PEA Stata Code Manual

I.   Introduction

Poverty and Equity Assessments (PEA) are one of the five core ASAs under the Knowledge Compact. PEAs are an evidence-based diagnostic on country-specific changes in poverty and inequality, and drivers behind recent changes. Under the PEA 3.0 program, assessments are revamped to provide in-depth analyses of countries' progress in advancing on the updated World Bank mission statement: *ending extreme poverty and boosting shared prosperity on a livable planet.*

Key to the PEA revamp is a harmonization effort. The PEA 3.0 framework provides a systematic analysis of poverty, shared prosperity, and livable planet. A new and harmonized structure will help to streamline the analysis and facilitate comparisons across countries.[1] A set of minimum core diagnostics guides the narrative on inclusiveness of growth around the key indicators on poverty, shared prosperity and livable planet. A standardized data annex eases cross-country comparison.

This manual describes newly developed Stata codes that produce a comprehensive set of outputs to analyze trends in poverty, shared prosperity and inequality, as well as their drivers and profiles. New Stata commands generate the tables and figures for the standardized data annex (`pea core`), a full set of tables (`pea tables`) and figures (`pea figures`) that can be used for the production of core diagnostics in PEAs. Annex 1: List of tables and figures lists all tables and figures that can be produced.

II.   Stata package

Producing the harmonized outputs consists of the following steps:

- Step 1: Downloading and preparing the pea Stata package
- Step 2: Accessing the survey data from the GMD, or harmonizing it according to the GMD guidelines
- Step 3: Creating additional variables
- Step 4: Running the pea commands

Step 1: Accessing and preparing the pea package

The PEA package can be downloaded from Github (click on "Code" on the top right and "Download ZIP") and needs to be stored (unzipped) on your device.[2] While the exact storage location is not relevant, it will need to be specified in the code later on, as specified in Annex 2: Example code (adopath).

The package consists of several ado-files, including the three main commands (`pea core, pea tables, pea figures`), which produce one Excel output file each, as well as a number of other ado-files that allow for each table and figure to be created separately. Some of the ado-files run calculations,

---

[1] See the PEA website for resources and tools for PEAs.
[2] An integration to be downloaded directly from Stata for automatic installation is planned.

prepare and download data in the background. The main commands will only produce the full outputs if the whole package is downloaded. Access to the internet is needed to run the pea commands, as additional data files are downloaded using Stata's pip command.

Aside from the country's survey data, there are some additional datasets that are required to produce the full set of pea tables and figures. These are listed in Table 1. For user convenience, these datafiles are included in the pea package on GitHub (from step 1). In future versions of the package, all additional data files will be automatically downloaded in order to ensure that all data is up-to-date.

Store the data in the correct location: The data files are stored in the folder "/pea/Stata/personal/**pea**". Importantly, this **pea** folder (including the Scorecard_Summary_Vision folder) is to be moved into the personal folder of your Stata system directory (which can be found by entering "sysdir" in Stata): e.g. "c:/ado/personal/**pea**". All other required files are downloaded automatically from PIP or datalibweb when the commands are run.

Table 1: Additional data

File name	Data description	Source
CLASS.dta	Country classifications as of FY25. Note that country income class is updated in July of every year.	https://github.com/GPID-WB/Class/tree/master/OutputData
exposure_vulnerability_2021.dta	This file contains data needed to calculate exposure and vulnerability to climate-related hazards	The file can be accessed from the reproducibility package of the Poverty, Prosperity and Planet Report 2024.
UNESCO.dta	This file is used in the calculation for the multidimensional poverty measure (MPM)	Datalibweb
Scorecard_Summary_Vision	Folder contains individual excel files for each scorecard indicator	World Bank Scorecard website.

The package also installs several additional Stata commands. The following Stata packages are installed automatically when running the pea code for the first time (and don't need to be installed separately): `apoverty, ineqdeco, svylorenz, fastgini, glcurve, alorenz, povdeco, fs, groupfunction, drdecomp, adecomp, pip, schemepack, colorpalette, geoplot, palettes, colrspace, moremata`. The command `sedecomposition` cannot be installed through Stata's SSC and is therefore included in the pea package (pea/Stata/plus/s).[3].

---

[3] The sedecomposition package has been accessed from here:
https://github.com/vavalomi/stata_tools/blob/master/sedecomposition/sedecomposition.ado

Step 2a: Accessing the survey data in GMD

The pea codes run on survey data that are prepared following the Global Monitoring Database (GMD) harmonization protocols.[4] If the survey data has already been harmonized and included in the GMD, it can be accessed in Stata using the `datalibweb` command. If this is your first time accessing datalibweb, visit https://datalibweb2.worldbank.org/home to create a token to access `datalibweb` through Stata.[5] Annex 2: Example code includes an example of how survey data can be accessed using `datalibweb`. If the survey data has not yet been harmonized for the GMD, please follow step 2b. If the survey data is already in the GMD, accessing it through `datalibweb` has the advantage that multiple years of data can be retrieved easily.

Aside from the variables which are part of the standard GMD datasets, only a minimum set of additional variables need to be defined by users, which is detailed in step 3. When accessing the survey data through the GMD, make sure that variables listed in Table 2 are included in the data set. For any questions on the harmonized data, please contact your regional stats team. Note that the pea commands will run also in cases where not all variables are in the dataset or are not correctly defined. However, tables using these variables as inputs will not be produced.

Step 2b: Harmonization of survey data (not in GMD)

This step is only required if the latest survey is not already harmonized in the GMD. If the survey is already harmonized, you may skip this subsection.

If the new survey has not yet been included in the GMD, country teams should reach out to the regional stats team for harmonization of the data (strongly recommended). Table 2 lists all core variables that are needed to run the pea commands in Stata.

Furthermore, make sure that all variables take on the correct variable type and only the allowed values (see Table 2). Importantly, all variables should be numeric with variable labels and label values, to ensure proper presentation in tables. Use -`encode`- to convert string to numeric variables (see Annex 2: Example code for an example on data preparation). Users should also make sure that the subnational region variable is in the format which they would like region names to be displayed in the table output (for example, removing numbering of regions). It is important that regional stats and country teams verify that all key variables, including variable labels and label values, have been correctly specified, in particular with respect to the variable format, values and labels (as described in Table 2).

If the (harmonized) survey data is not yet in the GMD, the newest survey data needs to be appended to other available years (which can be accessed through `datalibweb`) before running the pea command, to include trends and changes in the tables and figures. When appending data from the GMD and other survey data, make sure that variables are in the same format (e.g. numeric or string). Make sure that all variables are consistently specified over years. Household identifiers might be repeated between surveys over time, even though it is not panel data.

---

[4] For more information, visit the GMD website.
[5] See here for datalibweb guidelines, and the D4G website for more instructions.

Step 3: Creating additional variables

If the new survey has been harmonized under the GMD dictionary, only the poverty lines and PPP-adjusted welfare aggregates need to be created manually. Section Annex 2: Example code (under "Preparation of additional variables") provides an example code of the construction of the additional variables (welfareppp, pline215, pline365, pline685, natline, and natline2—in case that there is a second national poverty line).[6] Variables in this part, such as national poverty lines, need to be adjusted according to the PEA country. Table 2 furthermore lists code for other variables that need to be created if an option (`setting(GMD)`) is not specified, which will be explained in more detail in step 5.

Table 2: List of variables used by the PEA command

Description	Name	Values	Code
Household identifier	hhid	*integer*	
Individual identifier	pid	*integer*	
Year of survey	year	*integer (4 digits)*	
Poverty specific weights	weight_p	*integer*	
Welfare aggregate (LCU)	welfare	*integer*	
Welfare aggregate (divided by CPI (2017) and ICP (2017), and 365 days)	welfppp	*integer*	gen welfppp = welfare/cpi2017/icp2017/365 (note: welfare should be in per-capita terms)
Relation to household head	relationharm	1=Household head; 2=Spouse; 3=Child; 4=Parents; 5=Other relative; 6=Non-relative	
Household head dummy	head	0=Not household head; 1=Household head	gen head = relationharm==1 if relationharm~=. (note: only needed when setting(GMD) is not specified, see next sub-section)
Marital status of individual	marital	1=married; 2=never married; 3=living together; 4=divorced/separated; 5=widowed	
Married dummy	married	0=not married; 1=married	gen married = marital==1 if marital~=. (note: only needed when setting(GMD) is not specified, see next sub-section)
Age of individual	age	*integer*	
Sex of individual	male	0=female; 1=male	

---

[6] Importantly, verify that the correct deflators are used.

Household size	hsize	*integer*	
Highest level of education completed (4 categories)	educat4	1=No education; 2=Primary (complete or incomplete); 3=Secondary (complete or incomplete); 4=Tertiary (complete or incomplete)	
Urban indicator	urban	1=urban; 2=rural	
Currently in school dummy	school	0=not currently in school; 1=currently in school	
Labor force Status	lstatus	1=employed ; 2=unemployed; 3=not in labor force	
Not working dummy	nowork	0=working; 1=not working	gen nowork = lstatus==2| lstatus==3 if lstatus~=. (note: only needed when setting(GMD) is not specified, see next sub-section)
Sector/industry of employment (4 categories) [NOTE: For some countries this is industrycat4_year]	industrycat4	1=Agriculture; 2=Industry; 3=Services; 4=Other	
Employment status	empstat	1=Paid employee; 2=Non-paid employee; 3=Employer; 4=Self-employed	
Improved water access	imp_wat_rec	0=No; 1=Yes	
Improved sanitation facility	imp_san_rec	0=No; 1=Yes	
Access to electricity	electricity	0=No; 1=Yes	
Ownership of a television	tv	0=No; 1=Yes	
Ownership of a car	car	0=No; 1=Yes	
Ownership of a cellphone	cellphone	0=No; 1=Yes	
Ownership of a computer	computer	0=No; 1=Yes	
Ownership of a fridge	fridge	0=No; 1=Yes	
National poverty line (2017 LCU)	natline	*integer*	gen natline = X if year==
Second national poverty line (2017 LCU)	natline2	*integer*	gen natline2 = X if year==

Poverty line: $2.15 per day (2017 PPP)	pline215	2.15	gen pline215 = 2.15
Poverty line: $3.65 per day (2017 PPP)	pline365	3.65	gen pline365 = 3.65
Poverty line: $6.85 per day (2017 PPP)	pline685	6.85	gen pline685 = 6.85
Subnational ID. Make sure the variable is numeric and harmonized over all years.	subnatvar	integer (labeled)	e.g. encode subnatid, gen(subnatvar)
Comparability between survey rounds	Comparability	*Integer*	

Step 4: Running the pea codes

Once the data and pea package have been prepared, users can follow the example in Annex 2: Example code to run the commands. Note that the prepared data and the adopath to the pea package need to be called every time before a pea code is run (as in the example).

The syntax of the three main programs inside the `pea` code is outlined below. Each table can also be produced separately, using the table specific command, such as (see Annex 2: Example code). All tables and figures include standardized tables notes. These can be adjusted by the country teams for specifics on the underlying survey used for the analysis in the Excel file or text program. In Stata, help files are available for each code's syntax and option.

Table 3 provides details to the main options that need to be specified to produce output. Option `setting(GMD)` plays a special role. When specified (see syntax below and example code in the annex), a set of variables is created and cleaned in a harmonized style automatically. Specifying the options is recommended, to produce harmonized output. If `setting(GMD)` is not specified, additional inputs are needed, which are listed below and in Table 4. While not adding `setting(GMD)` is not recommended, it can be useful in certain circumstances. It adds flexibility, for instance when users want to use different categories and values than the default values listed in Table 2. An example is when the education variable should take more than 4 values, a different variable can be inserted in option `edu()`.[7] The example code in the annex provides an example for a code which does not use `setting(GMD)`.

Options that are specific to the `pea figures` command are listed in Table 5. Table 2 shows the code for some variables that need to be created separately (e.g. married).

1. `pea core`

This first code (in Stata: `pea core`) generates a standardized data annex that should be included in each PEA, as set out in the new PEA guidelines. The produced Excel file contains 4 tables and 2 figures. This annex is composed of main poverty and shared prosperity indicators, as well as multidimensional

---

[7] In the GMD, education by default follows the International Standard Classification of Education (ISCED) mappings.

and sub-group (e.g. by age or education) poverty rates. The code also produces core statistics for benchmark countries, the PEA country's region and income group, and profiles of the poor and non-poor. Growth incidence curves and the Datt-Ravallion decomposition complement the core outputs.[8]

```
pea core [weight] [if exp] [in exp] [, NATWelfare(varname numeric)
NATPovlines(varlist numeric) PPPWelfare(varname numeric)
PPPPovlines(varlist numeric) Year(varname numeric) SETting(string)
excel(string) save(string) BYInd(varlist numeric) age(varname numeric)
male(varname numeric) ONELine(varname numeric) ONEWelfare(varname
numeric) MISSING Country(string) LATEST WITHIN3 BENCHmark(string)
spells(string) urban(varname numeric)
```

2. `pea tables`

The second set of outputs is produced using the `pea tables` command in Stata. The generated outputs expand on the core outputs, providing in-depth statistics on the state of poverty and shared prosperity over the last years, the profiles of those in poverty and vulnerable to falling into poverty, as well as drivers of changes. Drivers include economic growth, and how inclusive to those on the bottom of the income distribution it has been, as well as decompositions of changes into within and between group shifts in population shifts.

```
pea tables [weight] [if exp] [in exp] [, natwelfare(varname numeric)
natpovlines(varlist numeric) pppwelfare(varname numeric)
ppppovlines(varlist numeric) year(varname numeric) setting(string)
excel(string) save(string) byind(varlist numeric) urban(varname
numeric) married(varname numeric) missing oneline(varname numeric)
onewelfare(varname numeric) country(string) latest within3
benchmark(string) spells(string)]
```

3. `pea figures`

The third output includes a series of figures that are produced using the `pea figures` command in Stata for the convenience of the user. Figures span a broad range of topics, including a comparison of poverty rates and GDP per capita with other countries, changes in poverty and inequality over time, population composition by income decile, profiles, or exposure and risk from climate-related hazards.[9]

All outputs are automatically created by running the three commands, with minimum manual inputs needed from users. Codes can nevertheless be tailored to county contexts, for instance by specifying country-specific education or industry groups. The codes enable harmonization of core analysis across countries, and easier reproducibility for other users. The codes can be run for any number of surveys for a country. That is, the output produced can include a single year or multiple years to allow for easy comparisons of indicators over time.

```
pea figures [if] [in] [weight], [natwelfare(varname)
natpovlines(varlist) pppwelfare(varname) ppppovlines(varlist)
year(varname) setting(string) excel(string)save(string) byind(varlist)
```

---

[8] See benchmarking tool and guidance note to identify benchmarking countries.
[9] See the Corporate scorecard website and CCG dashboard on the risk from climate-related hazards indicator.

```
urban(varname) oneline(varname) onewelfare(varname) missing
country(string) within(integer) combine nonotes comparability(varname)
benchmark(string) spells(string) equalspacing scheme(string)
palette(string)welfaretype(string)]
```

Additional options, if **setting(GMD)** is not specified:

```
hhhead(varname numeric) edu(varname numeric) married(varname numeric)
school(varname numeric) services(varlist numeric) assets(varlist
numeric) hhsize(varname numeric)  hhid(string) pid(string)
industrycat4(varname numeric) lstatus(varname numeric) empstat(varname
numeric)
```

Table 3: Main options

Setting	Description	*Entry* or `variable name`
Country()	Required. PEA country	*3-letter country code*
natwelfare()	Required. Welfare aggregate in current or constant LCU	`welfare`
natpovlines()	Required. National poverty line, multiple entries allowed	`natline`
pppwelfare()	Required. Welfare aggregate in PPP terms	`welfareppp`

ppppovlines()	Required. International poverty lines in PPP terms	`pline215`
		`pline365`
		`pline685`

year()	Required. Year variable in survey dataset	`year`
onewelfare()	Required. Main welfare variable used for comparisons (i.e. Table 1, Table 7, Table 14, Table A1, Table A2, Table A4, Table A4, Figure A1)	`welfareppp`
oneline()	Required. Main poverty line used for comparisons (i.e. Table 1, Table 7, Table 14, Table A1, Table A2, Table A4, Table A4, Figure A1)	`pline685`
byind()	Optional (but needed for many tables). Geographic units for disaggregation, multiple entries allowed	`urban subnatvar`
urban()	Optional. Variable that indicates the urban indicator	`urban`
benchmark()	Optional. Countries to be benchmarked against, multiple entries allowed	*3-letter country codes*
setting()	Optional. If GMD option is specified, harmonized variables are created, and additional options (Table 4) in do not need to be specified.	*GMD*
spells()	Optional. Years to be used when calculating or showing changes across periods (such as Growth Incidence Curves), multiple entries allowed separated by semicolon	*year1 year2; year2 year3*
missing	Optional. Missing values are reported separately. Otherwise, they are treated as non-existent.	
latest	Optional. Includes only the most recent available data.	

Within3	Optional. Limits analysis to data from benchmark countries within 3 years of the target year.	
excel()	Optional. Path of directory where and name under which output file is saved. If it is not specified, the file is saved in "Temp". Note that if the option is specified, the excel file needs to exist before the code is run (empty).	*Path and file name*
save	Optional. Specifies the file path for saving results.	

Table 4: Additional options if setting(GMD) is not specified

Setting	Description	Default under setting(GMD)
age()	Age of individual	`age`
male()	Sex of individual	`male`
married()	Marital status (variable is based on marital)	`married`
services()	Services to be included in the analysis. Default are access to improved water, access to improved sanitation, access to electricity.	`imp_wat_rec imp_san_rec electricity`
assets()	Assets to be included in the analysis	`tv car cellphone computer fridge`
hhsize()	Household size	`hsize`
head()	Household head indicator (Variable is based on variable relationharm)	`head`
hhid()	Household identifier	`hhid`
pid()	Individual identifier	`pid`
edu()	Education level	`educat4`
school()	Currently in school dummy	`school`
lstatus()	Dummy if respondent does not work (Variable is based on variable lstat)	`nowork`
industrycat4()	Industry categories	`industrycat4`
empstat	Employment status of respondent	`empstat`

Table 5: Additional figure options

Setting	Description	Entry or variable name
within()	Optional. Specifies the number of years before and after the pea survey year, to define which surveys from other countries should be used (e.g. in scatter plots on inequality). Default is 3, and value should be less than 10.	*integer*
combine	Optional. When specified, figures with multiple panels are combined to one figure with only one legend.	
nonotes	Optional. Suppresses that standard notes are added beneath figures.	

9

comparability()	Required. This variable denotes which survey rounds are comparable over time. Non-comparable survey rounds are not connected in figures.	`comparability`
equalspacing	Optional. When specified, figures display constant gaps between years, regardless of how far away years are. This can be useful if gaps between survey-years are large.	
scheme()	Optional. Sets the scheme, specifying the overall look of the figures. Default is "white_tableau".	*string*
palette()	Optional. Sets the color palette for figures. Default is "tab20". See Annex 2: Example code for an example of a custom set of colors.	Either *string* (e.g. *vividis)* or list of colors.
welfaretype()	Optional. Can be used to specify whether the survey uses income or consumption to calculate welfare. Figures showing scatters of inequality display different symbols for countries with consumption or income aggregates.	*CONS* or *INC*

III.   Frequent caveats and questions

- Poverty numbers are based on population numbers from the respective surveys. Note that this may cause slight discrepancies between poverty numbers reported in PIP and in PEBs, which are based on population from WDI.
- Note that labor force related variables can have different recall periods. For example, in GMD, some survey data use the variable *lstatus* (which has a 7-day recall period) and some use *lstatus_year* (which has a 1-year recall period). It is up to country teams to decide whether variables are comparable over time (e.g. even when recall periods are different). For the example of labor force status, if different recall periods are deemed to be comparable, one variable would need to be defined which has both variables' values (e.g. `replace lstatus = lstatus_year if lstatus==.`*)*
- When there are multiple survey periods and welfare in LCU is in current prices, make sure that either national poverty lines are also expressed in current prices or the welfare aggregate is converted to constant terms.
- Note that the code produces output, regardless of the number of observations for each cell. It may be that a subgroup indicator is based on very few observations. Country teams need to decide if these are to be shown or not.
- Note that the MPM is retrieved from the GMI. New surveys in the GMD might not be immediately included in the GMI (updated usually in April and October).
- For growth incidence curves disaggregated by urban and rural areas, the welfare distribution represents the whole distribution within that area.
- The "Other" sector in the variable industrycat4 is based on a country-specific harmonization from the GMD.
- It is important that the correct temporal and spatial deflators are used.

Annex 1: List of tables and figures

Standardized data annex (pea core):
- Table A1: Core poverty and equity indicators
- Table A2: Poverty indicators by subgroup
- Table A3: Benchmarking of poverty and inequality
- Table A4: Poverty profiles
- Figure A1: Growth Incidence Curves
- Figure A2: Datt-Ravallion decomposition

Full set of tables (pea tables):
- Table 1: core poverty indicators.
- Table 2: poverty rate and share of poor by area and region.
- Table 3: subgroup poverty rates.
- Table 6: Multidimensional Poverty Measure (World Bank)
- Table 7: vulnerability for poverty.
- Table 8: core inequality indicators.
- Table 9: vision indicators (corporate scorecard).
- Table 10: benchmarking of poverty and inequality.
- Table 11_1 / Figure 11_1: Growth Incidence Curve (All sample).
- Table 11_2 / Figure 11_2: Growth Incidence Curve (Rural).
- Table 11_3 / Figure 11_3: Growth Incidence Curve (Urban).
- Table 12b: decomposition of poverty changes: growth and redistribution - Datt-Ravallion decomposition
- Table 12c: decomposition of poverty changes: growth and redistribution - Shorrocks-Kolenikov decomposition
- Table 13: decomposition of poverty changes: Huppi-Ravallion decomposition.
- Table 14: profiles of the poor.
- Table 15: gender typology (to come).

Full set of figures (pea figures):

- Figure 1: Poverty rates by year lines.
- Figure 2: Poverty and GDP per capita scatter.
- Figure 3: Growth Incidence Curve.
- Figure 4: Decomposition of poverty changes: growth and redistribution: Datt-Ravallion and Shorrocks-Kolenikov.
- Figure 5: Decomposition of poverty changes: growth and redistribution: Huppi-Ravallion .
- Figure 6: GDP per capita GDP - Poverty elasticity.

- Figure 7: Welfare Figure with poverty line breakdowns.
- Figure 9a: Inequality by year lines.
- Figure 9b: GINI and GDP per capita scatter.
- Figure 10a: Prosperity gap by year lines.
- Figure 10b: Prosperity gap scatter (line-up).
- Figure 10c: PG (survey) and GDP per capita scatter.
- Figure 12: Decomposition of growth in prosperity gap.
- Figure 13: Distribution of welfare by deciles
- Figure 14: Multidimensional poverty: Multidimensional Poverty Measure (World Bank).
- Figure 15: Climate risk and vulnerability.

Annex 2: Example code

The following code produces the main tables, the standardized data annex and figure outputs for Guinea-Bissau.

```
datalibweb, country(GNB) year(2018 2021) type(gmd) mod(all) clear

// Preparation of additional variables
//Welfare aggregate to compute international poverty rates
gen welfppp = welfare/cpi2017/icp2017/365
gen pline215 = 2.15
gen pline365 = 3.65
gen pline685 = 6.85
//National poverty line: Adjust according to the PEA country
gen natline = 298083.5 if year == 2021
replace natline = 271071.8 if year == 2018
la var pline215 "Poverty line: $2.15 per day (2017 PPP)"
la var pline365 "Poverty line: $3.65 per day (2017 PPP)"
la var pline685 "Poverty line: $6.85 per day (2017 PPP)"
la var natline "Poverty line: 298,083.5 per year (2017 LCU)"
// Cleaning of region variable: Adjust according to the PEA country
split subnatid, parse("-") gen(tmp)
replace tmp2 = ustrlower( ustrregexra( ustrnormalize( tmp2, "nfd" ) ,
"\p{Mark}", "" ) )
replace tmp2 = " bolama/bijagos" if tmp2 == " bolama_bijagos"
replace tmp2 = proper(tmp2)
encode tmp2, gen(subnatvar)

save "${pea_path}/data/GNB_GMD_clean.dta", replace

// 1. Main tables code:
//Set ado path
adopath + "${pea_path}/ado/pea/Stata/plus"
//Call data
use "${pea_path}/data/GNB_GMD_clean", clear
//Run code with options
pea tables [aw=weight_p], c(GNB)  year(year)                   ///
        natw(welfare) natp(natline)                           ///
        pppw(welfppp) pppp(pline365 pline215 pline685)        ///
        onew(welfppp) oneline(pline365)                       ///
        byind(urban subnatvar)                                ///
        benchmark(CIV GHA GMB SEN)                            ///
        setting(GMD) missing                                  ///
        spells(2018 2021)

// 2a. Appendix tables code:
//Set ado path
```

13

```
adopath + "${pea_path}/ado/pea/Stata/plus"
//Call data
use "${pea_path}/data/GNB_GMD_clean", clear
//Run code with options

pea core [aw=weight_p], c(GNB) year(year)                       ///
        natw(welfare) natp(natline)                             ///
        pppw(welfppp) pppp(pline365 pline215 pline685)          ///
        onew(welfppp) oneline(pline365)                         ///
        byind(urban subnatvar)                                  ///
        benchmark(CIV GHA GMB SEN)                              ///
        setting(GMD) missing                                    ///
        spells(2018 2021)


// 2b. Appendix tables code without setting(GMD):
//Set ado path
adopath + "${pea_path}/ado/pea/Stata/plus"
//Call data
use "${pea_path}/data/GNB_GMD_clean", clear
//Run code with options
pea core [aw=weight_p], c(GNB) year(year)                       ///
        natw(welfare) natp(natline)                             ///
        pppw(welfppp) pppp(pline365 pline215 pline685)          ///
        onew(welfppp) oneline(pline365)                         ///
        byind(urban subnatvar)                                  ///
        benchmark(CIV GHA GMB SEN)                              ///
        missing spells(2018 2021)                               ///
        age(age) male(male) hhhead(head) edu(educat4)           ///
        urban(urban) married(married) school(school)            ///
        services(imp_wat_rec imp_san_rec electricity)           ///
        assets(tv car cellphone computer fridge)                ///
        hhsize(hsize) hhid(hhid) pid(pid)                       ///
        industrycat4(industrycat4)                              ///
        lstatus(nowork) empstat(empstat)


// 3. Figures code:
//Set ado path
adopath + "${pea_path}/ado/pea/Stata/plus"
//Call data
use "${pea_path}/data/GNB_GMD_clean", clear
//Run code with options
pea figures [aw=weight_p], c(GNB) year(year)                    ///
        natw(welfare) natp(natline)                             ///
        pppw(welfppp) pppp(pline365 pline215 pline685)          ///
        onew(welfppp) oneline(pline215)                         ///
        byind(urban) benchmark(CIV GHA GMB SEN AGA)             ///
        spells(2010 2018; 2018 2021)                            ///
```

```
            setting(GMD) urban(urban)   within(3)                    ///
            comparability(comparability) welfaretype(CONS)           ///
            combine nonotes
```

// 4. Running a single figure and defining own color palette
```
use "${pea_path}/data/GNB_GMD_clean", clear
adopath + "${pea_path}/ado/pea/Stata/plus"
local custom_palette = "#337ab7 #5cb85c #5bc0de #f0ad4e #d9534f #e6e6e6
#286090 #449d44 #31b0d5 #ec971f #c9302c"

pea figure2 [aw=weight_p], c(GNB) year(year)                        ///
            onew(welfppp) onel(pline215)                            ///
            benchmark(CIV GHA GMB SEN) palette(`custom_palette')
```

// 5. Running a single table with LCU poverty line
```
use "${pea_path}/data/GNB_GMD_clean", clear
adopath + "${pea_path}/ado/pea/Stata/plus"

// When running single tables, setting(GMD) is not called, and additional
variables need to be created
gen head = relationharm==1 if relationharm~=.
la def head 1 "HH head"
la val head head
gen nowork = lstatus==2|lstatus==3 if lstatus~=.
gen married = marital==1 if marital~=.

pea table14 [aw=weight_p], welfare(welfare) povlines(natline)       ///
            year(year) missing urban(urban)                        ///
            age(age) male(male) hhhead(head)                       ///
            edu(educat4) married(married) school(school)           ///
            services(imp_wat_rec imp_san_rec electricity)          ///
            assets(tv car cellphone computer fridge)               ///
            hhsize(hsize) hhid(hhid) pid(pid)                      ///
            industrycat4(industrycat4) lstatus(nowork)             ///
            empstat(empstat)
```