Reproducibility Check

AUTHOR

Ayush

What is this document?

Here in this doc I will try and reproduce some of the results in the World Bank paper "Intent to implementation: Summary of lessons from tracking India's social protection response to COVID-19".

Doing this will help establish if I am applying the weights correctly for analysing the CMIE's data.

Which datasets are is used?

> **Note**
> WB has confirmed that the numbers in the paper were generated by using responses to CSM 1 only.

We will be using the responses to the household frame to this module as the questions/results of interest are in HH frame.

Reproducing the results

In order to reproduce the results it is necessary to note the following from the paper:

> **Important**
> "For analytical tractability, households are classified as those that received food only, cash only and both food and cash benefits. Food benefits comprise of access to PDS rations while cash comprises of transfers received under all PMGKY schemes except PDS and provident fund withdrawals. NREGA related cash receipts were not included in the survey module and are excluded in this analysis. Additionally, EPF withdrawals are excluded in the analysis (even though the survey captures this information) because of the smaller proportion of households that are enrolled in the program."

Read CSM1

Here the data is imported from a `csv` file and the only observations with `response_status == "Accepted"` are preserved. Consequently, the new variable `w_cntry` is created as a product of `hh_weight_for_country_w` and `hh_non_response_for_country_w`.

A small peek into the data set provided below.

▶ Code

hh_id	state	hr	region_type	month_slot	reason_for_non_response	district	response_status	family
17673997	Jammu &	HR 1	RURAL	Jun-20	No Failure	Baramula	Accepted	N

hh_id	state	hr	region_type	month_slot	reason_for_non_response	district	response_status	family
39288720	Jammu & Kashmir	HR 1	RURAL	Jun-20	No Failure	Baramula	Accepted	N
66805895	Jammu & Kashmir	HR 1	RURAL	Jun-20	No Failure	Baramula	Accepted	N
23664907	Jammu & Kashmir	HR 1	RURAL	Jun-20	No Failure	Baramula	Accepted	N
15332843	Jammu & Kashmir	HR 1	RURAL	Jun-20	No Failure	Baramula	Accepted	N

Reproduce: "Across the country, nearly 74% of all households received food through PDS allocations, 40% of households received cash-transfers."

▶ Code

```
# A tibble: 1 × 2
  food_ben_perc cash_ben_perc
          <dbl>         <dbl>
1         0.740         0.403
```

> **Sucessful replication for:**
>
> Across the country, nearly 74% of all households received food through PDS allocations, 40% of households received cash-transfers. The reproduced results are not show in % format.

Reproduce: "While nearly 85% of all rural households' report accessing food or cash benefits following the pandemic, only 69.2% of urban households report the same."

Assuming here by "accessing food or cash" it is meant that HHs that can access either one or both.

▶ Code

```
# A tibble: 2 × 2
  region_type perc_food_or_cash
  <chr>                   <dbl>
1 RURAL                   0.851
2 URBAN                   0.699
```

> **Sucessful replication for:**

> While nearly 85% of all rural households' report accessing food or cash benefits following the pandemic, only 69.2% of urban households report the same.

Reproduce: Graph 1 from the WB Paper

The Graph 2 of the WB paper has to column charts. The one on the left if *Households Accessing Benefits as a share of population* and the one on the right is *Enrollment Pre-Covid*. I will try to generate results of these two graphs.

Graph 1a Households Accessing Benefits as a share of population

▶ Code

```
# A tibble: 2 × 4
  region_type  Food  Cash     PF
  <chr>       <dbl> <dbl>  <dbl>
1 RURAL        78.1  47.0  0.157
2 URBAN        65.6  26.8  1.07
```

> Sucessful replication for:
>
> Graph 1a Households Accessing Benefits as a share of population

Graph 1b Enrollment Pre-Covid

In order to calculate the numbers for this chart, I am making the following assumptions:

- Food, eligible if `type_of_ration_card` is "BPL/Antodaya/Annapurna" or "Other"
- Cash, enrolled in any one or more of Jandhan, LPG, P Kisan or Pension. This is determined by `hh_enrolled_in_*`, except for jandhan.
- PF, determined by variable `hh_enrolled_in_pf_scheme`

▶ Code

```
# A tibble: 2 × 4
  region_type perc_ele_food perc_ele_cash perc_ele_pf
  <chr>               <dbl>         <dbl>       <dbl>
1 RURAL               0.901         0.604      0.0283
2 URBAN               0.813         0.346      0.137
```

> Sucessful replication for:
>
> Graph 1b Enrollment Pre-Covid

Reproduce: Tabel 2 of the WB paper

Following are some of the estimates from the Table2 to reproduce.

Table 2 : Eligible Households Receiving Benefits under PMGKY of WB paper

Partial table is presented

Share of Eligible Households Who have Received Benefits	All-India	Rural	Urban
Food Only	84.6%	89.1%	77.3%
Cash Only	80.2%	81.5%	76.7%
Food and cash Only	70.5%	70.2%	71.8%

I will attempt to reproduce these partial results below:

▶ Code

REPRODUCED::Table 2 : Eligible Households Receiving Benefits under PMGKY of WB paper

Partial table is presented

Share of Eligible Households Who have Received Benefits	All-India	Rural	Urban
Food Only	0.8447841	0.8630615	0.8033929
Cash Only	0.7772320	0.7778194	0.7751396
Food and cash Only	0.7000666	0.6981874	0.7069682

> **Unsuccessful replication for Table 2**
>
> Need to cross check eligibility definitions with WB.
>
> This could be because of how Food only or Cash Only are defined. need to cross check if I have done it the same way as WB paper.