**SOLVING UTTAR PRADESH’S GROWTH PARADOX**

**Abstract**

Uttar Pradesh, India with a population of 200 million + (India Census, 2011), sixth most populous region of the world, and a per capita income of ~ USD 990, second from the bottom in India, represents a paradox that have left many economists wondering why the state, with its high demographic dividend, hasn’t been able to step out of the Malthusian trap. With a falling women labor force participation and high human capital migration outside, the state deserves our attention as an economic problem that only few represent. My intention from this paper is to understand what is happening in Uttar Pradesh from the lens of its economic factors and start exploring possible avenues of research. The paper is not at all causal and the only regression done is more to understand the role of power in attracting industries.

**Research Question**

Considering how broad the project is there are multiple avenues of approaching this problem. Few key questions that I had in mind when starting it were:

1. How has Uttar Pradesh developed economically over the years in comparison to other states?
2. What is the role of politics in affecting development in Uttar Pradesh? Particularly, who are the representatives, what does it take them to elect and what kind of policies do they enact?
3. Where is Uttar Pradesh placed in the wider investment sentiment towards India?

**Data Scraping and Cleaning**

All of the data that I have worked with has been scraped from myneta.info (a website that monitors elections in India) and Reserve Bank of India, India’s central bank. The data thus represents authentic sources that have been cited actively in research.

* **Scraping and Cleaning Political Data**

This was the first step in the project. I wanted to understand the assets and liabilities of the politician and then create a cross-walk of this data with data from [Data Development Lab](http://www.devdatalab.org/). Towards this I focused on three years for which the data is available, though the formats being different I couldn’t run a function for scraping all three. But while I was able to scrape the website and create formats that could have been used for analysis, I wasn’t able to move forward because as of now it is rather magnanmpouos task to create a cross-walk of political data and economic data. Only recently has research started focusing on it for India (Coordinating Voters against Criminal Politicians, 2018). The reason for this is that the constituencies are different for politics and economic census, thus a politician responsible for one constituency cannot be easily mapped to the outcomes of its constituency. Realising this I had to drop the idea, but I have still committed my code for it. My next step was to look at Uttar Pradesh’s economy from a more zoomed out lens.

* **Scraping RBI data and Cleaning**

Towards the process of understanding what is happening, I scraped RBI’s website for data on fiscal, industry, agriculture, socio-economic indicators, price and wages (https://rbi.org.in/Scripts/AnnualPublications.aspx%3Fhead%3DHandbook%20of%20Statistics%20on%20Indian%20States); this process was then followed up by data engineering to understand what are the relevant areas to look at and explore further. I first cleaned the data for datasets that state as there rows and then datasets that year as their rows. Finally I merged these two datasets together to create the final data set for the analysis.

I would like to mention here that this was the major part of my project and I had to spend extensive time on scraping and cleaning to get the data to a usable format.

**Analysis**

* **RBI Data**

For my analysis I focused on understanding what are the outcomes that are visible. In particular I looked at the rural wage rate and how it has changed over the years, how has the economic activity from manufacturing changed over the years and how has number of factories evolved. While I was doing this I wanted to keep the lens of what is happening in rest of India and thus I kept all states part of my analysis. The signs were clear, UP was way behind in all the indicators and these indicators in a way explain why migration outside of UP is so high, with low manufacturing and number of factories while neighboring states offering more opportunities and higher wages, it made sense.

* **IMF’s Sentiments**

The next step for me was to understand what has been business climate/ political risk sentiment towards India and thus I did a sentiment analysis of International Monetary Fund’s Article IV reports which provide consultation to countries. I wanted to place how have sentiments towards India changed over the last 10 years. The reports were not available for 2012 and 2014. This exercise didn’t seem that fruitful as the general sentiment distribution was similar in the last 10 years and was quite balanced.

* **Infrastructure Analysis**

In the end I moved on to analysis of the infrastructure in Uttar Pradesh and how it might have affected development. The proxy for development that I use is manufacturing and the indicators for infrastructure include highways, gap in power availability and requirement and installed capacity of power. The result is statistically significant for gap in power availability and the total r-square as would be obvious is .90+. I don’t take this as a causal estimate as there are huge confounding effects in terms of government policies, fixed effects, and other data generating processes that might not be visible. But what the analysis does allow is to surface one of the key areas that might be holding UP back. When I compare other states to Uttar Pradesh, I see how states like Maharashtra and Andhra Pradesh are doing well on these indicators.

**Next Steps**

A next step for me from here is to start exploring the fiscal angle for these dynamics and to further build a dataset which allows me to create an analysis of how politics impacts economic development.