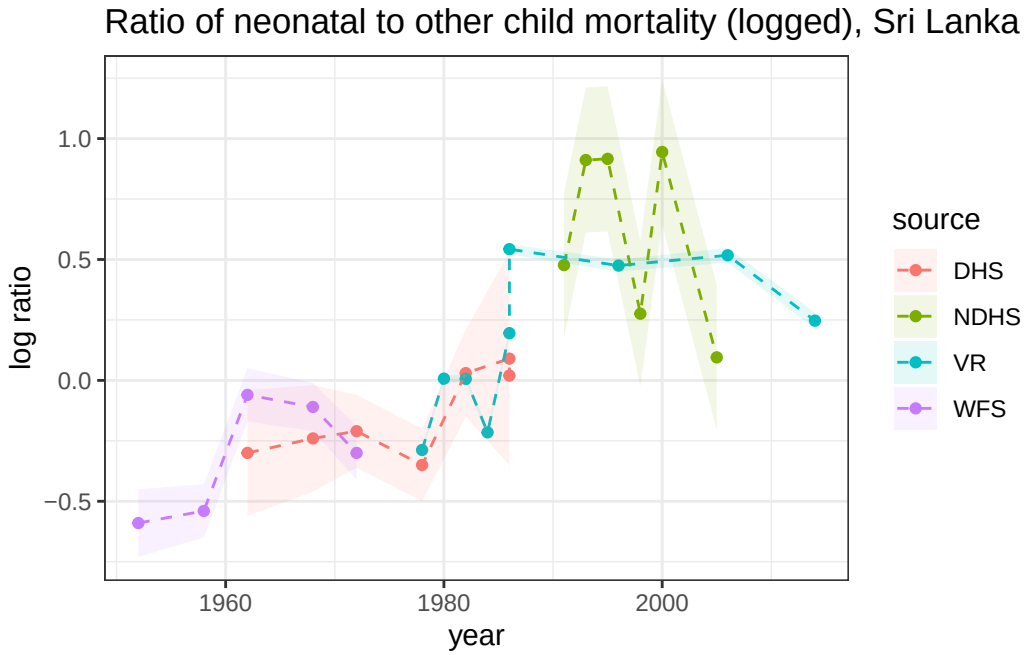
Week 10: Temporal data

Priyanka Verma

25/03/23

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:

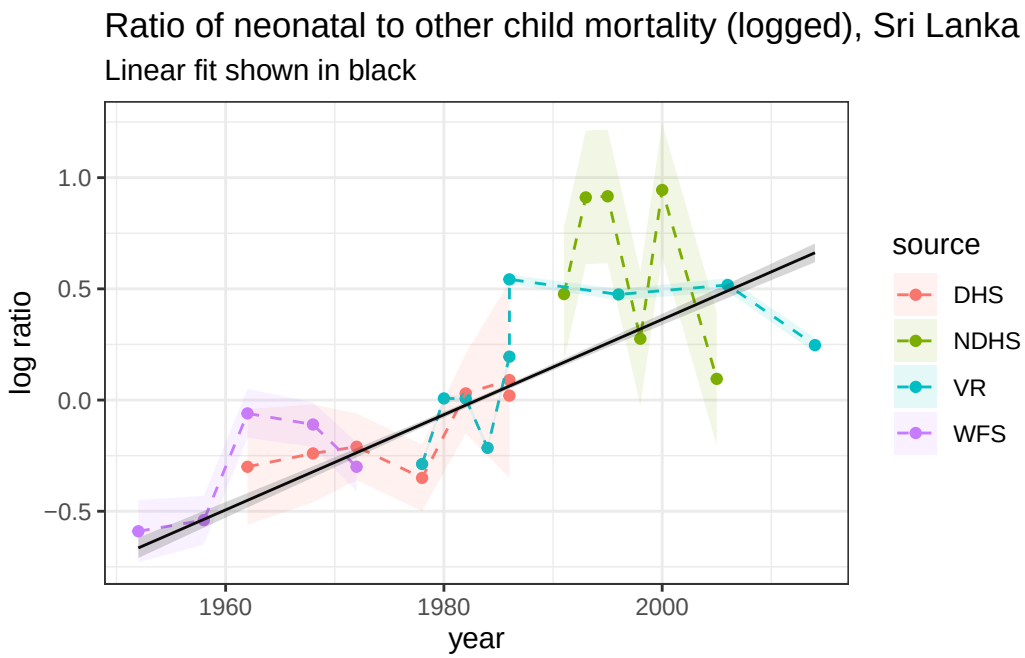


Fitting a linear model

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

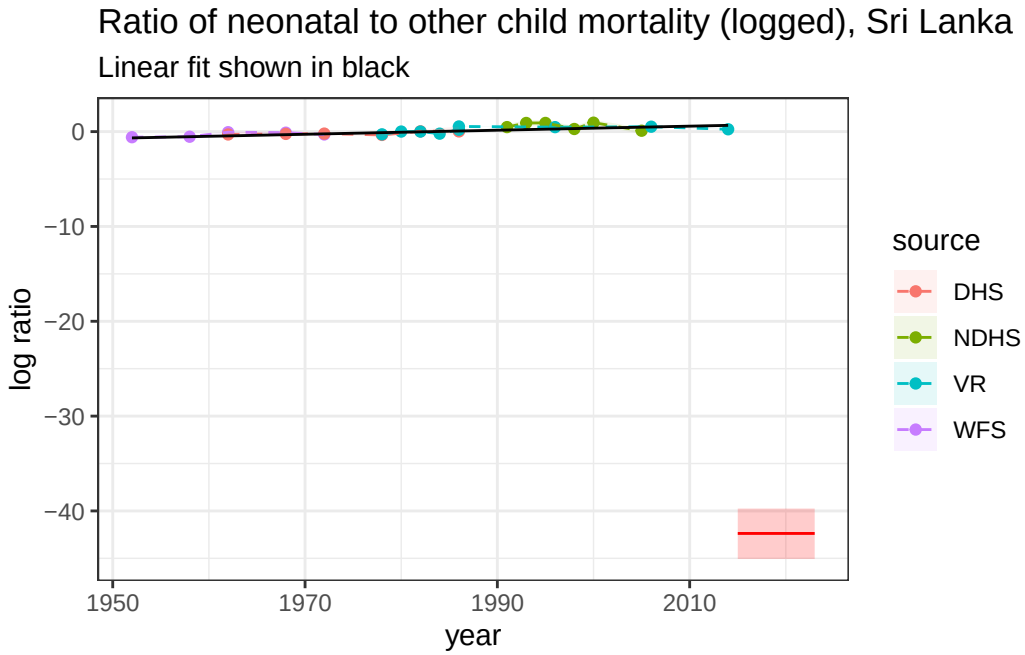
Extract the results:

Plot the results:



Question 1

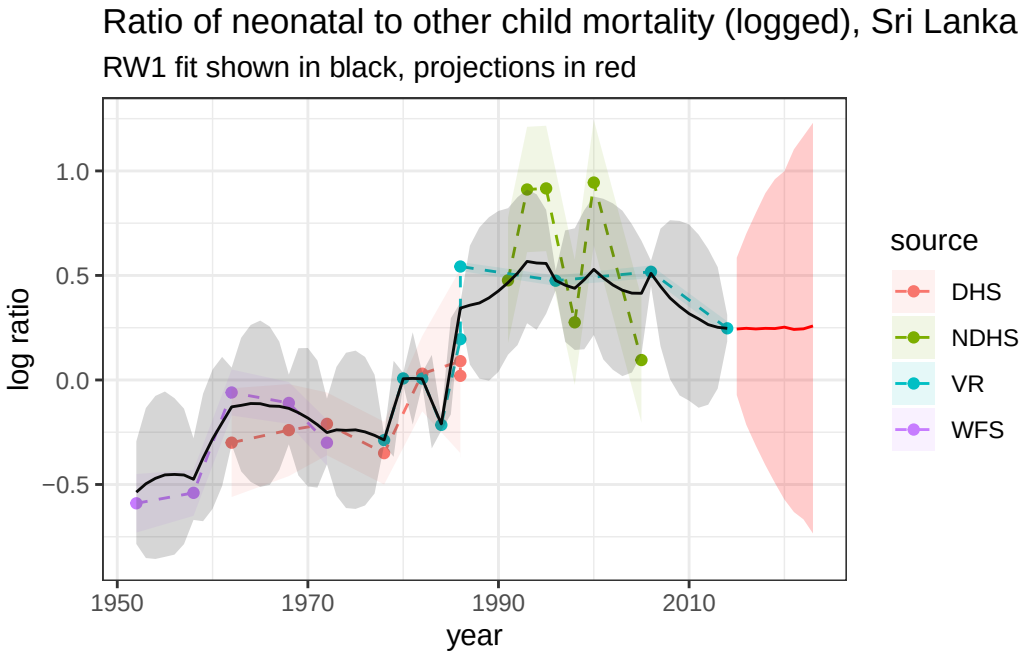
Project the linear model above out to 2023 by adding a `generated quantities` block in Stan (do the projections based on the expected value μ). Plot the resulting projections on a graph similar to that above.



Random walks

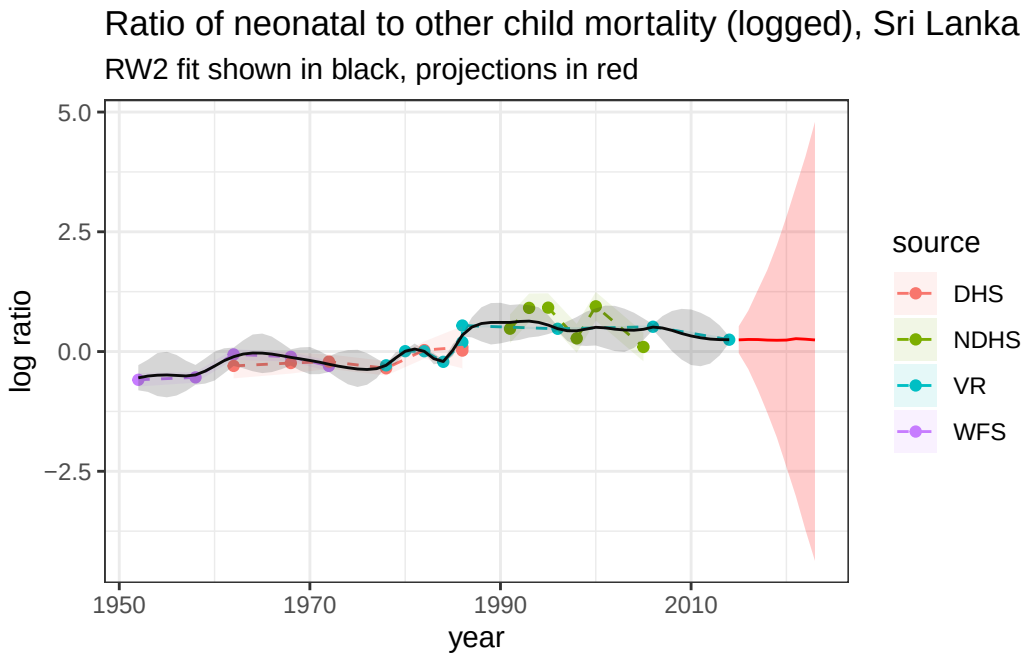
Question 2

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.



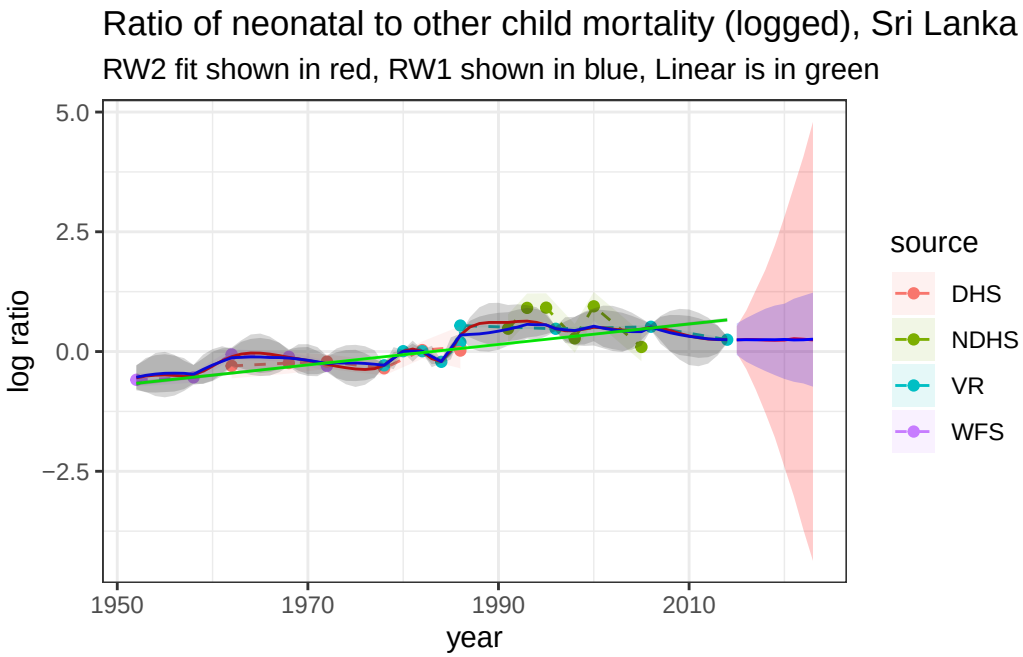
Question 3

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



Question 4

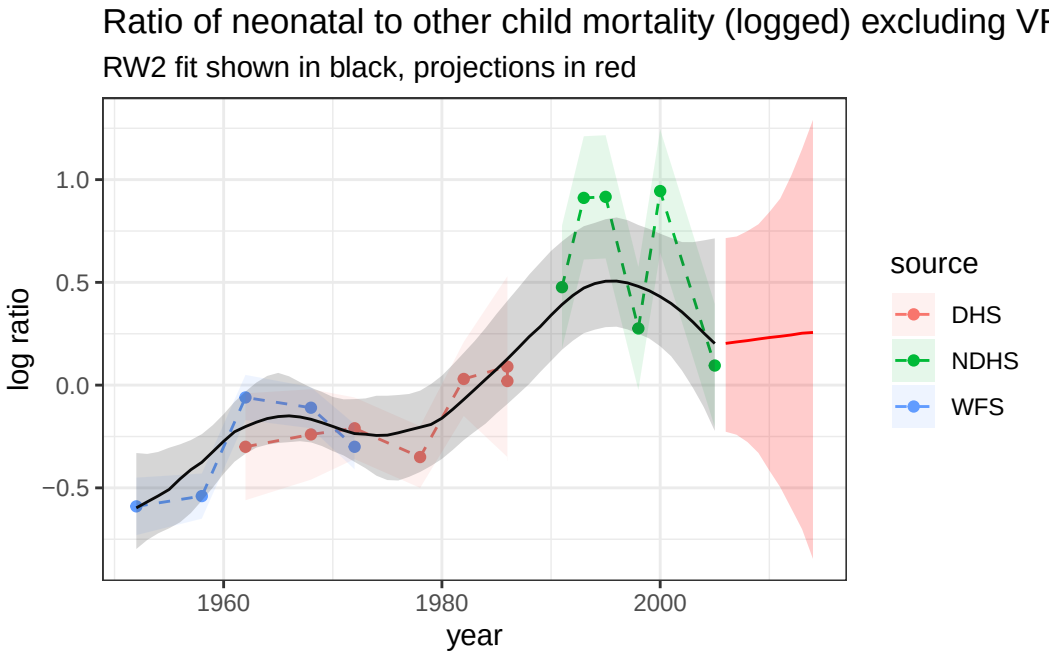
Run the first order and second order random walk models, including projections out to 2023. Compare these estimates with the linear fit by plotting everything on the same graph.



- The fit for linear model is off completely.
- RW models predict the data better as compared to linear model. RW2 has more uncertainity around prediction values, however the within data fit is smoother for RW2.

Question 5

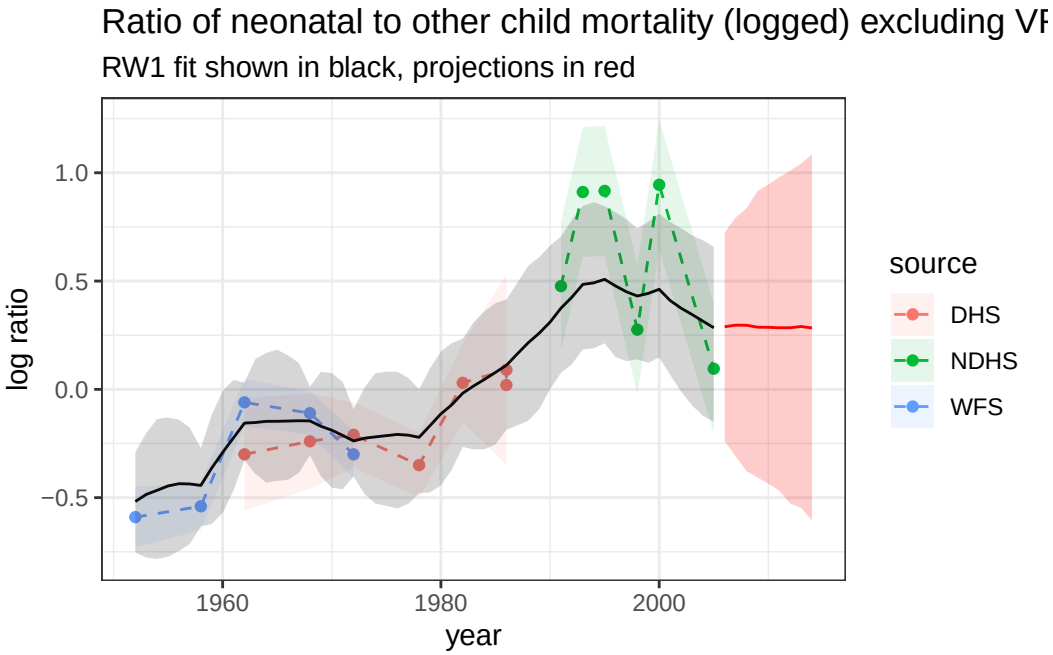
Rerun the RW2 model excluding the VR data. Briefly comment on the differences between the two data situations.



- For the context when VR data source is not included, the uncertainity in estimates produced by RW2 model are lesser than the case when RW2 is trained with VR data included. The model fit is also smoother when VR data is excluded. So, the model trained w/o VR data is better.

Question 6

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



- From comparing the plot of RW2 and RW1, we can see that the predictions of RW2 are more appropriate. The uncertainity in estimates produced by RW2 model are lesser than the case when RW2 is trained with VR data included. The model fit is also smoother when VR data is excluded. So, the model trained w/o VR data is better.
- Lastly, there are multiple observations in the data for the year 1986, so it might be a good idea to manually choose the data for that year to reduce the uncertainities around that point during model fitting.