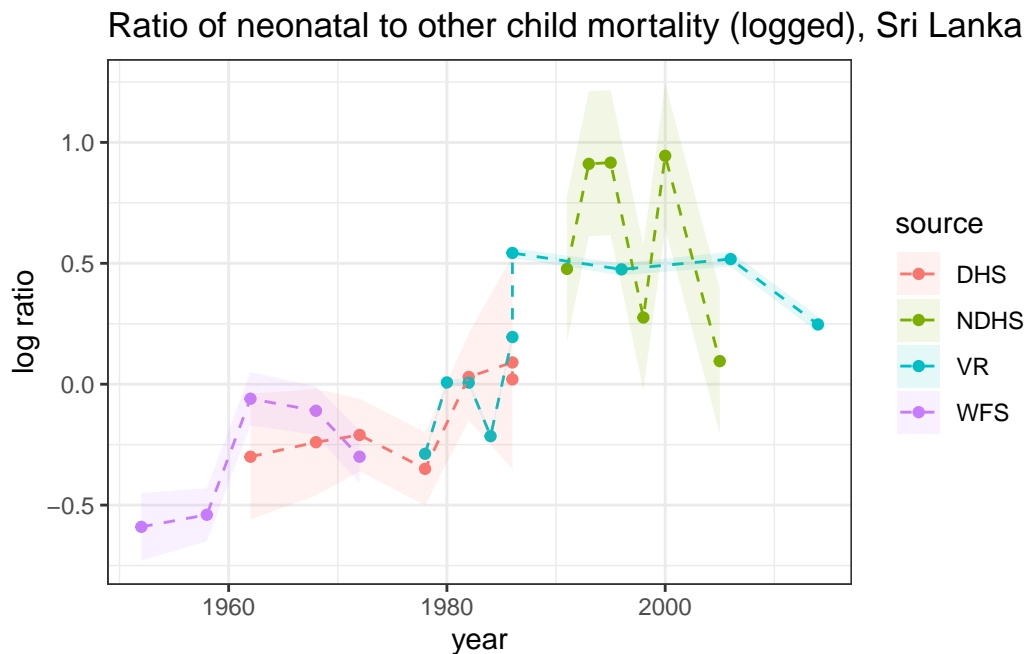


Week 10: Temporal data

26/03/23

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:

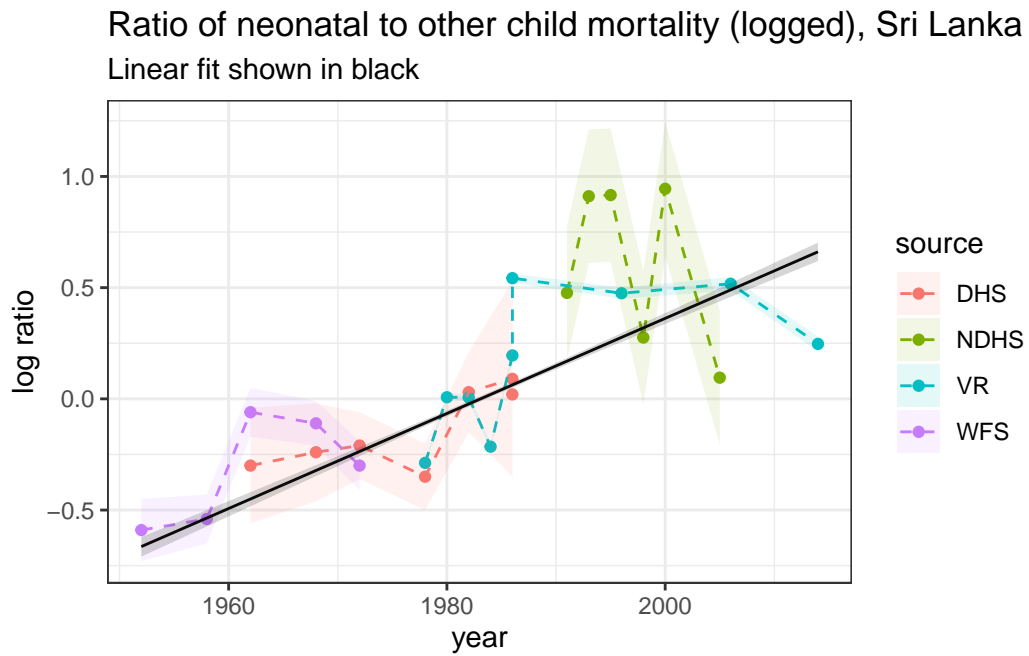


Fitting a linear model

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

Extract the results:

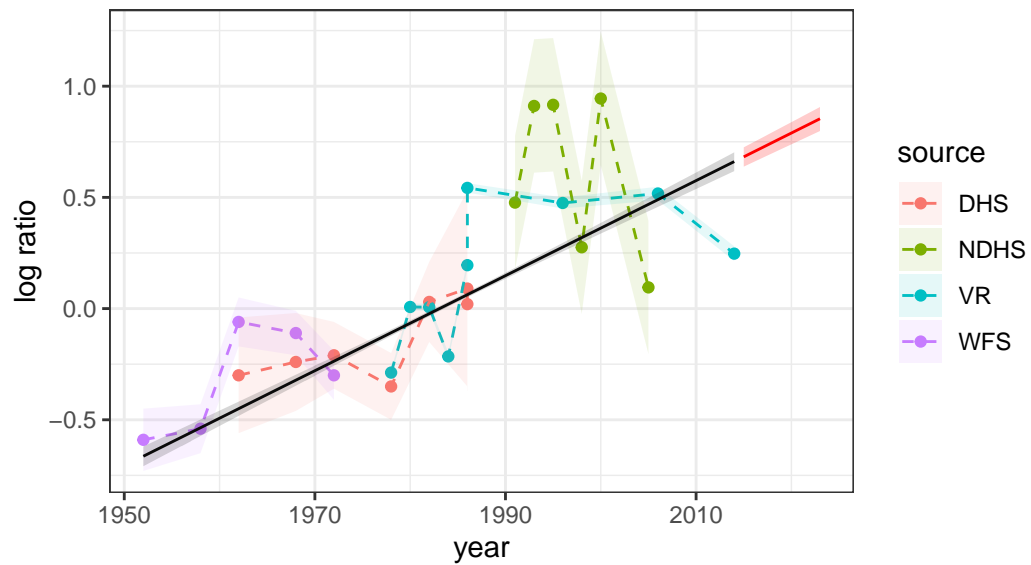
Plot the results:



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.

Ratio of neonatal to other child mortality (logged), Sri Lanka
Linear fit shown in black

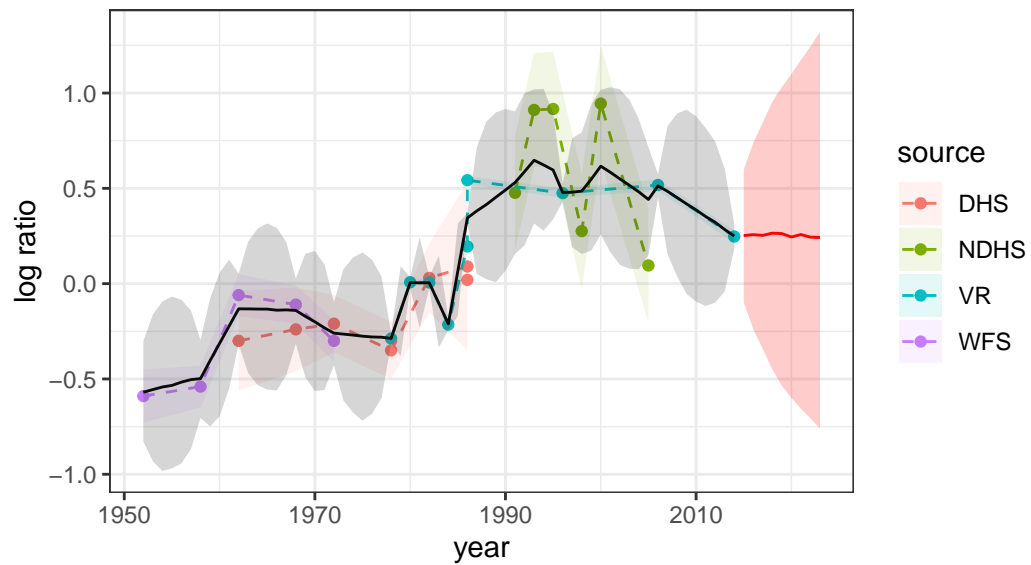


Random walks

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.

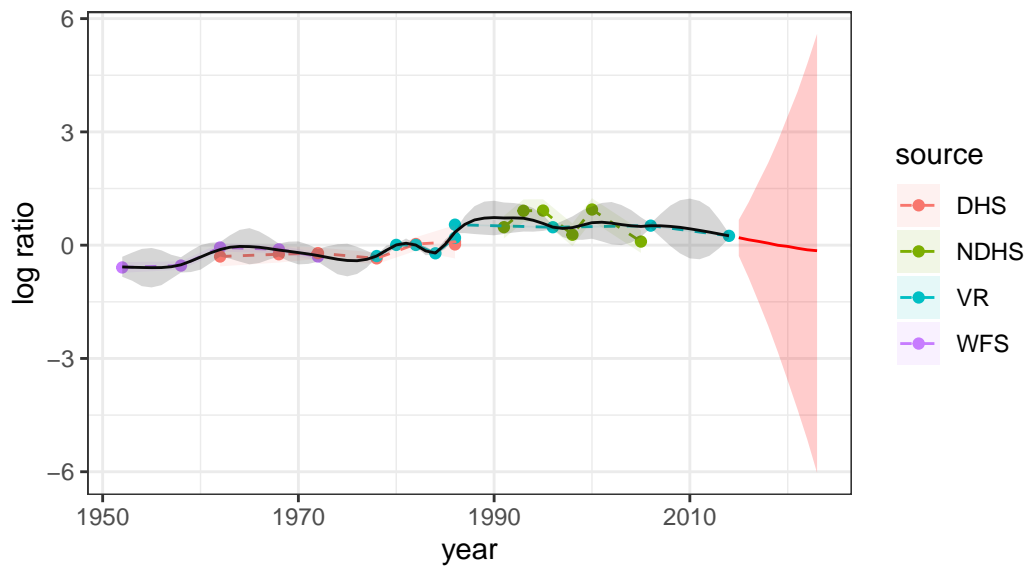
Ratio of neonatal to other child mortality (logged), Sri Lanka
RW1 fit shown in black



Question 3

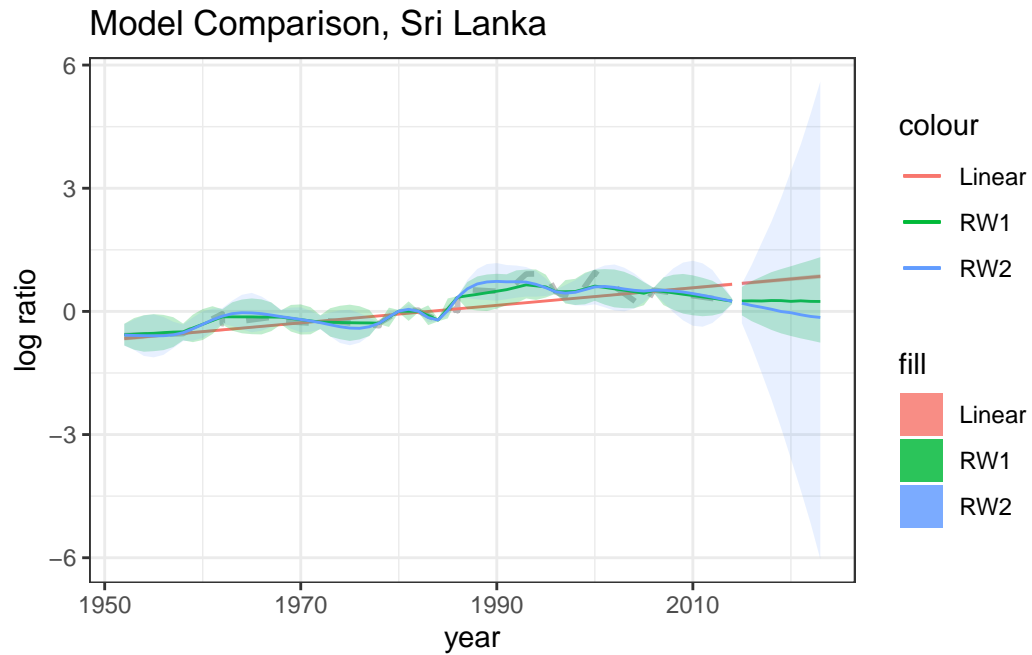
Now alter your model above to estimate and project a second-order random walk model (RW2).

Ratio of neonatal to other child mortality (logged), Sri Lanka
RW2 fit shown in black



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.

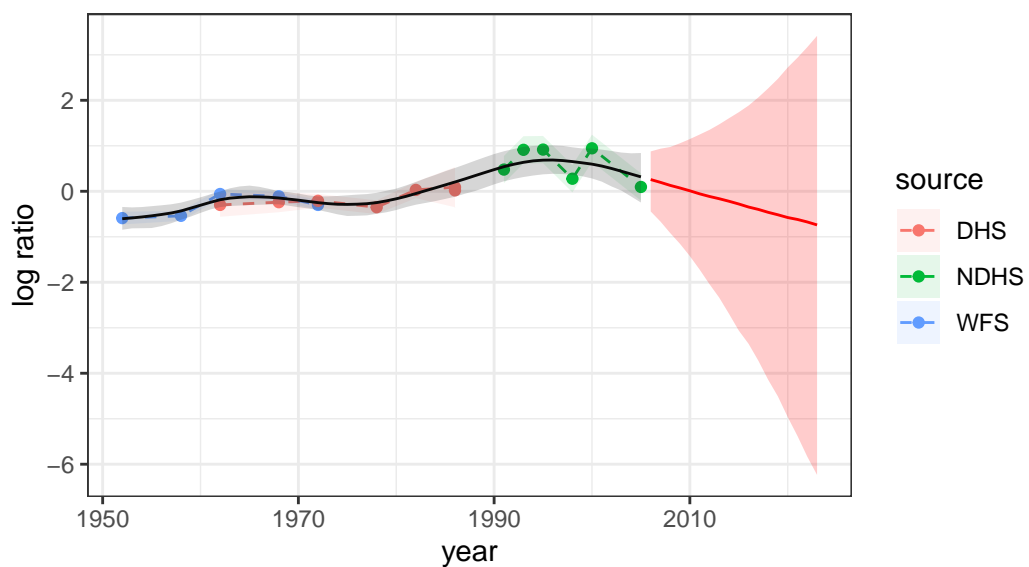


Question 5

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

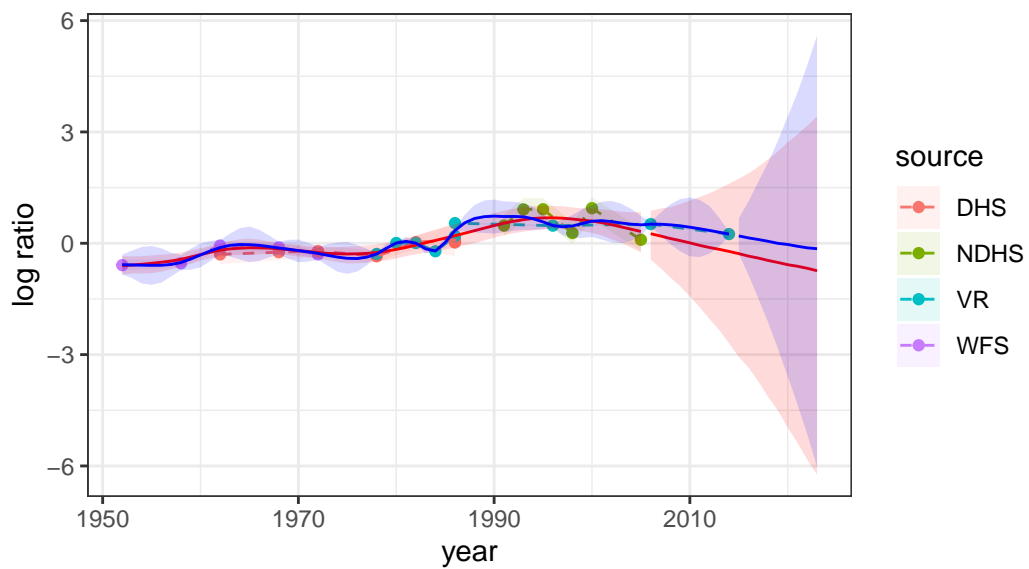
- Assuming that we are still predicting to 2023, here $P = 18$.

Ratio of neonatal to other child mortality (logged), Sri Lanka, Vt
RW2 fit shown in black



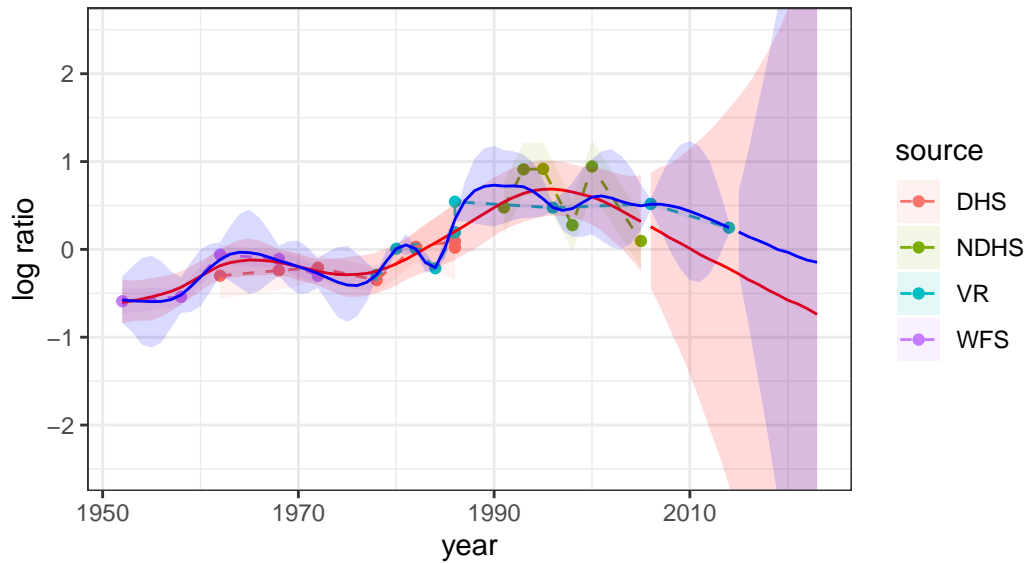
VR (Blue) vs No VR (Red)

RW2 fit shown in black



VR (Blue) vs No VR (Red)

RW2 fit shown in black



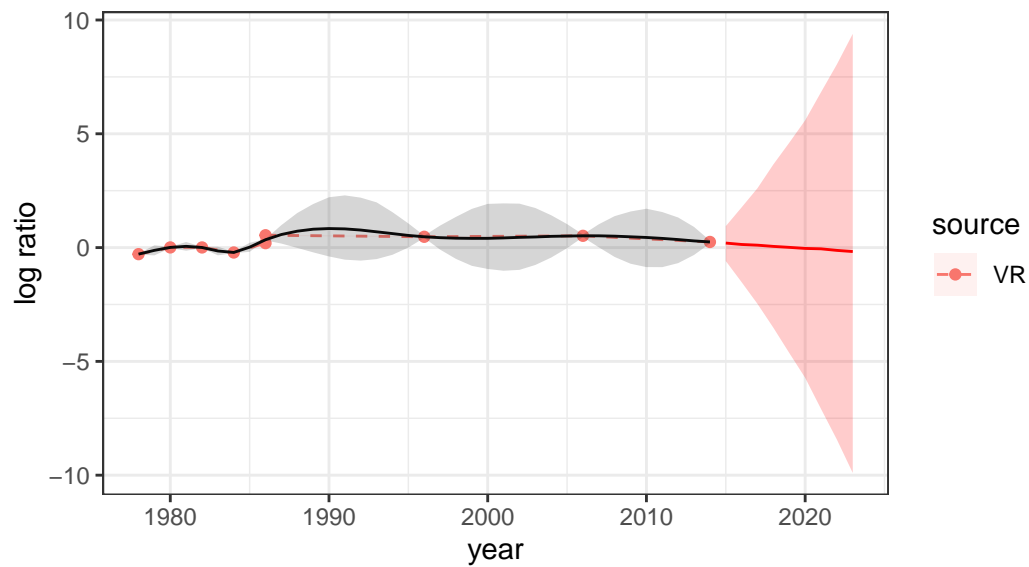
- When the VR data is removed, the fitted line is much smoother, and the CI band is much more even in width.
- When the VR data is added, since all VR data have small se, the fitted line seems to be forced to stay close to the VR data points. For the same reason the CI band is very narrow near the VR data points, but becomes very wide when not close to a VR data point.

Question 6

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

Model 3, VR data only

RW2 fit shown in black

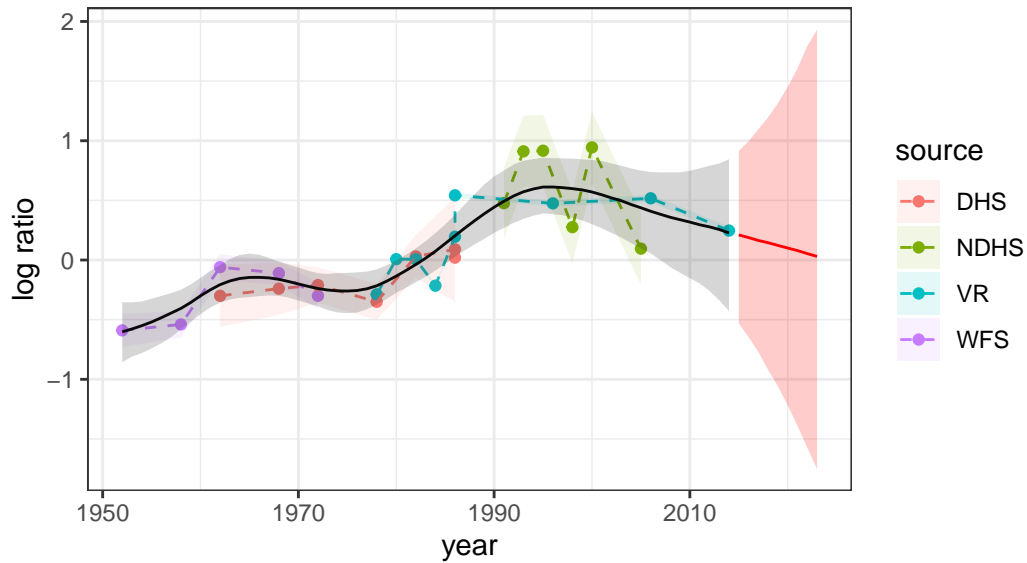


- When we use the RW2 model for VR data only, the model has the issue of uneven CI band (and likely not smooth enough as well).
- For experimenting purpose, relax the se in VR data to be similar to other data:

```
# A tibble: 2 x 2
  if_vr se_mean
  <chr>   <dbl>
1 N_VR    0.231
2 VR      0.0210
```

Model 3, VR data only

RW2 fit shown in black



- The fitted and prediction seems to be much more plausible comparing to before.
- I am not sure if this count as a new alternative model, but when we manually lose some information and widen the se, the result becomes better and the model seems to become more appropriate.
- As for the 3 original model, RW2 seems to be the best among 3, given that the se is not too small. Neither linear model's nor the RW1 model's prediction seems reasonable here, but the RW1 model does fit the existing data better.