

Issues and Feedback with the PIP data and the API

Michail Moatsos

09 Jun 2022

This document distills and highlights the key quirks/problems with the PIP API that I have discovered.

1 Monotonicity

Getting regional or country distributions at a high frequency (for example by requesting the API to fetch the complete set of household surveys at 10000 or more distinct poverty lines) one gets into trouble because quite often headcount rates are not always increasing or constant compared to the previous (lower) poverty line requested, but a decrease is observed. For example, EAP at 1981 at the poverty line of \$0.612/day. At 0.611 the headcount is 0.1429, but all values returned using poverty lines 0.612, 0.613, 0.614, 0.615, 0.616, 0.617 and 0.618, give a headcount **lower** than 0.1429. This can be verified with a direct call to the API:

```
res = GET(paste0("https://api.worldbank.org/pip/v1/pip-grp?country=all&year=1981&povline=",  
                0.611,"&group_by=wb&format=json"))  
Temp <- fromJSON(rawToChar(res$content))  
Temp <- subset(Temp,Temp$region_code=='EAP')  
res = GET(paste0("https://api.worldbank.org/pip/v1/pip-grp?country=all&year=1981&povline=",  
                0.612,"&group_by=wb&format=json"))  
Temp2 <- fromJSON(rawToChar(res$content))  
Temp2 <- subset(Temp2,Temp2$region_code=='EAP')
```

Which confirms the issue that EAP at 0.612 is 0.1385 which is lower than 0.1429 the reported headcount at 0.611.

Other regional cases, 286 in total, with monotonicity breaks are included in the appendix. Also in terms of country level monotonicity breaks I have counted 973 of them.

A similar situation is also observed, but not investigated, for example with poverty_gap, poverty_severity and watts index at a poverty line of 0.011 for BDInational1995consumption.

2 HHS with more than one mean values

For example, for FJI at national coverage and measuring consumption, for the years 2014 up to 2019, the data contain more than one mean values across various poverty lines. In 2014, those values are 8.8123 and 8.8731.

3 GNBnational1981consumption

GNBnational1981consumption's headcount only goes up to 0.8213 which corresponds to a poverty line of 4.748. This occurs to other GNB distributions as well in a similar manner.

4 Missing variable definitions

The exact definitions for each reported variable used should be easily findable. It is not the case at the moment. For example, the difference between `survey_coverage` and `reporting_level` is rather unclear. Another example is `survey_comparability`. To find out how this is defined one has to look through a number of pages at the worldbank portal, and finally to read “September 2019 PovcalNet Update : What’s New” pdf to understand what is precisely meant by 0, 1, 2, etc.

5 Parallel calls to API

If one attempts to request a high frequency version of the dataset (say by requesting the data from the API at 10000 distinct poverty lines), and this is attempted in parallel mode with 5 or more parallel “workers” or instances, the consequence of this is that several poverty lines do not return a result.

6 Non-unique variable values

EAC 2010 has two means 465.6219 464.6207, and WLD 1995 two population values 5705.720253 5706.720253. But this is not an exhaustive list, there must be plenty more like them on country level as well. Another example is the `reporting_gdp` column returns two unique values 526.3314 and 551.7318 for GIN, national, consumption in 1981.

7 Appendix

7.1 Regional cases of monotonicity breaks

```
## [1] "EAP1981@0.612" "EAP1982@0.666" "EAP1983@0.737" "EAP1985@0.602"
## [5] "EAP1985@2.662" "EAP1986@0.633" "EAP1986@2.798" "EAP1987@0.678"
## [9] "EAP1987@2.993" "EAP1988@0.741" "EAP1988@0.876" "EAP1988@3.276"
## [13] "EAP1989@0.424" "EAP1989@0.902" "EAP1989@1.905" "EAP1990@0.434"
## [17] "EAP1990@0.996" "EAP1990@1.951" "EAP1991@0.389" "EAP1991@0.441"
## [21] "EAP1991@1.05" "EAP1991@1.674" "EAP1991@2.052" "EAP1992@0.401"
## [25] "EAP1992@0.454" "EAP1992@1.062" "EAP1992@1.724" "EAP1992@2.231"
## [29] "EAP1993@0.414" "EAP1993@1.779" "EAP1994@0.389" "EAP1994@0.458"
## [33] "EAP1994@0.89" "EAP1994@1.971" "EAP1994@2.375" "EAP1995@0.424"
## [37] "EAP1995@0.5" "EAP1995@0.97" "EAP1995@2.148" "EAP1995@2.498"
## [41] "EAP1996@0.459" "EAP1996@1.049" "EAP1996@2.62" "EAP1997@0.459"
## [45] "EAP1997@0.466" "EAP1997@1.05" "EAP1997@1.093" "EAP1997@2.597"
## [49] "EAP1997@2.724" "EAP1997@2.808" "EAP1998@0.46" "EAP1998@0.467"
## [53] "EAP1998@1.052" "EAP1998@1.094" "EAP1998@2.72" "EAP1998@2.852"
## [57] "EAP1998@2.94" "EAP1999@0.468" "EAP1999@1.096" "EAP1999@2.909"
## [61] "EAP1999@3.145" "EAP2000@0.497" "EAP2000@0.998" "EAP2000@1.164"
## [65] "EAP2000@3.249" "EAP2000@3.513" "EAP2001@0.515" "EAP2001@1.036"
## [69] "EAP2001@1.208" "EAP2001@3.468" "EAP2001@3.749" "EAP2002@0.32"
## [73] "EAP2002@1.089" "EAP2003@0.573" "EAP2003@1.156" "EAP2003@1.594"
## [77] "EAP2004@0.621" "EAP2004@1.252" "EAP2004@1.726" "EAP2005@0.692"
## [81] "EAP2005@1.923" "EAP2006@0.721" "EAP2006@2.006" "EAP2007@0.766"
## [85] "EAP2007@2.105" "EAP2007@2.131" "EAP2008@2.192" "EAP2008@97.1"
## [89] "EAP2009@2.324" "EAP2010@2.442" "EAP2010@7.31" "EAP2011@3.049"
## [93] "EAP2011@7.17" "EAP2012@3.009" "EAP2012@7.08" "EAP2013@3.692"
## [97] "EAP2013@6.85" "EAP2014@4.058" "EAP2014@7.32" "EAP2015@4.399"
## [101] "EAP2015@4.791" "EAP2015@7.71" "EAP2016@5.08" "EAP2016@7.25"
## [105] "EAP2017@5.75" "EAP2017@7.36" "EAP2018@5.84" "EAP2018@6.74"
## [109] "EAP2019@0.379" "EAP2019@6.23" "EAP2019@8" "ECA1990@4.6"
```

[113] "ECA1990@13.22" "ECA1990@22.08" "ECA1991@4.311" "ECA1991@21.99"
 ## [117] "ECA1992@4.066" "ECA1993@3.858" "ECA2002@0.08" "ECA2003@0.091"
 ## [121] "ECA2004@0.083" "ECA2008@0.011" "LAC2003@36.21" "LAC2011@48.16"
 ## [125] "MNA1985@0.112" "MNA1985@2.406" "MNA1986@0.099" "MNA1986@2.436"
 ## [129] "MNA1987@2.477" "MNA1988@2.526" "MNA1989@2.575" "MNA1989@6.55"
 ## [133] "MNA1990@2.622" "MNA1990@7.42" "MNA1991@2.658" "MNA1991@8.45"
 ## [137] "MNA1992@2.685" "MNA1992@8.34" "MNA1992@154.5" "MNA1993@2.726"
 ## [141] "MNA1994@0.066" "MNA1994@2.772" "MNA1994@3.637" "MNA1994@7.09"
 ## [145] "MNA1995@0.066" "MNA1996@0.063" "MNA1997@0.06" "MNA1998@0.061"
 ## [149] "MNA1999@0.062" "OHI2016@14.74" "OHI2017@15.91" "OHI2018@16.62"
 ## [153] "OHI2019@18.58" "SAS1981@1.179" "SAS1981@1.571" "SAS1982@1.177"
 ## [157] "SAS1982@1.525" "SAS1983@1.19" "SAS1983@1.558" "SAS1984@1.22"
 ## [161] "SAS1985@0.995" "SAS1985@1.248" "SAS1985@1.914" "SAS1985@2.885"
 ## [165] "SAS1986@1.02" "SAS1986@1.268" "SAS1986@1.963" "SAS1986@2.959"
 ## [169] "SAS1987@1.027" "SAS1987@2.978" "SAS1989@47.84" "SAS1993@0.008"
 ## [173] "SAS1994@0.008" "SAS1995@0.008" "SAS1996@0.008" "SAS2002@0.02"
 ## [177] "SAS2003@0.021" "SAS2004@0.021" "SAS2005@0.022" "SAS2006@0.032"
 ## [181] "SAS2007@0.038" "SAS2008@0.065" "SAS2009@0.158" "SAS2010@0.197"
 ## [185] "SSA1990@0.259" "SSA1991@0.259" "SSA1992@0.254" "SSA1999@39.15"
 ## [189] "SSA2004@0.018" "SSA2005@0.019" "SSA2006@0.019" "SSA2007@0.019"
 ## [193] "SSA2008@0.019" "SSA2009@0.019" "SSA2010@0.019" "SSA2011@0.019"
 ## [197] "SSA2012@0.011" "SSA2013@0.008" "SSA2014@0.003" "SSA2015@0.003"
 ## [201] "SSA2016@0.003" "SSA2017@0.003" "SSA2018@0.003" "SSA2019@0.003"
 ## [205] "WLD1981@0.612" "WLD1981@1.179" "WLD1982@0.666" "WLD1982@1.177"
 ## [209] "WLD1983@0.737" "WLD1983@1.19" "WLD1984@0.268" "WLD1984@1.22"
 ## [213] "WLD1985@0.602" "WLD1986@0.269" "WLD1986@0.633" "WLD1986@2.798"
 ## [217] "WLD1986@2.959" "WLD1987@0.678" "WLD1987@2.993" "WLD1988@0.741"
 ## [221] "WLD1988@3.276" "WLD1989@0.424" "WLD1989@0.902" "WLD1989@1.905"
 ## [225] "WLD1989@25.81" "WLD1990@0.434" "WLD1990@0.996" "WLD1990@1.951"
 ## [229] "WLD1991@0.389" "WLD1991@0.441" "WLD1991@1.05" "WLD1991@1.674"
 ## [233] "WLD1991@2.052" "WLD1992@0.401" "WLD1992@0.454" "WLD1992@1.724"
 ## [237] "WLD1993@0.414" "WLD1993@1.779" "WLD1994@0.458" "WLD1994@1.971"
 ## [241] "WLD1995@0.424" "WLD1995@0.5" "WLD1995@0.97" "WLD1995@2.148"
 ## [245] "WLD1995@2.498" "WLD1996@0.459" "WLD1996@1.049" "WLD1996@2.62"
 ## [249] "WLD1997@0.459" "WLD1997@0.466" "WLD1997@1.05" "WLD1998@0.467"
 ## [253] "WLD1998@1.094" "WLD1998@2.72" "WLD1998@2.94" "WLD1999@0.468"
 ## [257] "WLD1999@1.096" "WLD1999@2.909" "WLD1999@3.145" "WLD2000@0.497"
 ## [261] "WLD2000@1.164" "WLD2000@3.249" "WLD2000@3.513" "WLD2001@0.515"
 ## [265] "WLD2001@1.036" "WLD2001@1.208" "WLD2002@1.089" "WLD2003@0.573"
 ## [269] "WLD2003@1.156" "WLD2003@1.594" "WLD2004@0.621" "WLD2004@1.726"
 ## [273] "WLD2005@0.692" "WLD2005@1.923" "WLD2006@0.721" "WLD2006@2.006"
 ## [277] "WLD2007@0.766" "WLD2007@2.131" "WLD2011@3.049" "WLD2012@3.009"
 ## [281] "WLD2013@3.692" "WLD2014@4.058" "WLD2014@7.32" "WLD2015@4.399"
 ## [285] "WLD2015@4.791" "WLD2015@7.71"