

Why I support same language subtitling

2020-06-07

In the fall, India's ministry of information and broadcasting announced that it would require all major TV channels to add "same language subtitles" (SLS) – i.e. subtitles in the same language as the content to the programming. The rules stipulate that channels must have at least one program with subtitles in 2019 and ramp up to 50% of programming with subtitles by 2025.

The motivation for the rule is that it will (hopefully) help children learn to read. The change in rules was the result of years of tenacious advocacy by PlanetRead, an organization which pushes the cause of same language subtitles. (For the story behind this advocacy, check out this recent piece by PlanetRead's founder Dr. Brij Kothari in SSIR.)

I think that this is great news. While I am not entirely convinced by the existing evidence for SLS (see below), I don't think it matters – SLS is cheap and easy enough that countries should just do it.

Evidence for SLS

This page by turnonthesubtitles, a group dedicated to promoting SLS, summarizes the evidence for SLS. Several studies have found that people watching SLS videos pay attention to the subtitles (as measured by tracking subjects' eye movements). But there are few studies which estimate the impact of SLS on reading ability and those that do don't have particularly strong study designs. According to the review, one of the strongest pieces of evidence of SLS's impact on reading comes from Maharashtra where Zee Talkies, with the support of PlanetRead and USAID, added subtitles to all of the songs of many of the Marathi movies it broadcast between 2013 and 2015. In The Huffington Post, Kothari uses ASER data to argue that Zee's use SLS likely increased reading skills.

Re-analysis of the Maharashtra data

SLS may have caused an increase in ASER reading scores in Maharashtra, but I don't think the evidence supports this claim. The first graph below shows the proportion of rural fifth graders in Maharashtra who can read a grade 5 text by year. The second graph shows rural grade 5 reading scores over time for all states. (While I focus on grade 5 scores to avoid clutter, grade 3 scores, the only other scores I have for the full time period, exhibit very similar trends.) ASER data used for this analysis can be found [here](#).

The figures show that in most states, there are a **lot** jumps in ASER scores from year to year. Regardless of whether these jumps represent noise or true changes (more on that later), it is clear that using ASER state data to estimate the impact of a policy is dangerous business. Either there is too much noise in the data for the effects to be precisely estimated or else there are so many other changes from year to year that isolating the effect of just one policy is a fool's errand. (Note that ASER sampling sizes for nearly all states are quite large and thus sampling error can only explain a small portion of these jumps.)

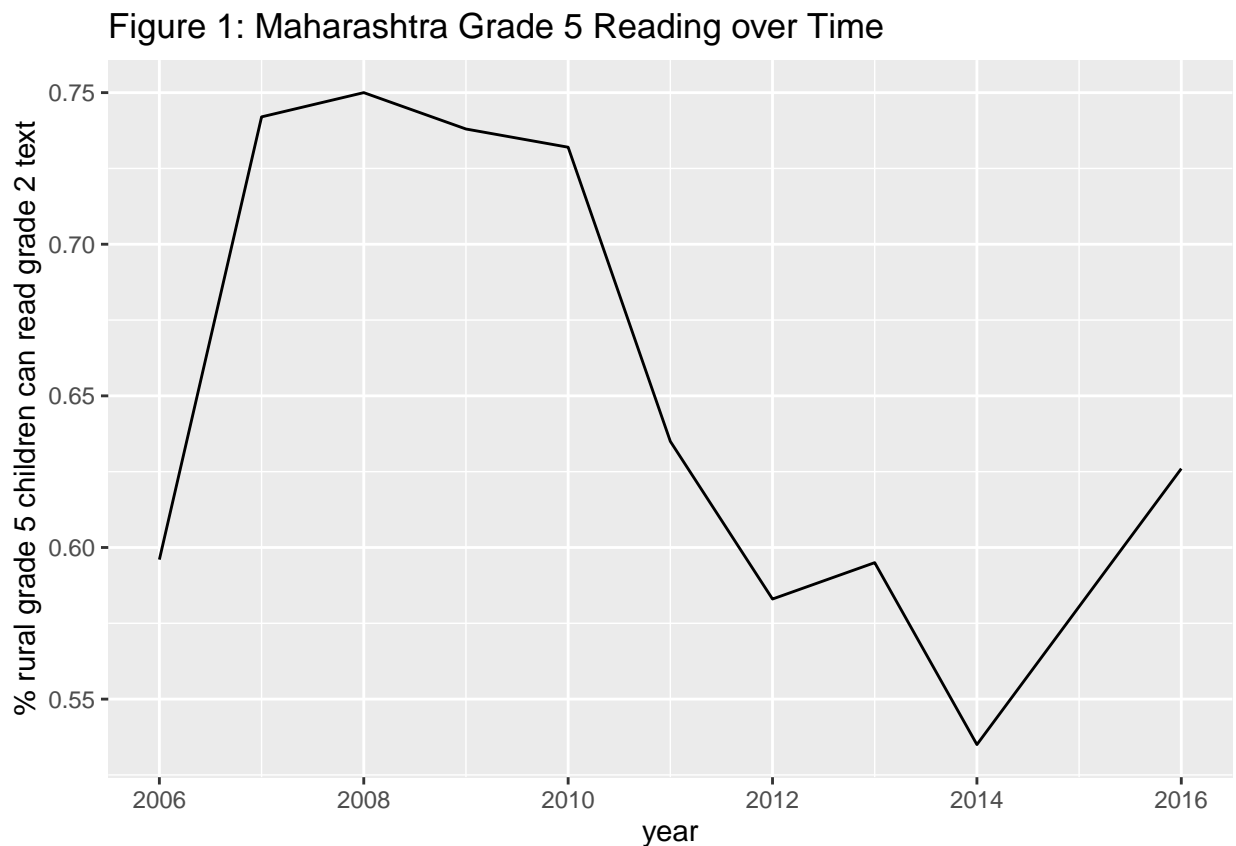
In case you need more evidence, the third graph zooms in on the period from 2012 to 2016 and shows the difference in math and reading scores for grade 3 and grade 5 students by state. This graph shows that Maharashtra's slight increase in reading scores between 2012 and 2016 is not that unusual. In addition, I have performed a more rigorous analysis of these data using the synthetic control methods (I will share that

soon), but the key takeaways are the same as you would get just from looking at these graphs: the data is too noisy to show that SLS led to an increase in reading in Maharashtra over this time.

This is not to say that SLS in Maharashtra didn't have an effect. It may very well have. But, in my opinion, ASER data don't support that claim. I should also mention that PlanetRead commissioned a one-off ASER survey in select districts to estimate the effects of this program. Those data may very well reveal a different story.

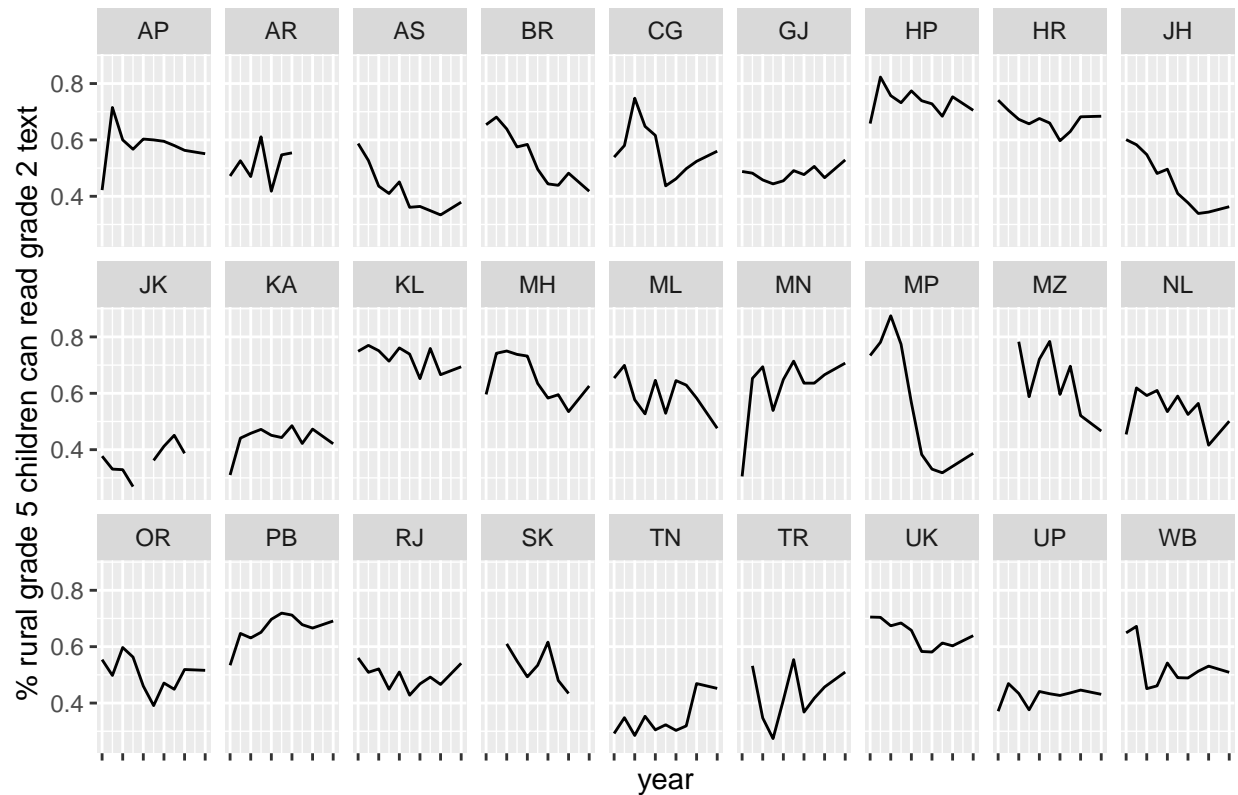
```
library(tidyverse); library(Synth)
aser <- read_csv("https://raw.githubusercontent.com/dougj892/public-datasets/master/ASER%20trends%20over%20time.csv")
mutate_at(vars(-state_abbr, -State), as.numeric) %>%
  select(year, State, state_abbr, std3_read = std3_read_std1_all, std5_read = std5_read_std2_all, std3_read_2018 = std3_read_2018)
  filter(year != 2018)

ggplot(subset(aser, state_abbr == "MH"), aes(x = year, y = std5_read)) +
  geom_line() +
  labs(title = "Figure 1: Maharashtra Grade 5 Reading over Time", y = "% rural grade 5 children can read grade 2 text")
```



```
ggplot(aser, aes(x = year, y = std5_read)) +
  geom_line() +
  theme(axis.text.x = element_blank()) +
  facet_wrap(~ state_abbr, nrow = 3) +
  labs(title = "Figure 2: Grade 5 reading scores over time by state", y = "% rural grade 5 children can read grade 2 text")
```

Figure 2: Grade 5 reading scores over time by state



```
# Generate dataset of deltas from 2012 to 2016
diffs <- aser %>% filter((year == 2012) | (year == 2016)) %>% group_by(State) %>%
  mutate_at(vars(starts_with("std")), ~.x-lag(.x, order_by = year)) %>%
  filter(year == 2016) %>%
  ungroup()

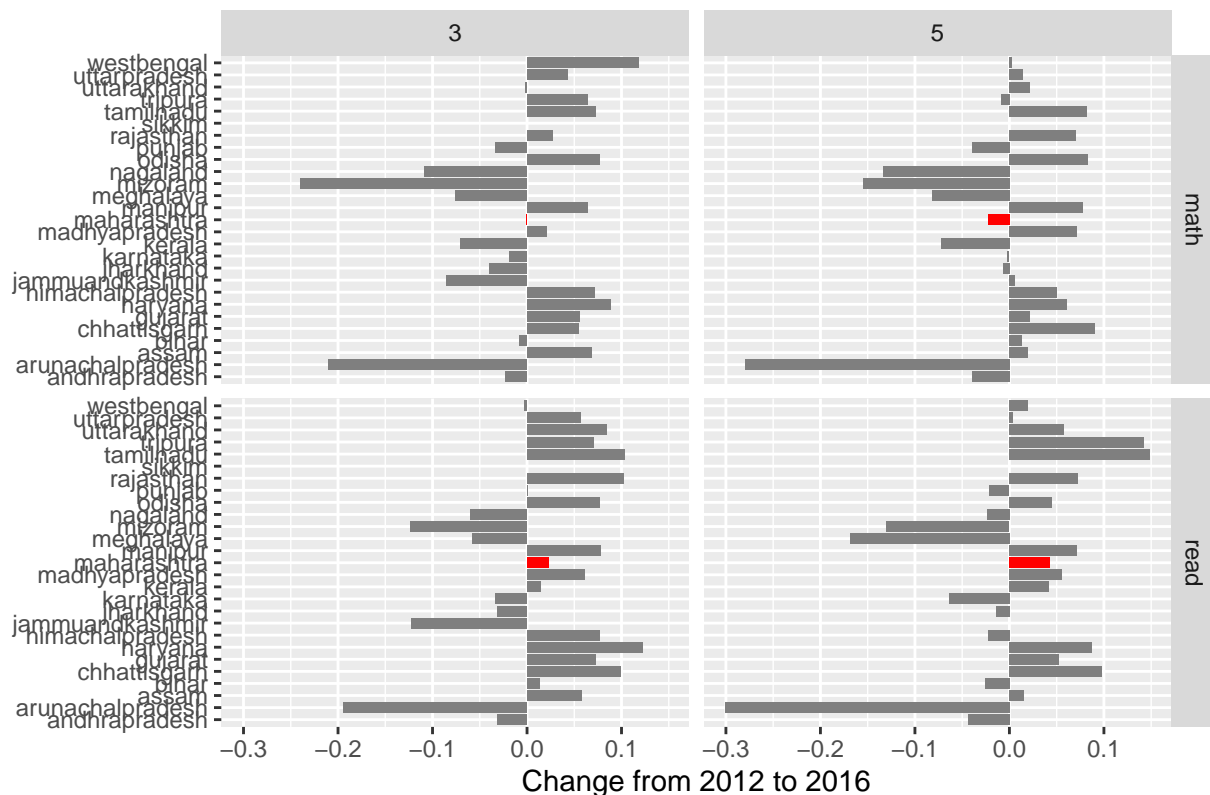
diffs_long <- diffs %>% pivot_longer(starts_with("std"), names_to = "temp", values_to = "diff") %>%
  mutate(class = str_sub(temp, 4,4), subject = str_extract(temp, "(math|read)"))

# plot the reading changes for 2016
ggplot(diffs_long, aes(x = State, y = diff, fill=factor(ifelse(State=="maharashtra","Highlighted","Normal")))) +
  scale_fill_manual(name = "area", values=c("red","grey50"), guide= FALSE) +
  geom_bar(stat = "identity") +
  facet_grid(subject ~ class) +
  coord_flip() +
  labs(y = "Change from 2012 to 2016", x =element_blank(), title = "Figure 3: Change in ASER score 2012-2016")

## Warning: Removed 5 rows containing missing values ('position_stack()').

## Warning: The 'guide' argument in 'scale_*()' cannot be 'FALSE'. This was deprecated in
## ggplot2 3.3.4.
## i Please use "none" instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```

Figure 3: Change in ASER score 2012–2016 by state, subject and



Why I like SLS despite the (imo) lack of evidence

The reason I really like SLS despite (imo) the relatively weak evidence is that it is easy and cheap, two great features in an education intervention.

First, SLS is very easy (for the government at least). Mandating SLS is, in Lant Pritchett’s useful terminology, a prototypical “thin” task. That is, the government mandates that it be done, checks that TV channels are complying, and hands out penalties if channels don’t comply. I’m over-simplifying somewhat (there are a lot of TV channels in India), but as long as the government takes its own rule at all seriously, we can be fairly confident that SLS will reach children. In addition, because it is relatively easy to enforce, it doesn’t crowd out other education interventions by taking up the time of teachers, BEOs, DEOs, and state-level officials. As Pritchett shows, there are very few education interventions like this. For example, if you want to change some aspect of classroom pedagogy, you need to figure out how to train the teachers, motivate the teachers to actually make the change, motivate BEOs to support and monitor teachers as they make this change, and motivate the district officials to coordinate and lead this effort. This is not impossible, but it takes a lot of effort and leaves little space for other important reforms.

Second, in terms of dollar per instructional time, SLS is ridiculously cheap. Kothari estimates that the initial first year cost of adding SLS to programming across India would be around USD 15 million and that recurring annual costs would be about USD 1 million (since, in the future, producers could simply demand scripts for content aired). About two thirds of all Indian households, and one third of the poorest Indian households, own a TV. (See our Scroll article for more info.) There are roughly 200 million primary age children in India so even if the SLS programming only reaches 20% of these children it’s steady state cost is only \$1 per child per year. (This assumes that SLS only benefits primary school children but there may be benefits for low literacy students in higher grades and adults as well.)

If you agree with the argument that SLS is relatively easy to implement and that it doesn't crowd out other efforts to improve education and you also agree that there is at least a sliver of a chance that SLS works, then SLS is almost certainly cost effective. Hanushek and Woessman (2010) show that differences in education level are highly predictive of future economic growth. In particular, they estimate that going from the education level of Brazil to Mexico or Mexico to Romania results in 2 percentage points more growth over the next 40 years. These exact figures (like any results from cross country regressions) should be taken with a grain of salt but they clearly show that any program which at least has a hope of increasing learning at a national level and which doesn't impose significant costs (either financial or in terms of crowding out other reforms) is likely to be cost effective.

Let's say you believe that there is a one in a hundred chance that SLS leads to learning gains which is one thousandth of this gap and that, absent SLS, the growth rate would be 5% over the next 30 years. Then the expected increase in GDP 30 years from now due to SLS is

$$E(SLS) = (2.72e12 * (1.05 + .02 * .001)^{30} - 2.72e12 * (1.05)^{30}) * .01 = USD67.2million$$

Again, these are very crude calculations, but if you think there is a sliver of a chance that SLS works it is almost certainly cost effective.

But shouldn't we wait for results from an impact evaluation?

I like impact evaluations. In fact, I have spent most of the past 10 years or so conducting impact evaluations. And I believe that more research on SLS is desperately needed (so that, for example, we can figure out which children / adults it should be targeted to and what types of programs, if any, it works best with). But I believe that it would be a mistake to wait on results from an impact evaluation before scaling up SLS.

Given the extremely low cost of SLS (if implemented at scale) it would take an enormous impact evaluation to determine whether SLS has sufficient impact. The smaller the effect size you want to detect, the larger the sample size you need for an impact evaluation and SLS would be cost effective even if its effect is miniscule. Running an impact evaluation of that size would take a lot of time and be very expensive (probably more than USD 1 million). Rather than ex ante conduct an impact evaluation and wait for the results, a better option would be to scale it up and then conduct a large survey to try to gauge whether SLS has had an effect on learning. These studies are much less rigorous than an RCT, but, if done well (in particular, if at least 3 data points in time are collected) can do a reasonably good job of estimating impact.