# Macroeconomics of Development

Data Assignment 2 (15 points)

Amit Basole

March 7, 2022

## World Development Indicators data

The dataset contains data for a maximum of 266 countries over 60 years. However, many variables are available for a much shorter period of around 30 years. Data are available on sectoral shares in output and employment as well as GDP per capita. Two additional variables are worth noting - "not tiny" defines countries that have over 1 million in population. And "not high inc" defines countries with annual GDP less than USD 13,000 per capita.

## Structural change

- Cross-sectional relationship between agriculture's share in output or employment versus GDP per capita
    1. For the year 2015, make a cross-country scatter plot of the share of agriculture in GDP versus GDP per capita in PPP terms. Remember to log transform GDP per capita. Limit your sample to not tiny and not rich countries.
    2. Regress agricultural share of GDP on ln GDP per capita (PPP), once again limiting your sample to not tiny, not rich countries. What is the interpretation of the coeffieicent you get on log GDP?
    3. Fit a regression line to the scatter plot you made in 1. Locate India on the plot visually.
    4. Repeat the above steps, 1 through 3 for share of agriculture in employment.
- Cross-sectional relationship between self-employment share in total employment and GDP per capita
    1. For the year 2015, make a cross-country scatter plot of the share of self-employment in total employment versus GDP per capita in PPP terms. Remember to log transform GDP per capita. Limit your sample to not tiny, not rich countries.
    2. Regress self-employment share on ln GDP per capita (PPP), once again limiting your sample to not tiny, not rich countries. What is the interpretation of the coeffieicent you get on log GDP?
    3. Fit a regression line to the scatter plot you made and locate India on the plot visually.
- What difference do you notice with respect to India's location in the self-employed share graph versus the one you made for agriculture's share in employment versus the one for agri share in GDP? Reflect on your finding.
- Changes in agri and self-employment shares over time for selected countries.
    1. Plot a line graph of agri share of employment over time (since 1991) for India and Vietnam.
    2. Plot a line graph of self-employment share of employment over time (since 1991) for India and Vietnam.
    3. The time trends for the two variables differ in one key respect. What is that? Can you link what you observe here to your reflections in the previous question?
- Panel regression with country fixed effects
    1. Now we will take advantage of the panel aspect of the dataset.
    2. Run a regression of agriculture's share of employment on ln GDP per capita (PPP) with country fixed effects (dummies). Use India as the base category.
    3. Run an identical regression with self-employment share instead of agriculture's share.
    4. How will you interpret the coefficient on ln GDP per capita in each case? How can you compare India's performance to Vietnam's in this framework?

## Premature deindustrialisation

- Let us try to replicate Figure IV from Amirapu and Subramanian (page 10). Instead of the years they choose (1988, 2000, 2010) we will choose 1991, 2001 and 2011.
- To replicate this figure you will need to make three scatter plots and fit a quadratic function to each. Interpret what you see.
- Next we will run a panel country fixed effects regression with period dummies as in Rodrik (page 9). On the LHS, you should have industry share of employment. On the RHS you will have log population, square of log population, log GDP pc (PPP), square of log GDP pc (PPP), and country dummies, with decade dummies for the 1990s, 2000s and 2010s (we don't have data prior to 1991).
- Compare the coefficients you get for GDP and GDP square with what Rodrik gets (Table 2). How do you interpret both sets of coefficients?