Great gatsby curve

**Feedback from Saez**

Multiply each household by a weight to make the total distribution.

Replace the top so many percent with tax data entries.

Derive income of the tax data people.

**Types of Data**

Income

Expenditure / Consumption (Do pensions count??)

Tax Data (there may be nuances – did people pay taxes but not file returns??)

(tax data, national accounts, surveys)

GDP Data ( see 2.1.3 in India paper). This

IHDS data equivalent (other sources like the survey which serve as backup. Note that government benefits do NOT count)

UN statistics population data

Census data

🡺 First copy the data / strategies that are used in the paper – then look for alternatives

- Poverty and Growth in India Database of the World Bank (Tabulations for non-HEIC / HIIC data

* National Accounts (WID)
* National Accounts (PBS)
* Tax Data…
* Other Survey Data…

Goal for tomorrow: Get all the data sources in PANDAS notebook, cleaned a formatted and explainable.

Next step: Learn how to combine the data ‘with national accounts data’ to get percentages. What’s wrong with existing percentages on income within the survey? Is there some way we transform / get rid of the units for each of the datasets.

🡪 Also spend some time figuring out where World Bank got their numbers from

WID National Accounts

PBS National Accounts (double check)

**Data Sources to Explore**

World Bank

<https://data.worldbank.org/indicator/SI.DST.10TH.10?end=2015&locations=PK&start=1987&view=chart> \*\*\*\*\*\*\*\* Double check against HEIS / HIIC data

WID

Luxembourg Income Study database

HEIS / HIIC Data

Tabulations + GPinter

Forbes ‘Rich Lists’

Poverty and Growth in India Database of the World Bank (but for Pakistan)

NSSO = PBS lower level organization. What is that?

Existing openly available data sources

Data sources from other papers

UN-WIDER WIID

World Bank Pakistan Surveys (based off of PBS - insertdata)

Beaconhouse National University

Lahore School of Economics

<https://library.fes.de/pdf-files/bueros/pakistan/14113.pdf> (xi, xiii, xiv)

India Paper - [in particular data from the National Council for

Applied Economic Research (NCAER) and from the Inter University Consortium for Applied Political and Social Sciences Research

(ICPSR). These data sources however do

not enable comparison before and after 2000.]

World Bank, Development Research Group. Data are based on primary household survey data obtained from government statistical agencies and World Bank country departments. Data for high-income economies are from the Luxembourg Income Study database. For more information and methodology, please see PovcalNet (http://iresearch.worldbank.org/PovcalNet/index.htm).

Ravallion and Chen (1996) on estimation for low and middle income economies



The above surveys.

Was the Central Statistical Office (1947), then in 1950 a part of the Economic Affairs Division. In 1972 became full government department. In 1981, Federal Bureau of Statistics. IN 2000s became the

Mahbub ul Haq (the 22 families). **READ ALL HIS RESEARCH**

<https://paycheck.pk/salary/celebrity-income/richest-families-in-pakistan/the-22-families/>

Formal vs informal sector is 30:70…

<https://en.wikipedia.org/wiki/Category:Pakistani_economists>

Burki Uncle??? The other person mom took me to meet.

**WORDS**

FORMAT (INDIA PAPER)

1 ENTERING THE DIGITAL AGE WITHOUT INEQUALITY DATA

* 1. ECONOMIC POLICY SHIFTS SINCE THE 1980s
  2. THE INCOME INEQUALITY DATA GAP

The shortcomings of household survey data in monitoring the evolution of inequality are well known; because of underreporting and undersampling issues, surveys fail to properly capture inequality dynamics at the top of the distribution (Atkinson and Piketty, 2007, 2010).

1. DATA SOURCES AND METHODOLOGY
   1. DESCRIPTION OF THE DIFFERENT DATA SOURCES
      1. TAX DATA
      2. NSSO CONSUMPTION DATA
      3. NATIONAL ACCOUNTS DATA
      4. IHDS INCOME AND CONSUMPTION DATA
      5. UN STATISTICS POPULATION DATA
   2. METHODOLOGY
      1. ESTIMATION OF TOP FISCAL INCOMES
      2. ESTIMATION OF BOTTOM SURVEY INCOMES
      3. INTERPOLATING SURVEY AND TAX DATA FOR MISSING YEARS
      4. COMBINATION OF TAX AND SURVEY DATA
      5. FROM TOTAL FISCAL INCOME TO NATIONAL INCOME
      6. DEFINITION OF A BENCHMARK SCENARIO
2. RESULTS
   1. SHARP RISE IN TOP INCOME SHARES SINCE MID-1980S
   2. FALL IN MIDDLE 40% AND BOTTOM 50% SHARES
   3. TOTAL GROWTH RATES BY INCOME GROUP
   4. GROWING SHARE OF INCOME GAP EXPLAINED BY TOP INCOMES
   5. MEASUREMENT ISSUES AND ROBUSTNESS TESTS
3. DISCUSSION
   1. THE MID 1908S TURNAROUND
   2. SHINING INDIA FOR THE RICH MOSTLY?

5 CONCLUSION