Enrollment Plots

Neha Narayan

Process

Some of the enrollment variables have to be constructed. The data from 2018-19 onwards doesn't include a total enrollment variable for any categories. Therefore, I use the following process for this later data:

- For each class (grade), the data includes counts for each age. For instance, I have counts of number of 6-year-old boys in grade 1, number of 7-year-old boys in grade 1, etc.
- So to get the total number of boys or girls in each class, I add up the counts for each age group within a given class. I do the same to get counts for, for instance, total SC (scheduled caste) boys in class 1.

To calculate total primary + upper primary enrollment for girls/boys, I add the number of boys in each class from 1-8, and same for girls. To calculate total regardless of sex, I add the girl total and the boy total.

Since the treatment happens at the state-year level, I collapse to the mean enrollment across schools in a given state-year to generate my plots.

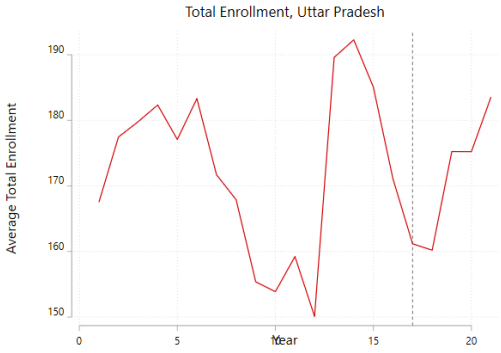
All plots are generated on a 1 percent sample (N = 275,760).

Results

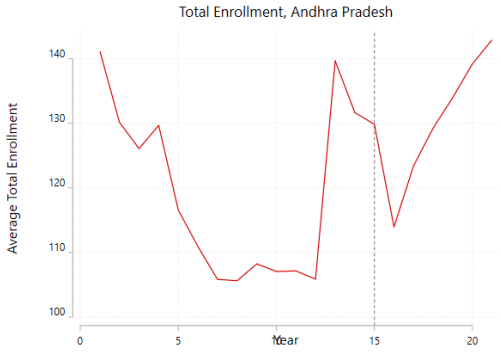
In **Figure 1**, I plot mean enrollment for all students across all schools in four different states: Uttar Pradesh, Andhra Pradesh, Maharashtra, and Assam. These states represent the north, south, east, and west of India respectively. Key takeaways include:

1. As expected, estimates are very noisy.
2. Noticing a huge spike at 2013-14, I plotted a few more randomly chosen states, and noticed the same spike.
3. Because I remembered that variable names changed a bit from 2012-13 to 2013-14, even though the 2005-06 through 2017-18 data comes from the same source, I went back to look at the spreadsheets. I realized that I don't have data for classes 9, 10, 11, and 12 for the years 2005-06 to 2012-13, so the years prior and after are incorporating more classes.
4. Since classes 1-8 are the ones who are eligible for the MDM, I am deciding to restrict enrollment in my preliminary graphs to these classes only across years.

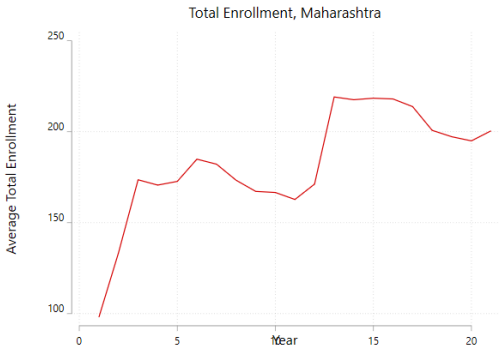
Figure 1: First Pass Enrollment Graphs



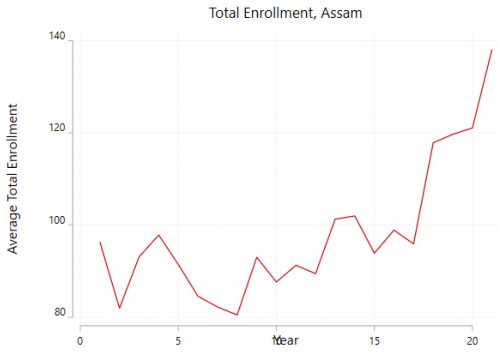
Uttar Pradesh is in Northern India.



Andhra Pradesh is in Southern India.



Maharashtra is in Western India.



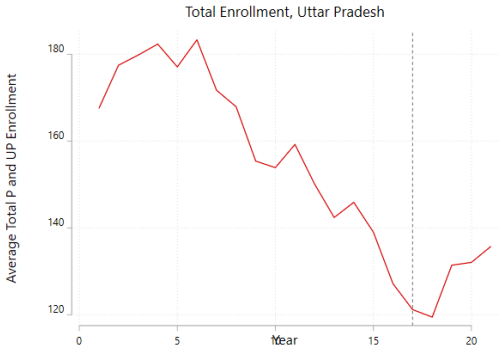
Assam is in Eastern India.

Notes: The grey dotted vertical line represents treatment timing. Maharashtra and Assam are control states, so no dotted line is present.

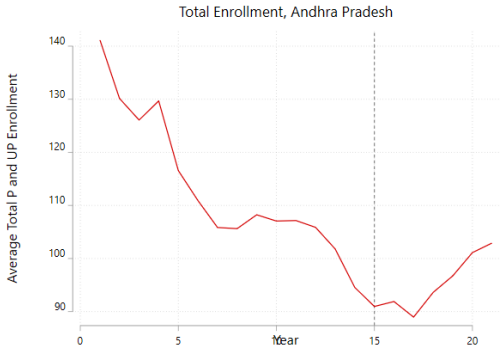
In **Figure 2**, using my new definition of total enrollment as all kids in primary and upper primary grades (1-8), I create the same plots as in Figure 1. Key takeaways include:

1. These graphs look a bit more sensible to me in that there's no sudden spikes in the middle of the graph.
2. However, I'm pretty confused about the consistent downward trend in enrollment, especially because I don't think that's what was actually happening during this time.
3. By that I mean India's population consistently increased in the relevant period, so I don't think this is a reflection of population. Also, other pro-education policies were passed during this time period – for example, the midday meal became national policy in 2001-02, and the right to education act was passed in 2009 – so I'd expect the opposite trend of what I see here.
4. I tried plotting a few more states to see if this was a pervasive issue; Tamil Nadu, Himachal Pradesh, Madhya Pradesh, Odisha all have similar trends. For context, I didn't plot every single state, and these states were chosen because they fall in the middle of the population distribution.

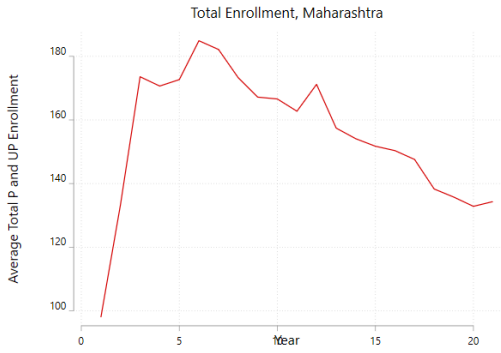
Figure 2: Second Pass Enrollment Graphs



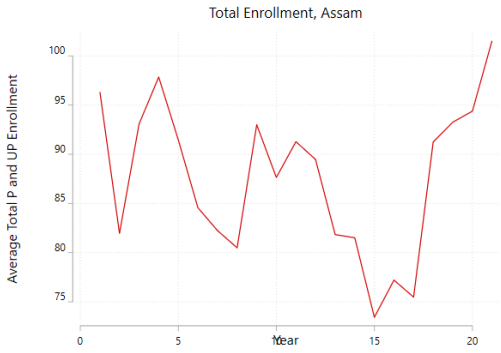
Uttar Pradesh is in Northern India.



Andhra Pradesh is in Southern India.



Maharashtra is in Western India.

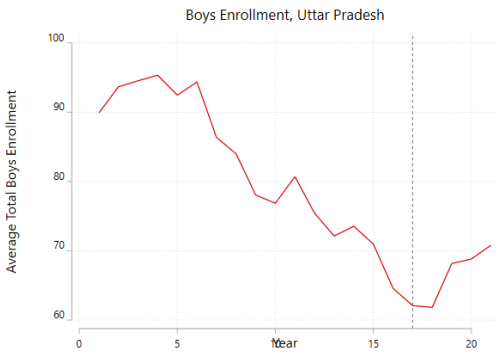


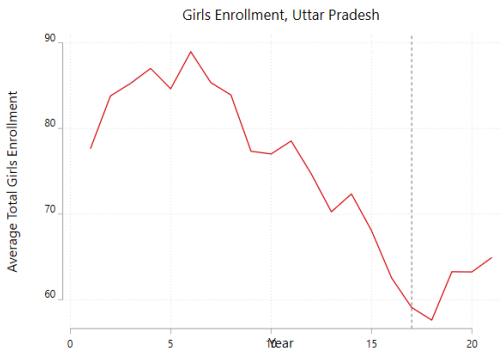
Assam is in Eastern India.

Notes: The grey dotted vertical line represents treatment timing. Maharashtra and Assam are control states, so no dotted line is present.

Finally, to confirm that one sex was not driving the downward trend, I split the enrollment plot for Uttar Pradesh by gender in **Figure 3**. The graphs are basically identical and both have a downward trend, so there's not some gender specific shock causing the aggregate graph to slope downward.

Figure 3: Enrollment Graphs by Gender





Next Steps

1. Once I dig deeper into the causes of some of these unusual enrollment trends, I'll want to make plots by caste and tribal status as well. The challenge is that my data from 2001 to 2005 doesn't go as granular as enrollment within a caste-gender-class; it only has caste-class or gender-class. However, a lot of schools are missing from this earlier data anyway, so I don't think I'll spend too much time trying to accomodate the limitations of the pre 2005 data.

2. I'll put together population figures for the states and years, split by relevant groups, so that I can plot versions of these graphs scaled by population. However, as mentioned earlier, I'm not sure if population scaling will resolve my questions about the strange trends.