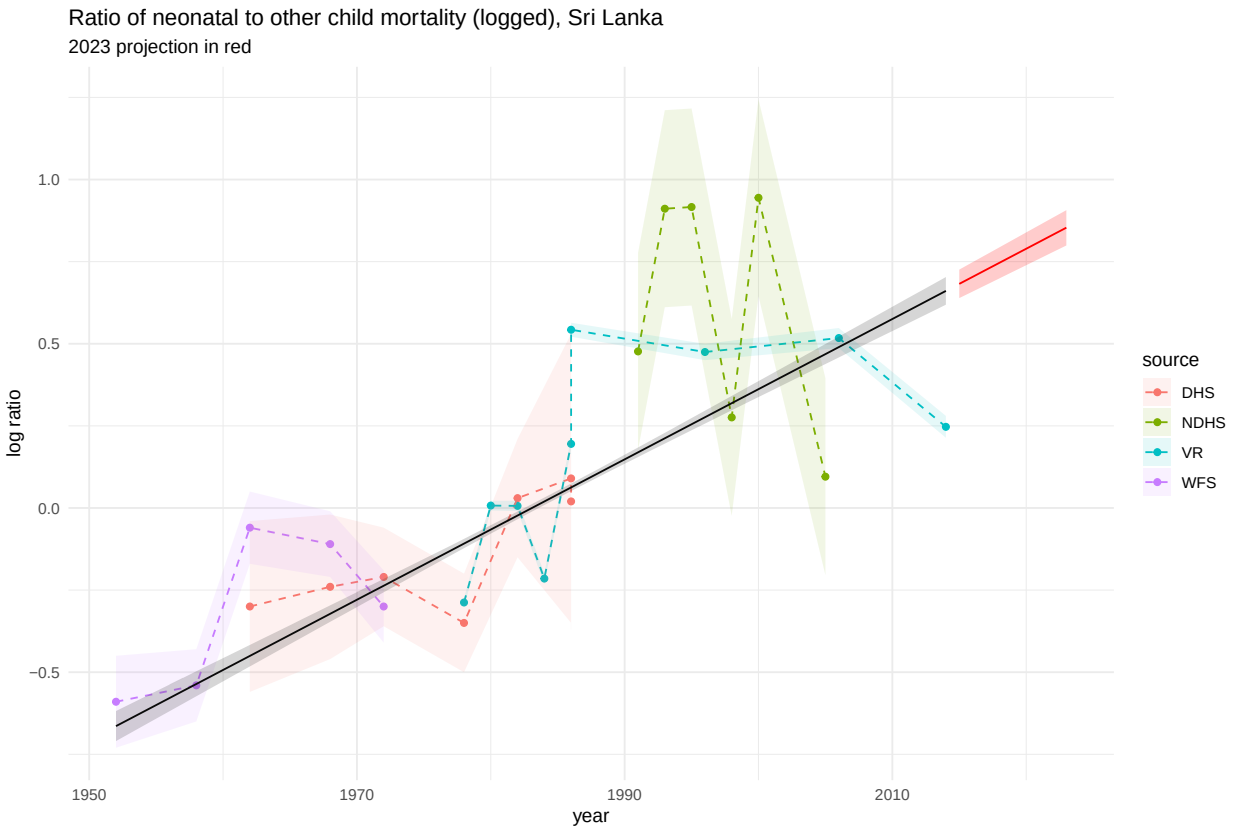
```
## 4     4 mu        -0.600 -0.643 -0.560   0.95 median qi         1955
## 5     5 mu        -0.579 -0.620 -0.540   0.95 median qi         1956
## 6     6 mu        -0.557 -0.597 -0.520   0.95 median qi         1957
```

Plot the results:



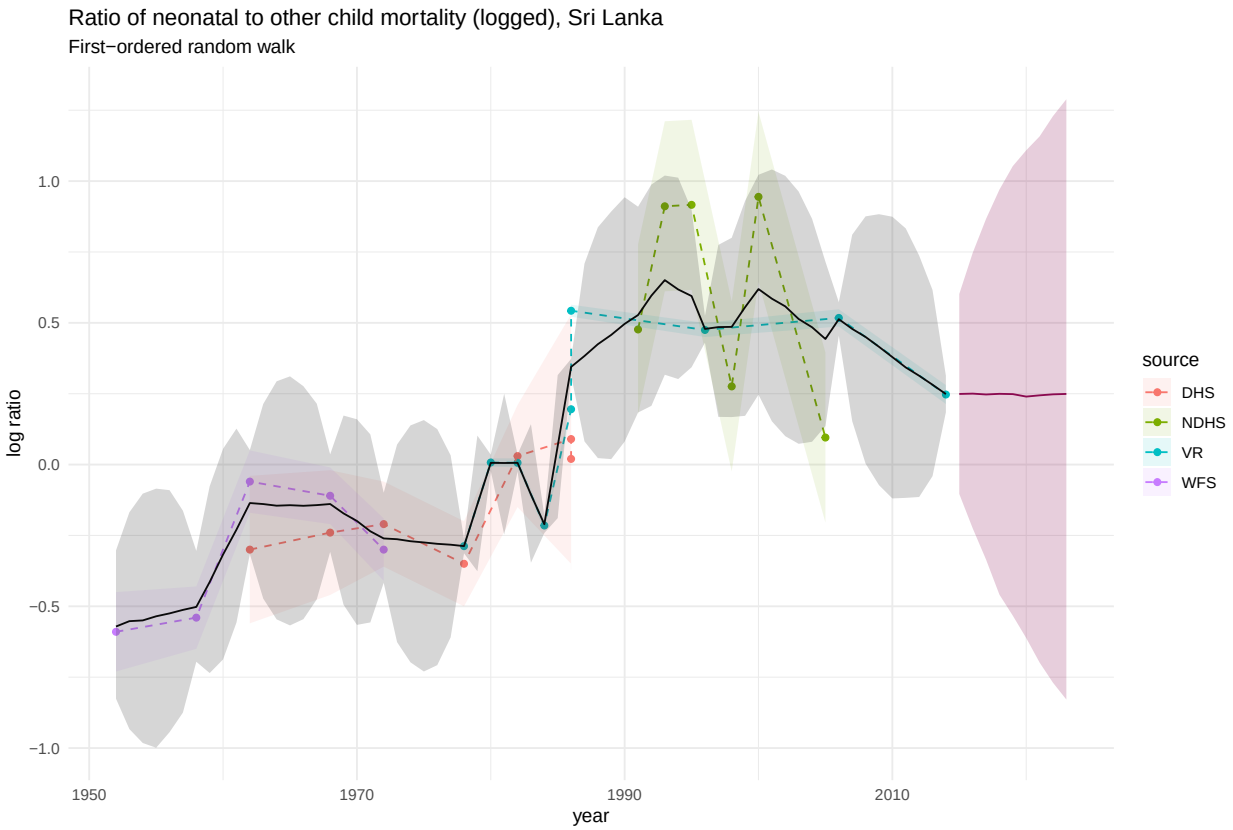
Question 1: Project the linear model above out to 2023 by adding a `generated quantities` block in Stan (do the projections based on the expected value μ). Plot the resulting projections on a graph similar to that above.

```
## # A tibble: 6 x 9
##       p .variable .value .lower .upper .width .point .interval  year
##   <int> <chr>      <dbl>  <dbl>  <dbl>  <dbl> <chr>  <chr>     <int>
## 1     1 mu_p       0.682  0.639  0.726   0.95 median qi         2015
## 2     2 mu_p       0.704  0.659  0.748   0.95 median qi         2016
## 3     3 mu_p       0.725  0.679  0.771   0.95 median qi         2017
## 4     4 mu_p       0.746  0.699  0.794   0.95 median qi         2018
## 5     5 mu_p       0.768  0.719  0.816   0.95 median qi         2019
## 6     6 mu_p       0.789  0.739  0.839   0.95 median qi         2020
```



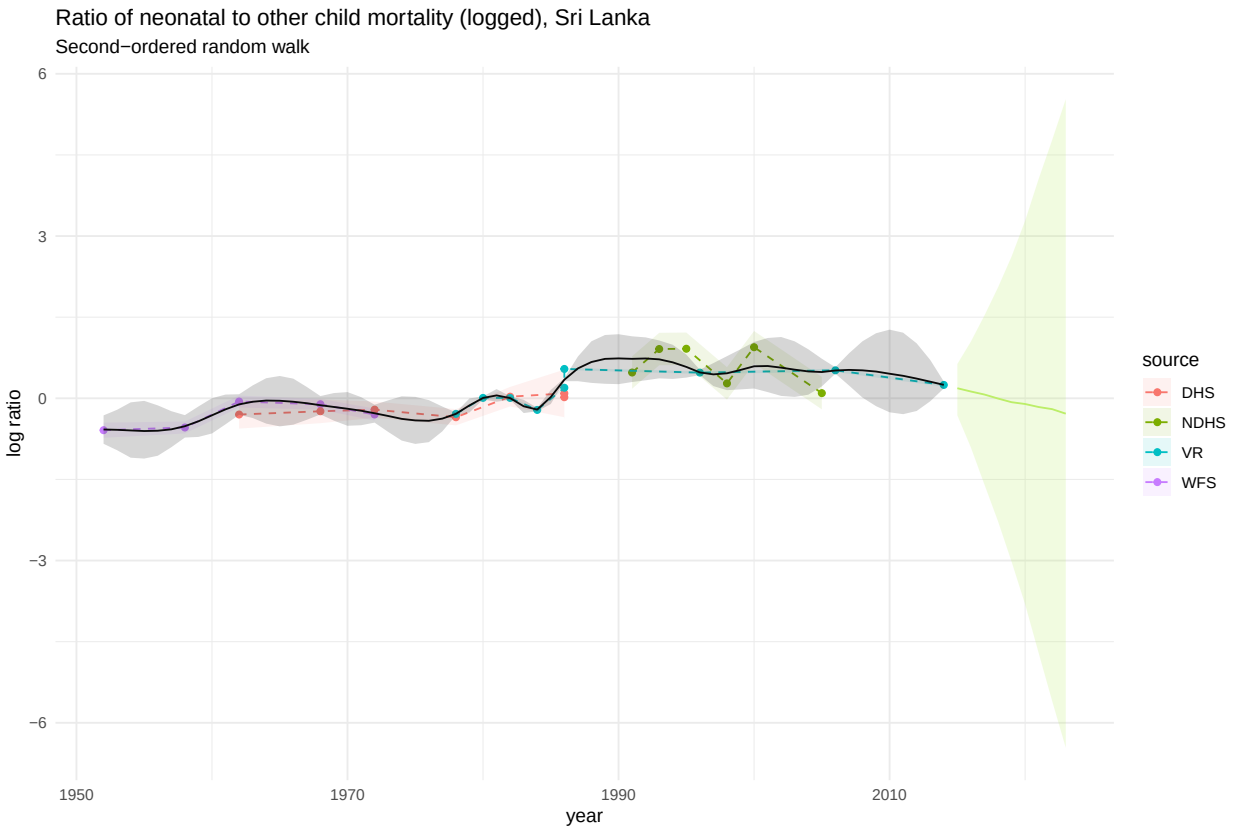
Question 2: Code up and estimate a first order random walk model to fit to the Sri Lankan data, taking into account measurement error, and project out to 2023.

```
## # A tibble: 6 x 9
##       p .variable .value .lower .upper .width .point .interval  year
##   <int> <chr>      <dbl>  <dbl>  <dbl>  <dbl> <chr>  <chr>     <int>
## 1     1 mu_p       0.249 -0.104  0.602   0.95 median qi         2015
## 2     2 mu_p       0.250 -0.225  0.746   0.95 median qi         2016
## 3     3 mu_p       0.247 -0.335  0.867   0.95 median qi         2017
## 4     4 mu_p       0.250 -0.459  0.970   0.95 median qi         2018
## 5     5 mu_p       0.249 -0.533  1.05    0.95 median qi         2019
## 6     6 mu_p       0.240 -0.612  1.11    0.95 median qi         2020
```



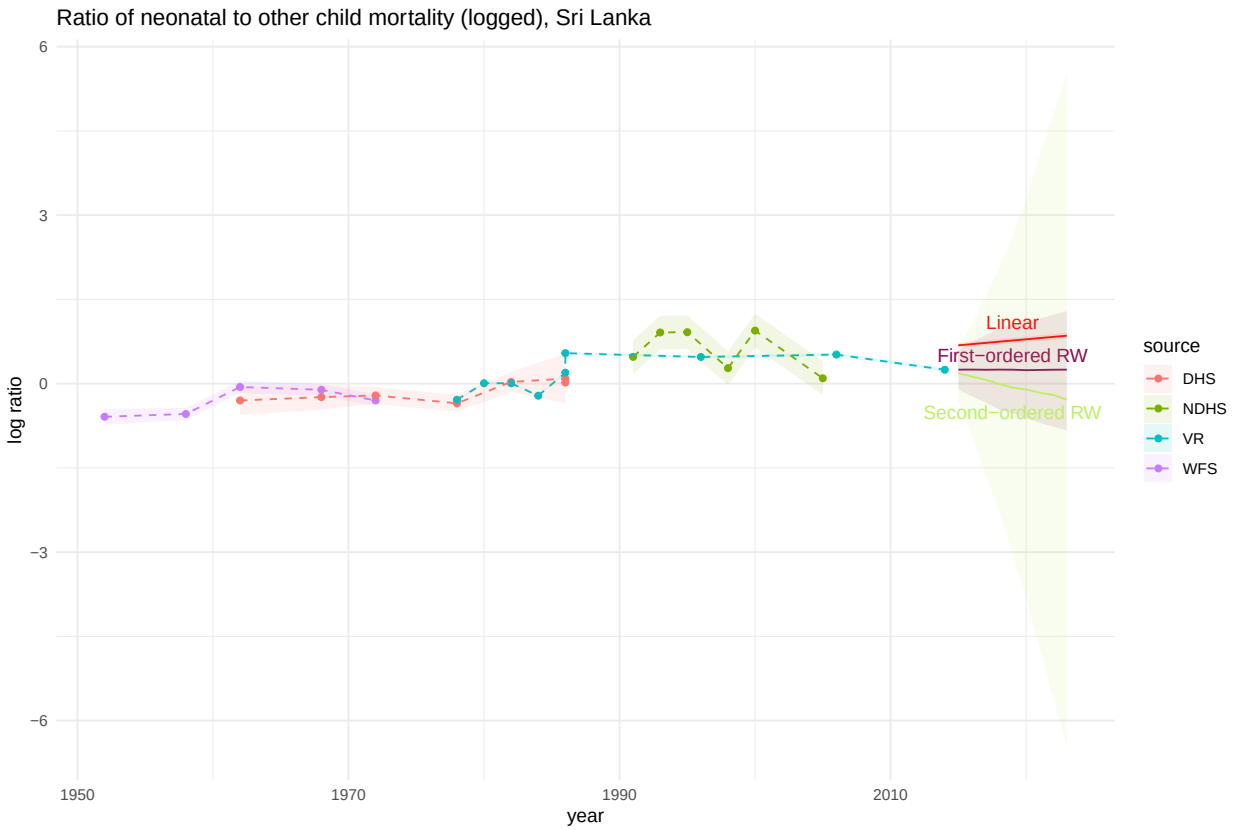
Question 3: Now alter your model above to estimate and project a second-order random walk model (RW2).

```
## # A tibble: 6 x 9
##       p .variable   .value .lower .upper .width .point .interval  year
##   <int> <chr>        <dbl>  <dbl>  <dbl>  <dbl> <chr>  <chr>     <int>
## 1     1 mu_p       0.186   -0.316  0.634   0.95 median qi         2015
## 2     2 mu_p       0.125   -0.910  1.05    0.95 median qi         2016
## 3     3 mu_p       0.0694  -1.60   1.53    0.95 median qi         2017
## 4     4 mu_p      -0.00445 -2.27   2.05    0.95 median qi         2018
## 5     5 mu_p      -0.0723  -3.02   2.62    0.95 median qi         2019
## 6     6 mu_p      -0.105   -3.81   3.28    0.95 median qi         2020
```



Question 4: Run the first order and second order random walk models, including projections out to 2023. Compare these estimates with the linear fit by plotting everything on the same graph.

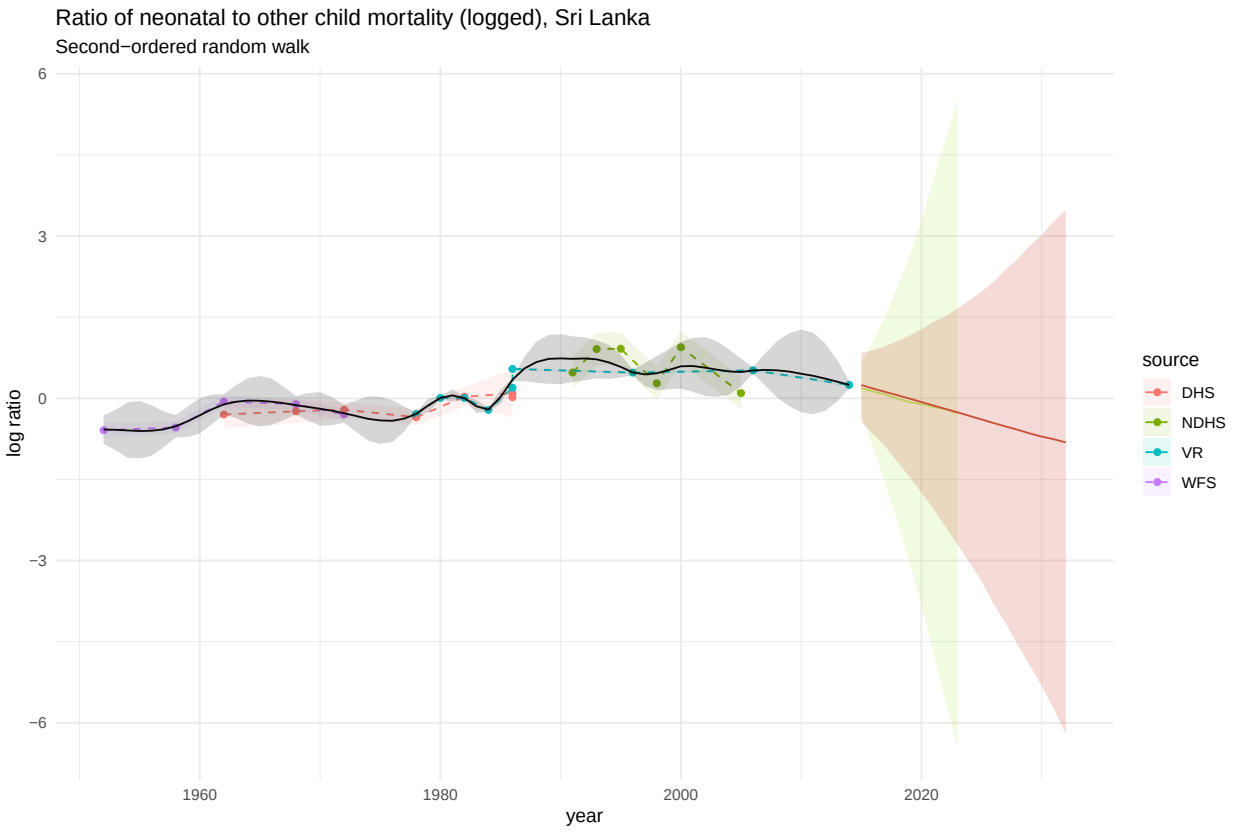
The linear projection reflects an upward trend with the smallest prediction interval, the first-ordered random walk projection seems to go flat, and the second-ordered one indicates a downward tendency after 2014 with the widest prediction interval.



Question 5: Rerun the RW2 model excluding the VR data. Briefly comment on the differences between the two data situations.

```
## # A tibble: 6 x 9
##       p .variable   .value .lower .upper .width .point .interval  year
##   <int> <chr>        <dbl>  <dbl>  <dbl>  <dbl> <chr>  <chr>     <int>
## 1     1 mu_p       0.245   -0.440  0.838   0.95 median qi         2015
## 2     2 mu_p       0.179   -0.662  0.891   0.95 median qi         2016
## 3     3 mu_p       0.118   -0.893  0.963   0.95 median qi         2017
## 4     4 mu_p       0.0571  -1.18   1.06    0.95 median qi         2018
## 5     5 mu_p      -0.00560 -1.45   1.16    0.95 median qi         2019
## 6     6 mu_p      -0.0691  -1.74   1.27    0.95 median qi         2020
```

Without VR data, the second-ordered random walk model yields smaller estimates of the log ratios than the observed values from 2006 to 2014 and predicts lower log ratios for the period from 2015 to 2023.



Question 6: Briefly comment on which model you think is most appropriate, or an alternative model that would be more appropriate in this context.

The second-ordered random walk model fitted to all available data seems to be the most appropriate since it reflects the gradual downward trend of the read data. That said, it is not doing a great job of capturing the real situation in Sri Lanka since the prediction interval is quite large.