

Time Mixed Effects Model

Aissata Bah

Brice Laurent

Linh Vu

2023-12-07

Mixed Effects Linear Model for Time

To include the data points from all the years collected in both datasets, we decided to incorporate time effects into a mixed effects model. For this model, we used the variables manually selected based on correlations: **x1.6**, **x3.2**, **x5.1**, **x6.4**, and **x7.3**. This mixed effects model is based on the equation:

$$HDI_{time, country, region} = \alpha_{region} + \beta_{region} \times X_{time, country, region} + a_{country} + b \times X_{time, country} + \varepsilon_{time, country, region}$$

where

$$\alpha_{region} \sim N(\mu_{\alpha}, \sigma_{\alpha}^2)$$

$$\beta_{region} \sim N(\mu_{\beta}, \sigma_{\beta}^2)$$

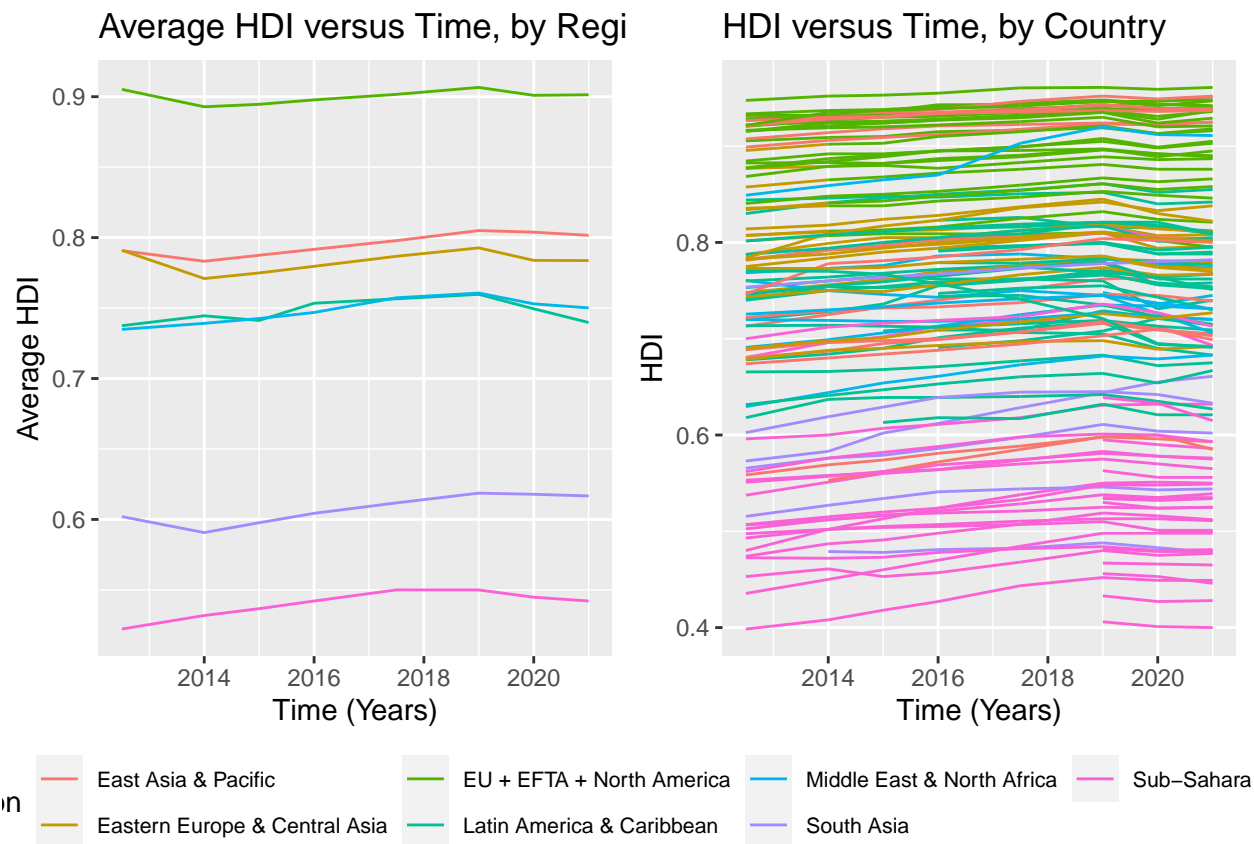
$$a_{country} \sim N(\mu_a, \sigma_a^2)$$

$$b_{country} \sim N(\mu_b, \sigma_b^2)$$

$$\varepsilon_{time, country, region} \sim N(0, \sigma_{HDI}^2)$$

$X_{time, country, region}$ is the design matrix for the variables **x1.6**, **x3.2**, **x5.1**, **x6.4**, **x7.3**, and intercept, while $X_{time, country}$ is the design matrix for the time variable (in years) and intercept.

This equation above assumes independence between random effects. Though this assumption is reasonable for the random effects of Rule of Law variables between regions, it is not as reasonable for the random effects of the time variable between countries. This can be seen from the chart below since the HDI of different countries mostly follow similar trajectories over times, and therefore are associated. There is no closed form solution when we allow correlation between random effects, so the actual mixed effects model will not exactly follow the equation above.



The coefficients from the fixed effects are as follows:

	Estimate	Std. Error	df	t value	Pr(> t)
(Intercept)	0.7517782	0.0224802	173.8784	33.4418193	0.0000000
years_from_2000	0.0020507	0.0002651	113.3425	7.7355831	0.0000000
x1.6	0.0092192	0.0065835	756.7572	1.4003514	0.1618180
x3.2	0.0042736	0.0039213	710.9956	1.0898585	0.2761447
x5.1	-0.0018105	0.0065328	710.4018	-0.2771414	0.7817522
x6.4	-0.0078659	0.0047829	719.6871	-1.6445824	0.1004927
x7.3	0.0160220	0.0077260	772.5638	2.0737721	0.0384313
regionEastern Europe & Central Asia	0.0570565	0.0236452	124.3715	2.4130289	0.0172794
regionEU + EFTA + North America	0.0626189	0.0235676	122.7752	2.6569930	0.0089322
regionLatin America & Caribbean	-0.0617601	0.0246805	123.0787	-2.5023842	0.0136475
regionMiddle East & North Africa	-0.0470149	0.0344864	122.4413	-1.3632880	0.1752933
regionSouth Asia	-0.1820306	0.0380781	122.7845	-4.7804555	0.0000049
regionSub-Saharan Africa	-0.2567541	0.0245990	123.7915	-10.4375705	0.0000000

The only fixed variable that was significant (both unadjusted and adjusted) in this model was the `years_from_2000` variable. The other variables were very far from significance.

```
## Data: selected_columns
## Models:
## time_mixed_effects_noYear: hdi ~ years_from_2000 + x1.6 + x3.2 + x5.1 + x6.4 + x7.3 + region + (1 | country)
## time_mixed_effects: hdi ~ years_from_2000 + x1.6 + x3.2 + x5.1 + x6.4 + x7.3 + region + (1 + years_from_2000
##
##               npar      AIC      BIC logLik deviance  Chisq Df
## time_mixed_effects_noYear    15 -5086.4 -5014.1 2558.2 -5116.4
## time_mixed_effects           17 -5536.4 -5454.5 2785.2 -5570.4 454.03  2
##
##               Pr(>Chisq)
## time_mixed_effects_noYear
## time_mixed_effects      < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
## Data: selected_columns
## Models:
## time_mixed_effects_noMixed: hdi ~ years_from_2000 + x1.6 + x3.2 + x5.1 + x6.4 + x7.3 + region
## time_mixed_effects: hdi ~ years_from_2000 + x1.6 + x3.2 + x5.1 + x6.4 + x7.3 + region + (1 + years_from_2000
##
##               npar      AIC      BIC logLik deviance Chisq Df
## time_mixed_effects_noMixed    14 -2888.7 -2821.3 1458.3 -2916.7
## time_mixed_effects           17 -5536.4 -5454.5 2785.2 -5570.4 2653.7  3
##
##               Pr(>Chisq)
## time_mixed_effects_noMixed
## time_mixed_effects           < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

We also analyzed the importance of including random effects of the variables x1.6, x3.2, x5.1, x6.4, and x7.3 based on region, and the random effects of the time variable based on country using an extra-sum-of-squared F test. These F-tests

