Final Term Project - Child Mortality and Girls' Education

A. BOGNAR

Introduction

Education, specifically education of girls is cited by several international NGOs as driving force for reducing child mortality. UNICEF highlights 3 mechanisms for the connection:

1. More educated women means more healthcare professionals.
2. Women with degrees can join the workforce, reducing poverty and risk factors of poor health.
3. Educated mothers have access to information to take better care of their children.

In this analysis I look at UN data to see if I can validate the claim of this association.

Data

My main data source is the The State of the World's Children 2019 Statistical Tables which aggregates the main indicators gathered by different UN agencies about issues affecting children worldwide. I also used UNPD country population data.

These indicators are generally gathered from two primary sources:

1. Civil Registrations from the country's vital registration system.
2. Household surveys by UN agencies in countries where vital systems are inadequate.

These sources can introduce measurement errors by humans and machines. In principal, this would make my findings understate associations, but the UN agencies took steps to minimize these effects so we can expect them to be low. Similarly, the different methods of collection may produce different level of errors between the countries, we can also expect this effect to be minimal.

Table 1: Descriptive statistics

	N	Missing	Mean	Median	SD	Min	Max	P05	P95
Mortality under 5 (in thousand births)	131	0.0	30.2	20.0	27.8	2.0	119.0	3.0	84.5
Out of Primary school rate, Girls (%)	131	0.0	10.9	6.0	13.5	0.0	72.0	0.0	38.0
Out of Primary school rate, Boys (%)	131	0.0	10.4	6.0	12.3	0.0	64.0	1.0	30.0
GDP per capita (USD)	131	0.0	10 393.0	4104.6	16 170.2	283.5	104 498.7	625.3	41 292.8
Institutional delivery (%)	131	0.0	84.8	93.0	19.7	12.0	100.0	43.0	100.0

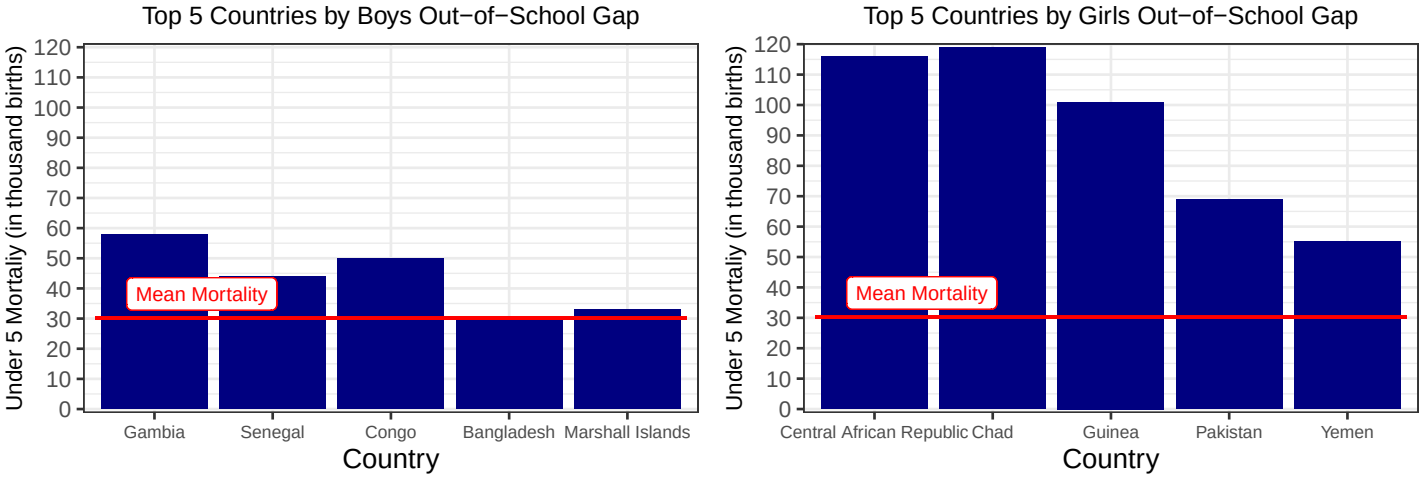
From the UNICEF tables I chose the above 5 measurements for my models. Already, you can see that I only had the relevant data points for 131 countries (202 total countries have any data listed). Even with UN surveys it is difficult to get information about some countries, I considered several other measures that I could not include due to lack of data (for example welfare measures). Also, these are not perfect indicators of what I'm interested in about the countries but they are the best approximations I could find:

- Under 5 Mortality is the most widely used number for child mortality as worldwide 70% of under 25 deaths occur in the first 5 years.
- For Education, I chose % of children out of Primary schools. Alternatives were out of secondary school and school completion rates or a combination of them. The latter would have been an arbitrary number, as for completion it may differ more country to county while being in school is a more objective category even if we don't learn what actually happens in schools.
- GDP per capita is used to represent wealth of a country, albeit not perfectly.
- Institutional delivery stands for the % of births that took place in hospitals or other medical institutions. This is very relevant for child deaths and for the state of healthcare in a country.

Based on *Table 1* all measures except Institutional delivery have higher means than medians. This means many values below the average and few high values increasing the average. Generally, for better models such measures can be changed by taking the logarithm of their values and interpreting the model results as percentage differences. In my case, Mortality and Education measures are rates and percentages so for easier understanding I only transformed GDPP. This will make my models worse approximations.

Gender gaps in education

My main focus is not gender differences in education, but I include a section about it as it is relevant to my findings. In my sample about 75% countries have low, 2% or less, difference in out of primary school rates between genders. But while no countries have more than 10% higher absence rate for boys than girls, there are 3 countries with more than 15% higher out-of-school rate for girls. In addition, consider the below graphs about the countries with the highest gender disparities.



Child mortality is high, above average, in almost all countries with gender gaps, but in countries where girls are disadvantaged it is about twice as much. The main takeaways:

- Most countries don't have out-of-school rate gender gaps. Those that do, have higher gaps for girls.
- Where there is a big gender gap child mortality is high.
- But it is much higher where there is a gender gap for girls.

Models

With the data described, let me present my findings. First, the general patterns of association.

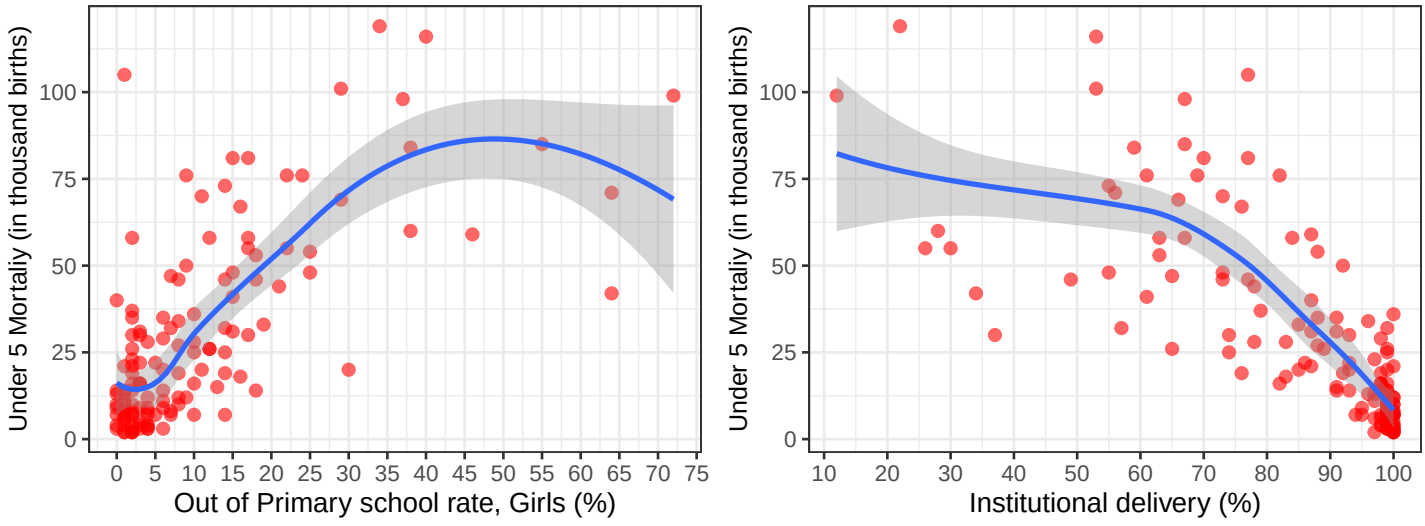
Below you can see scatterplots of two measurements compared with child mortality. In each graph the pattern of association is represented by a line, the best approximation that the data allows. Here we see the issue of skewed data, most countries fall very close to one end of each graph so it is difficult to find a straight-forward pattern. Still, we can describe what we see.

Low out-of-school rate for girls seem to have no clear pattern with child mortality. After a certain point however (around 3%) higher child mortality starts to be associated with more girls left out of education. At a higher point where out-of-school rate is high (around 45%) this pattern seems to turn around. But there are few countries above that rate and they have high mortality rates, even if not increasing. So this pattern is likely not a robust finding.

For hospital deliveries the pattern is also interesting. While higher institutional delivery is associated with lower mortality, this pattern is especially strong after a point where the majority of births are delivered by professionals.

Therefore, in my models I will:

- Make a distinction in the association between mortality and girls' education above and below 3% out-of-school rate.
- And between countries where at least 2/3s of children are delivered in hospitals.



Finally, you can see the results of my models in *Table 2*.

Models 1 and 2 show that for countries with 1% higher out-of-school rate we expect 1.3 more deaths before age of 5 from 1000 births on avg. This is true for both genders with the caveat that for girls this is only true above 3% primary school absence. Below that variation among countries is too high to draw a conclusion.

In **Model 3** we want to see the pattern between the education of a gender and child mortality with the difference in the education of the other gender accounted for. We find that lack of girls education is associated with higher child mortality while lack of boys education is associated with lower child mortality. Does this mean that lower education for boys would be associated with less child mortality? No. As I argued, level of education of genders tend to go together, it is simply the case that countries most affected by child mortality also have lower levels of education for girls. In other words this shows that with current rates of education, girls out-of-school rate explains changes in child mortality more than boys'.

Adding to the complexity **Model 4** accounts for the other measurements I described previously. Higher GDP per capita is strongly associated with child mortality (1% higher GDPP on avg. with about 0.08 lower child mortality rate) along with institutional delivery (above 2/3 hospital deliveries, 1% higher hospital delivery associated with 0.85 lower mortality rate). Girls education is still associated with lower mortality even controlled for these 2 additional measures.

Next, I wanted to see how widely these findings can be applied. In **Model 5** I used the same measurements, but instead of comparing countries, I weighted each state by their population. I compared people who happen to live in different countries. The findings are relatively similar. The biggest difference is that Institutional Delivery no longer shows a statistically significant association. That means that countries with big population and most births in hospitals tend to have high mortality rates. India has an above avg. mortality and 80% institutional delivery. Since I didn't have data for China, this makes most of the difference.

Table 2: Models to uncover relation between child mortality and girls education

	(1)	(2)	(3)	(4)	(5)
Intercept	13.0**	16.4***	14.3**	92.0***	108.6***
	(4.75)	(2.46)	(4.97)	(18.3)	(17.3)
Girls Out-of-School rate (<3 , primary)	2.31		3.95	-0.864	-1.48
	(1.91)		(2.06)	(1.56)	(1.61)
Girls Out-of-School rate (>=3, primary)	1.37***		2.90***	2.04***	1.64***
	(0.224)		(0.555)	(0.363)	(0.309)
Boys Out-of-School rate (primary)		1.33***	-1.76**	-1.52***	-0.867*
		(0.198)	(0.624)	(0.424)	(0.428)
Ln GDP per capita				-7.69***	-9.84***
				(1.45)	(1.91)
Institutional delivery (<67)				0.338	0.166
				(0.197)	(0.103)
Institutional delivery (>=67)				-0.851***	-0.392
				(0.156)	(0.211)
Observations	131	131	131	131	131
R2	0.461	0.342	0.512	0.797	0.856

Generalization

Let me clarify where the previous results are applicable. Findings for under 3% girls out-of-school rate and under 67% institutional delivery are not significant. They can't be generalized to the sampled countries. Other findings were significant with high confidence, therefore they can be applied to the countries sampled in 2019.

131 countries had data for primary school absence rates. To see how my findings hold I used the same **Model 4** with secondary school absence and primary school completion rates. I had data for different countries in these cases, thus they represent a different sample.

Model 6 shows the results for secondary out-of-school and **Model 7** for primary school completion rates. Overall the patterns are the same as in **Model 4**. Note that in these models I did not see a difference between low and bigger out-of-school rates for girls. Also, **Model 7** is about school completion, not absence so the sign of education patterns will flip but have a similar meaning. The gender gap is smaller in both models, we have no relevant findings for boys' education but girls' schooling is associated with lower mortality in both cases. Due to lower sample size we see less significant results especially in **Model 7**.

Table 3: Models for secondary school abscence and completion of primary school

	(6)	(7)
Intercept	65.2***	144.2***
	(14.8)	(17.0)
Girls Out-of-School rate (secondary)	1.28***	
	(0.302)	
Boys Out-of-School rate (secondary)	-0.599	
	(0.350)	
Ln GDP per capita	-6.02***	-5.95*
	(1.17)	(2.30)
Institutional delivery (<67)	0.265	-0.042
	(0.216)	(0.220)
Institutional delivery (>=67)	-0.627***	-0.658***
	(0.162)	(0.189)
Girls Primary School Completion		-0.695*
		(0.344)
Boys Primary School Completion		0.141
		(0.353)
Observations	113	97
R2	0.842	0.739

Conclusion, Further Analysis

In conclusion the findings are in line with the expectations. Lower out-of-school rate for girls (likely both genders) is associated with lower child mortality even after controlling for factors that are also associated with reduced mortality such as economic growth and child births in professional environments. GDPP itself could absorb some of the association of girls' education due to the mechanism I mentioned at the beginning about women in workforce. Finding controls that exclude this effect could help further analysis. More importantly, this analysis focused only on association, not on causality. We can't conclude that increasing girls' education will decrease child mortality since we couldn't have accounted for all relevant measurements behind the pattern such as effect of social welfare or access to proper nutrition on child mortality. Still, the findings affirm that international programs aiming to decreased child mortality by getting girls into schools follow a real pattern.

Appendix

Exact Data Sources

The variables I used from the The State of the World's Children 2019 Statistical Tables came from the following tables:

- Under-5 mortality rate from Table 2 - Child mortality.
- Institutional delivery from Table 3 - Maternal and newborn health.
- Out-of-school rates and school completion rates from Table 10 - Education.
- GDP per capita from Tables 12 and 15 - Social protection and equity; Economic indicators.

For the list of variables, their descriptions and original UN sources see the variable description file.

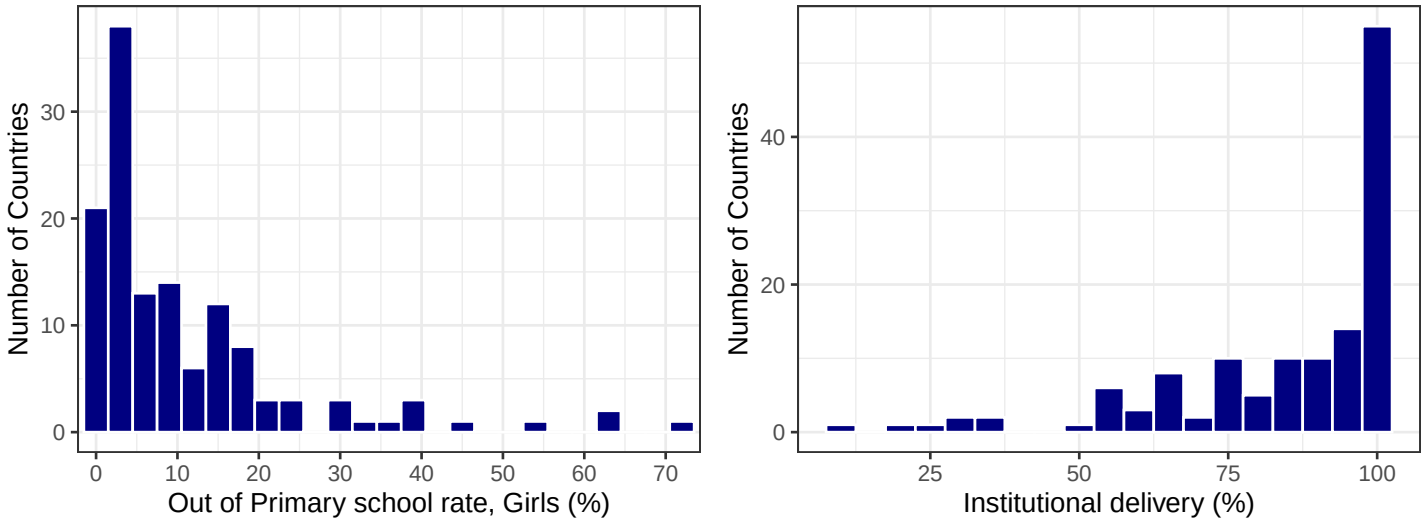
Missing data, Correlation

Table 4: Descriptive statistics

	N	Missing
Female_Prim	154	48
Male_Prim	154	48
Female_Low_Sec	134	68
Male_Low_Sec	134	68
Female_Hi_Sec	148	54
Male_Hi_Sec	148	54
Female_Prim_Co	99	103
Male_Prim_Co	99	103
Female_Low_Sec_Co	99	103
Male_Low_Sec_Co	99	103
Female_Hi_Sec_Co	99	103
Male_Hi_Sec_Co	99	103
Birth_Hosp	168	34
Mother_Cov	68	134
Child_Cov	66	136

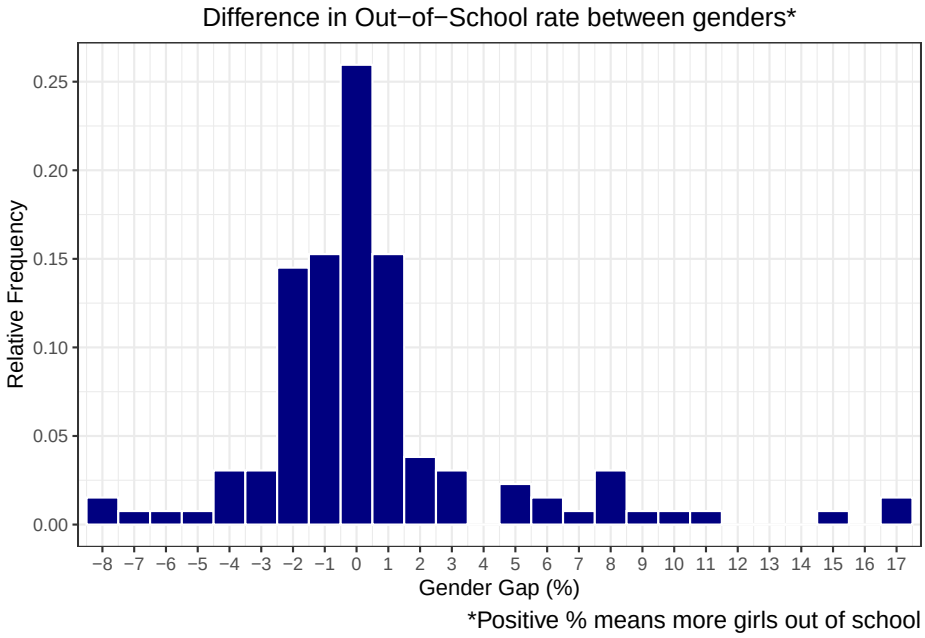
Missing data was certainly a challenge in selection, you can see in *Table 4* that from the total 202 countries listed over 100 don't have data available about school completion. Combining the different levels of education in my model would be even more challenging since not only would I work with a really low sample size but the different levels of education are closely related, including them in the same model could significantly decrease the confidence of findings. For instance in my final table the correlation coefficient between girls and boys primary out-of-school rates was 0.94. The last two variables refer to Mother's social coverage and children's benefits. I wanted to use these to measure state of social welfare in a given country for further controls. However, not many countries keep track of them. I thought about subsituting previous years' data for missing values, however, it turns out that countries with missing data, generally also didn't have data in previous years' reports.

Histograms



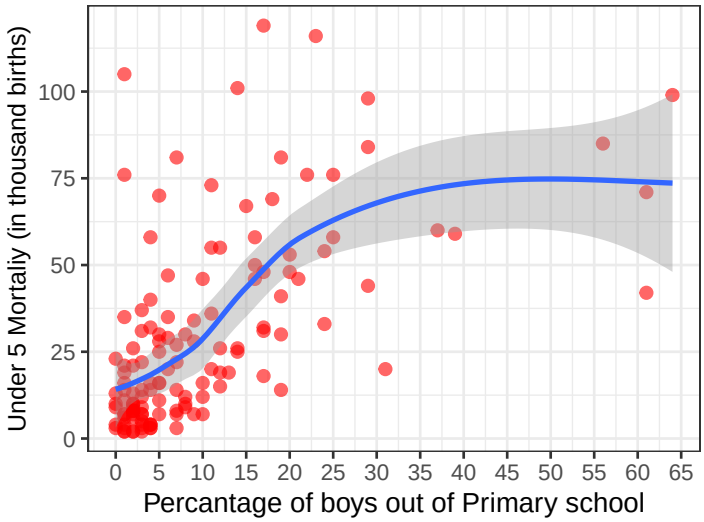
Above graphs show the distribution of female out-of-school rate (primary schools) and institutional delivery. The former is heavily right skewed with most countries having below 10% out-of-school rate. The latter is heavily left skewed with most countries having close to full institutional delivery. The most extreme value is South Sudan in both graphs with 77% out-of-school rate for primary school age girls and 12% of births taking place in healthcare facilities.

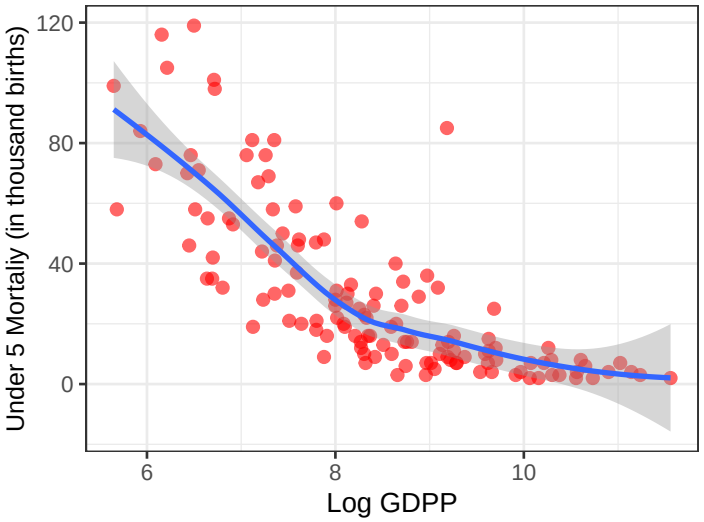
Gender Gap in Education



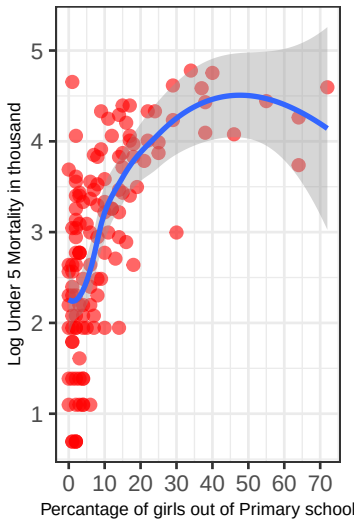
Above graph shows gender difference in out-of-school rate, calculated by subtracting male out-of-school rate (primary schools) from female out-of-school rate. As mentioned in about 75% of countries the difference is 2% or less. There are slightly more countries with gender gaps for girls but more importantly the gender differences in those countries are much higher.

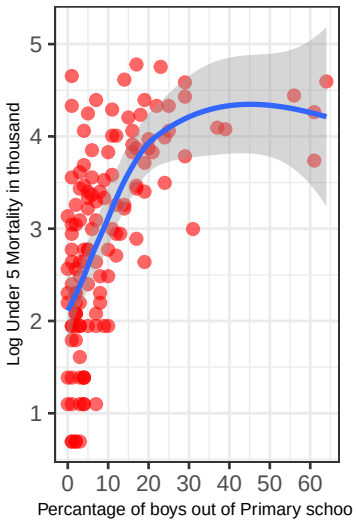
Loess graphs

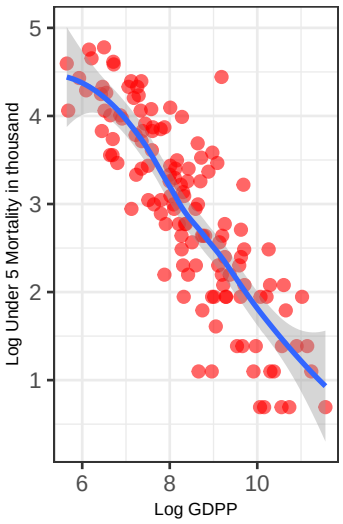


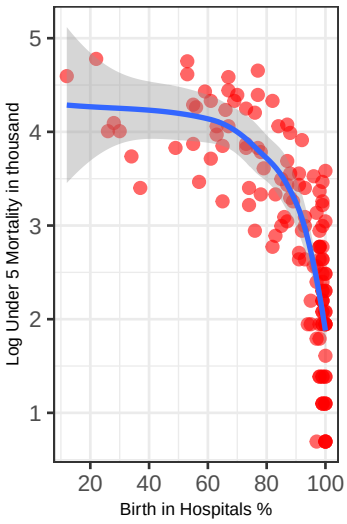


Above are the general patterns compared to Under 5 Mortality of the two variables I didn't include in my main report: Boys' out-of-school rate and GDP per capita. Boys' absence from school shows very similar pattern to girls' but even with the skewed group of countries at lower rates there already seems to be an increasing pattern. Once again, only a couple countries above 45%. Taking the logarithm of GDP makes the data less skewed and the pattern is fairly clear, higher GDP is associated with lower Mortality, although the pattern is getting weaker with higher log GDPP.









Above are the general patterns if Under 5 Mortality values is transformed with logarithm. This is technically possible since the lowest value is 2%. Education data is somewhat more spread out along y axis but the general patterns are very similar. Notably GDPP trend is even more linear while birth hospitalization is even more left-skewed. These are the expected outcomes for log transformation.

Other Models: logged Mortality, GDPP-Girls Education interaction

Below are the same models from the main report with logarithm of Under-5-Mortality as explained variable. You will see that the sign of associations is the same. However, the significance of findings is different. Since the transformations resulted in a strong linear relationship between GDPP and Mortality, this finding will be the most significant and education variables become less significant. Indeed, boys' education rate is only significant in model 2 with no control variables and in model 4 even girls education rate is not significant. Relationship between hospitalization is stronger in this model. Since Education variables are rounded percentages, they may also be further transformed with logarithms after substituting 0 values to a low

value. I decided not to do that since that decision would be arbitrary as many countries really do have virtually no absence from primary schools. Another approach I considered was to transform the percentage values into absolute number of kids out of from school and taking the logarithm of such a measurment. Here subsituting would cause less distortion, however, I couldn't find realiable source for number of primary school age children in the countries.

Table 5: Models with log Under-5-Mortality rates

	(1)	(2)	(3)	(4)	(5)
Intercept	2.04***	2.43***	2.05***	6.41***	7.56***
	(0.230)	(0.112)	(0.232)	(0.482)	(0.544)
Girls Out-of-School rate (<3 , primary)	0.217*		0.228*	-0.009	-0.076
	(0.090)		(0.095)	(0.052)	(0.051)
Girls Out-of-School rate (>=3, primary)	0.042***		0.051**	0.012	0.015*
	(0.007)		(0.016)	(0.009)	(0.007)
Boys Out-of-School rate (primary)		0.048***	-0.011	0.0007	-0.002
		(0.007)	(0.018)	(0.010)	(0.011)
Ln GDP per capita				-0.457***	-0.582***
				(0.044)	(0.078)
Institutional delivery (<67)				0.013**	0.012***
				(0.005)	(0.003)
Institutional delivery (>=67)				-0.027***	-0.020**
				(0.005)	(0.007)
Observations	131	131	131	131	131
R2	0.370	0.311	0.372	0.793	0.870

In addition, it is hard to say how much a country's GDPP depends on the education of women, it is possible that association is not shown here because GDPP is not an unbiased control variable for this question.

To investigate this connection, I considered building a model with an interaction term between GDPP and girls' education. However, such a model would have too much correlation between its variables to produce significant results.

Based on this difference I find the logarithmic results important to include here. Further analysis is definitely required. However, girl's education is still close to significance here and with the potential issues of this model I am still more convinced by the findings of my main report.

List of Countries in samples

Finally, find here the list of countries included in the 3 different samples:

- Sample 1 is the main model (models 1 to 5).
- Sample 2 is the model for secondary school absence rates (model 6).
- Sample 3 is the model for primary school completion (model 7).