`readme_WBMapGPS_replication.pdf`

Replication Data and Code for "Greening Prosperity Stripes across the Globe" (DP 2023/2024 with 3 versions, as below)

and for the Related Short *Economics Observatory* Article (September 2024)

Alexander Mihailov (University of Reading)

8 September 2024

This readme file provides guidance for the replication in Stata SE 18.5 of the world maps with 16 colors for the countries, defined using the same RGB triplets as in my MATLAB code underlying my DP "Greening Prosperity Stripes across the Globe"

https://research.reading.ac.uk/economics/wp-content/uploads/sites/87/2023/11/emdp202317.pdf

(1st version of November 2023; 1st revision of April 2024; 2nd revision of September 2024)

This world map version of the earlier cross-section color stripes in my DP figures was prepared during the 2nd revision of the DP. The world map version is, thus, only a small and added part to all other earlier figures that were generated in MATLAB and whose replication zip is on my GitHub webpage since November 2023:

https://github.com/AlexanderMihailov/GreeningProsperityStripes_DP23_replication

I have based my subsequent Stata SE 18.5 code (explained here) for my added world maps of GDP per capita, CO2 emissions per capita, and their resulting ratio. The later was proposed as a 'first pass' or a 'first approximation' indicator for what I termed 'greening prosperity' – ratio (GPR) or stripes (GPS) – per capita: see the DP as well as the subsequent short article for *Economics Observatory*:

[to add URL once it is online]

on a helpful post by Asjad Naqvi at the World Bank (which is acknowledged inside my richly annotated code):

https://blogs.worldbank.org/en/impactevaluations/making-visually-appealing-maps-stata-guest-post-asjad-naqvi

as well as a couple of other helpful Stata references on colormap palettes by Ben Jann (University of Bern) and a few other notes I was able to find online. My zip replication archive described here contains as `pdfs` these mentioned useful references by Naqvi (of 7 February 2022) and Jann (of 2018 in *The Stata Journal* and of 2 June 2022) as well as – and

most importantly – the `zip` archive by Naqvi (again, on the webpage of his post linked above):

`wb_countries_admin0_10m.zip`

with the important shape and data files from the World Bank inside, which I have then developed and adapted for my purposes. I have first unzipped this initial archive in a folder with the same name inside my own replication `zip` archive described here. One more stata.com note:

`g-4colorstyle.pdf`

is contained in it as well outside of the folder:

`WB_countries_Admin0_10m`

as well as 4 Excel `xlsx` files that contain the World Bank data I used in the DP and MATLAB codes: these 4 files have the same content but their respective filenames show the sorting of the data by a respective indicator/column: GDP pc, CO2 pc or GPR pc, whereas the 4th file has the data sorted alphabetically by 3-letter ISO codes for the countries and country groups the World Bank uses (see Table 1 in the DP or in the online note for *Economics Observatory*). These are all auxiliary files, and none of them resides in the main (and only) replication folder with the name as above, keeping the original Naqvi/WB `zip` archive name.

Inside the main (and only) `WB_countries_Admin0_10m` folder,

ALL files with names that begin with `WB_` come from the original Naqvi/WB `zip` archive and contain data and shapes. The Excel version of this dataset obtained from the original `dbf` file with the same name in Naqvi/WB's `zip` is in file `WB_countries_Admin0_10m.xlsx`, whereas my WB dataset used in the MATLAB code and the DP outside this same only folder was merged manually by matching ISO 3-letter codes (can be seen inside, using colors) to align information by country in both data sources to a single one used next, namely:

`WB_countries_Admin0_10m_ISO3Matching_240908am.xlsx`

This is the main dataset I used further, first importing the `xlsx` into Stata and saving it under the same name as a `dta` file.

Another auxiliary `xlsx` file inside this folder is:

`RGBRecode.xlsx`

which I created and used to recode the decimal expression of the RGB colors in my MATLAB code regarding the same blue-to-red colormap stripe definition as in the climate warming stripes by Ed Hawkins (University of Reading) into a corresponding fraction definition (where 255 is the denominator), as in the brown-to-green and reversed my green-to-brown RGB color definitions I created in MATLAB. There are 3 yellow columns in this `xlsx` file, each corresponding to the R(ed) G(reen and) B(lack) decimal expression as in my MATLAB code, and 3 corresponding white columns where the corresponding numerator of the fraction expression of the RGB definition colors is obtained. These fraction definitions

are then embedded into respective vectors used in 2 of the 3 Stata `do` files, i.e., those which use my own definition of the 16 colors in the spectrum, each accounting for 1/16 or 6.25% as 1 stripe in the whole spread/range between respective max and min values used (see the DP or my *Economics Observatory* note).

The other auxiliary `xlsx` file inside the folder is:

`RangesForWorldMapAsInStripes_AM240907v08.xlsx`

which I created and used to obtain the respective vectors with the border values splitting the color palette into 16 colors, the same ones as those I had used in my DP and underlying MATLAB code. Its content should be clear and self-explanatory, and there are 3 columns on the left and 3 on the right of a central blank light-red column: on the left, I have the color definitions for the cross-section of 2020 data only for the 3 indicators I use (denoted in the top row); and on the right, I have the color definitions for the world min-max range in the whole panel data set, 1990-2020 (not just for 2020). These 2 definitions are consistent with my DP and *Economics Observatory* note, where I exploit 2 versions of the 16-color stripe images: a local or country-specific one over time and a global or panel-determined analogue – these affect the choice of colors, as shown in the DP and the note.

There are 3 main `do` files, or Stata programs (and I ran them in Stata SE 18.5, as was already indicated). These can be ran in an arbitrary order.

The baseline `do` file is:

`WorldMapMercator2020GDPpcCO2pcGPSpc_240908am.do`

It produces the 3 world map color images that I first used (indeed, published) in my *Economics Observatory* online article. The color definitions by stripe within the range correspond exactly – by construction – to these used in my DP and MATLAB code.

There are 2 variations of this baseline code, as follows:

`WorldMapMercator2020GDPpcCO2pcGPSpc_240908amGlobMinMaxSince1990.do`

It produces the same 3 world maps using the same color definitions but now taking into account the global/panel min and max to form the range (itself divided into 16 stripes) over 1990-2020 for all countries rather than the local/cross-section min and max for 2020 only.

`WorldMapMercator2020GDPpcCO2pcGPSpc_240908amStataCrit.do`

It produces the same 3 world maps but this time using different color definitions, seemingly via an in-built default Stata criterion (that may be allocating a similar number of countries into each ‘bin’ defined by the 16 colors, which I do force in this code too, as well as imposing my own color definitions, for consistency with the many other figures in the DP and the note) within the whole range,, where the bins/stripes/intervals have an increasing length each marked on the respective legend/scale (mine were defined as if being equal bins in a density function – marked, again, on the respective legend/scale). As can be seen by running and comparing, this program results in more diverse colors captured on the world map.

Each of these 3 `do` files has an unusually (for posted online codes) rich annotations, notes and comments – also intended for myself (since we tend to forget what we have coded, why and how) as well as to facilitate the understanding and work of users/replicators.

The output of the 3 do files are world map graphs for the 3 main indicators studied (GDP pc, CO2 pc and GPRpc or also, equivalently, GPSpc), whose filenames are defined and clear from inside the code. Each graph is saved in 3 standard formats: `gph`, `png`, and `eps`.

Each of the 3 `do` codes runs for about a minute or two on standard laptops – I have used MacBook Pro of 2017 with macOS Monterey 12.6.2, with 2.9 GHz Quad-Core Intel Core i7 processor and memory of 16 GB.

Please let me know in case of found bugs, typos, errors or suggestions for improvements, writing to a.mihailov@reading.ac.uk, copied to mihailov.alexander@gmail.com or a_mihailov@hotmail.com, or messaging me via GitHub.