

Estimating Global Poverty in Stata

The povcalnet command

*R. Andrés Castañeda Aguilar, Christoph Lakner, Espen B. Prydz,
Jorge Soler Lopez, Ruoxuan Wu and Qinghua Zhao*

September 2019

Keywords: Global Poverty, Inequality, World Bank, Sustainable Development Goals.



Development Data Group
Development Research Group
Poverty and Equity Global Practice Group

Abstract

This note describes how poverty measures reported by the World Bank can be replicated using the Stata command **povcalnet**. Users can estimate poverty at any poverty line for the world, regions or sets of countries, by directly querying the World Bank's database of household surveys. The command also retrieves inequality statistics provided by the database.

All authors are currently affiliated with the World Bank, or have previously been. Corresponding author: acastaneda@worldbank.org. The authors would like to thank Joao-Pedro Azevedo, Francisco Ferreira, Tony Fujs, Dean Jolliffe, Aart Kraay, Kihoon Lee, Daniel Mahler, Minh Cong Nguyen, Marco Ranaldi and Prem Sangraula, as well as seminar participants at the World Bank, for comments received on earlier versions of this code and this note. This note has been cleared by Francisco Ferreira.

The Global Poverty Monitoring Technical Note Series publishes short papers that document methodological aspects of the World Bank's global poverty estimates. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent. Global Poverty Monitoring Technical Notes are available at <http://iresearch.worldbank.org/PovcalNet/>.

1 Introduction

PovcalNet reports the World Bank’s official global, regional and country-level poverty estimates.¹ It is not only a source of information on countries’ trajectories towards the Sustainable Development Goals, as well as the earlier Millennium Development Goals, but also an important public resource for researchers, international organizations, national governments, journalists and civil society to track indicators of poverty and inequality for countries at all stages of development.

PovcalNet is an interactive computational tool that allows users to estimate poverty at any poverty line, and for individual countries or groups of countries, by querying the World Bank’s database of household surveys. Default estimates are provided at several poverty lines, but users are free to choose their own poverty line.² Using PovcalNet, it can be shown for example that in 2015, half of the world’s population lived on less than \$6.17 per person per day. Finally, users can modify assumptions made by the World Bank, such as the use of the 2011 purchasing power parity (PPP) exchange rates for international price comparisons.

The main method for accessing PovcalNet has been through its web interface at <http://iresearch.worldbank.org/PovcalNet/home.aspx>. The website is based on a web API, as explained in more detail in Zhao (2018). This means that every query to the PovcalNet website (e.g. estimate the poverty headcount ratio at \$2 per day in Bangladesh in 2016) generates a URL that returns the results in a machine-readable format.³ In this paper we present the Stata command `povcalnet`, which calls the PovcalNet API (where all computations occur) and loads the results directly into Stata. It offers the same functionality as the web interface, but presents a major improvement to the usability of PovcalNet for Stata users.⁴ This allows for simple merging with other Stata datasets or the application of custom assumptions. For example, the PovcalNet data can then be easily combined with the World Development Indicators, which can be queried directly into Stata using the `wbopendata` command.⁵

1. The term country, used interchangeably with economy, does not imply political independence but refers to any territory for which authorities report separate social or economic statistics.

2. As stated on the PovcalNet website, the tool has been primarily designed for the public replication of the World Bank’s poverty estimates at the widely used international poverty lines, including \$1.9 a day and \$3.2 a day (in 2011 PPP). It should be noted that estimates of the densities near the bottom and top tails of the distribution could be less precise, so users should bear this in mind when setting their own poverty lines. For example, a sizable literature has shown that the coverage of household surveys tends to be incomplete at the top tail (see World Bank (2016), p. 80 for a summary).

3. For example, the following URL returns the poverty headcount ratio at \$2 per day in Bangladesh in 2016 (in .csv format):

```
http://iresearch.worldbank.org/PovcalNet/PovcalNetAPI.ashx?C0=BGD\_3&PL0=2&Y0=2016&format=csv
```

4. Zhao (2018) describes how to call the API from R, and download the results into Excel. An R package with the same features as the `povcalnet` Stata command is available in CRAN (`install.packages("povcalnetR")`).

5. The coding of `povcalnet` closely follows `wbopendata`, developed by Joao-Pedro Azevedo.

This short paper is structured as follows. [Section 2](#) provides a general description of the PovcalNet tool. The `povcalnet` Stata command is introduced in [Section 3](#). [Section 4](#) illustrates the use of the `povcalnet` command with several examples, and [Section 5](#) concludes.

2 PovcalNet

PovcalNet reports the World Bank’s official global, regional and internationally comparable country-level poverty estimates. These estimates are also published in the [World Development Indicators](#) and the [Poverty and Equity Data Portal](#), as well as the biennial Poverty and Shared Prosperity report (World Bank (2016) and World Bank (2018a)). PovcalNet is managed jointly by the Data and Research Groups within the World Bank’s Development Economics Division. It draws heavily upon a strong collaboration with the Poverty and Equity Global Practice, which is responsible for gathering and harmonizing the underlying survey data.

PovcalNet reports two types of estimates: [1] *survey-year* estimates, directly from surveys available, and [2] *reference-year* estimates, which align survey estimates to common reference-years for global and regional aggregations. Since many countries do not have household surveys every year, it is necessary to extrapolate or interpolate the welfare aggregate(s) of the closest survey year(s) to a common point in time (i.e., reference year) in which poverty can be estimated across all countries. These extrapolations and interpolations to the reference years use growth rates from national accounts to project forward (backward) the welfare aggregate. The reference-year estimation thus requires additional assumptions. In particular, it assumes that the growth in the survey mean can be approximated by growth in national accounts, and that all parts of the distribution grow at the same rate. Further details on the World Bank’s global poverty measurement methodology are provided in Jolliffe et al. (2015) (Chapter 6) and World Bank (2018a) (Data Appendix).

In sum, reference year estimates are appropriate for examining regional and global aggregates. For comparing trends for a particular country over time, users are advised to use survey-year estimates. Furthermore, inequality estimates are only reported for survey years, since the reference-year extrapolations assume distribution-neutrality. The defaults built into the PovcalNet web interface recognize this: aggregations over regions or sets of countries are only possible when a reference-year is selected, whereas queries at the country level return survey-year estimates by default. The `povcalnet` command mimics this functionality.

PovcalNet reports the following poverty measures at the poverty line chosen: headcount ratio, poverty gap, squared poverty gap and Watts index. For inequality measures, the Gini index, mean log deviation and decile shares are reported. The underlying welfare aggregate is

6. Welfare aggregates are income or consumption depending on data availability in each country. When both measures are available in the same year, PovcalNet uses consumption in the regional and global poverty aggregates, while reporting the survey-year estimates for both measures. World Bank (2018a) (Data Appendix) discusses the important differences between poverty measures based on consumption or income.

per capita household income or consumption expressed in 2011 PPP-adjusted USD.⁶ Poverty lines are expressed in daily amounts, while means and medians are monthly.

PovcalNet uses unit-record (micro) survey data whenever possible. For country-years without access to the complete micro data, grouped data, such as deciles or ventiles, are used in combination with a parametric Lorenz curve. This applies particularly to the earlier years and continues to apply to some countries with limited access to micro data, such as China. The PovcalNet website provides further details and documentation on the methodology and data sources underpinning PovcalNet.

3 The povcalnet command

The `povcalnet` command allows Stata users to compute poverty and inequality indicators for more than 160 countries in the World Bank's database of household surveys. It replicates the functionality of the [website interface](#), while offering a more convenient means of access for Stata users.

3.1 Syntax

The syntax for `povcalnet` is the following:

```
povcalnet [subcommand], [country(country list | all) region(region list | all)
  year(numlist | all | last) povline(numlist) aggregate
  fillgaps ppp(#)] coverage(national | urban | rural)
```

3.2 Basic use

The `povcalnet` command executed by itself retrieves the poverty headcount at \$1.9 a day for all the surveys available in the PovcalNet server. As explained in more detail in the next section, executing `povcalnet` by itself is equivalent to

```
povcalnet, countries(all) year(all) povline(1.9)
```

Options `countries()`, `years()`, and `region()` are used by the `povcalnet` command to create the appropriate API query which includes only the relevant countries, years and regions. In other words, `povcalnet` does not query the entire database every time the API is called. Thus,

```
povcalnet, clear
keep if (countrycode == "ARG" & year == "2015")
```

is equivalent to

```
povcalnet, country(ARG) year(2015) clear
```

Yet, the second option is not only less verbose, but also more efficient and faster.

3.3 Main options

`country(country list, all) | region(region list, all)`: Specifies the country (or economy) or region for the request. See Annex 1 for a list of economies and regions, and their three-letter country code used by PovcalNet. Users can specify either country or region, but not both. Annex 1 also includes the regional classification used by PovcalNet. In brief, “Other High Income” includes mostly high-income countries (see p. 152 of World Bank (2018a) for a list, as well as the PovcalNet website); all other economies are included in geographic regions, as defined by the World Bank.

For example, `povcalnet, country("ALB")` retrieves the default estimates for all the surveys available in Albania. Similarly, `povcalnet, country(COL ARG IND MEX)` retrieves information for Colombia, Argentina, India and Mexico. Note that Mexico has surveys that use income and consumption, so both types welfare are included in the output.

When specifying a region, all countries belonging to that region are returned. For example calling `povcalnet, region(LAC)` will return all countries in Latin America and the Caribbean for all the years available in each country.

Notice that providing `region(all)` and `country(all)` will return the same output, loading all countries in the PovcalNet database for the specified year(s).

`year(numlist | all | last)`: Specifies the year(s) for the selected countries or regions. Four-digit years are required. Multiple years need to be separated by spaces. The option `all` is a shorthand for calling all surveys available in each country. `all` is also the default value of option `years()`. Option `last` downloads the latest available year in each country.

Following the example [above](#), `povcalnet, country(COL ARG IND MEX) year(2010)` limits the results to 2010. Since India does not have a survey in 2010, no result is reported for India. Mexico, in contrast, has both an income and a consumption survey in 2010, which are reported as two lines of output. Similarly, `povcalnet, country("ALB") year("2002 2030")` yields results for 2002 and not for 2030. When no surveys are available for the years requested (e.g., `povcalnet, country("ALB") year(2030)`), `povcalnet` yields an error.

`povline(numlist)`: The poverty line(s) for which poverty will be calculated. When selecting multiple poverty lines, users must use less than 4 decimal precision and separate

each line with a space. If left empty, the default poverty line of \$1.9 per day is used. Poverty lines are expressed in 2011 PPP-adjusted USD per capita per day.

For example, `povcalnet, country("ALB CHN") povline(5 10)` estimates poverty headcounts for all the years available in Albania and China at \$5.0 and \$10.0 a day.

3.4 Reference-year options

So far, the options explained above refer the estimates at the survey-year level. In this section, we present the main functions to calculate global and regional aggregate poverty measures. Recently, the World Bank has been adding a reference year every two years. At the time of writing, the reference years available are 1981, 1984, 1987, 1990, 1993, 1996, 1999, 2002, 2005, 2008, 2010, 2011, 2012, 2013 and 2015.

fillgaps: Loads all country-level estimates that are used to create the aggregates in the reference years. In other words, **fillgaps** requests the underlying data to do the aggregation in option **aggregate**. Only reference years (see list [above](#)) can be combined with the **fillgaps** option. It is important to note that countries without a survey in the reference year are extrapolated or interpolated using national accounts growth rates, as described above (also see chapter 6 in Jolliffe et al. (2015)). Therefore, changes at the country-level from one reference year to the next need to be interpreted carefully and may not be the result of a new household survey.

In contrast to the example [above](#), `povcalnet, country(COL MEX ARG IND) year(2010) fillgaps` presents only one observation for Mexico because the consumption survey is used in the aggregation. Three observations are reported for India because its 2010 national estimate is aggregated from the urban and rural estimates, which are interpolated separately.

aggregate: Calculates aggregated poverty measures as the population-weighted average of the countries requested by the user. For example, `povcalnet, country(col idn vnm egyptur zaf) year(2010 2015) aggregate` estimates the *aggregated* poverty measures for the CIVETS countries (Colombia, Indonesia, Vietnam, Egypt, Turkey and South Africa) for 2010 and 2015: Since this aggregation of countries is specific to the user's request, the region variable will be labelled as "XX". Users can choose any set of countries for aggregation.

If option `country(all)` is combined with option **aggregate**, `povcalnet` executes instead `povcalnet wb`, which returns the default global aggregates used by the World Bank as explained in more detail in the next section. In contrast, to aggregate all countries in a particular reference year, users need to list them all in the `country(all)` option. One way to do so is as follows:

```
povcalnet info, clear
levelsof country_code, local(all) clean
```

```
povcalnet , country('all') year(2015) clear aggregate
```

As explained in more detail below, the World Bank's global aggregation is estimated as a population-weighted average of the regional headcount ratios. The regional population weights include countries without any household survey data, while the regional headcount ratios are estimated over those countries in the region that have survey data. In other words, the World Bank aggregation assumes that the headcount ratio in the missing countries equals the regional average. In contrast, aggregating over `country(all)` estimates a global aggregate over only those countries that have a household survey.

3.5 Sub-commands

`povcalnet` provides several tools for the convenience of the user. These features are invoked through sub-commands in the form `povcalnet subcommand`, where *subcommand* could be `info`, `cl`, `wb`, or `test`.

`povcalnet info`: Returns a clickable version of the available surveys, countries and regions. Selecting countries from the menu loads the survey-year estimates. Choosing regions loads the regional aggregates in the reference years. If option `clear` is added, the data in memory is replaced with a PovcalNet guidance database. If option `clear` is not included, `povcalnet` preserves data in memory but displays a clickable interface of surveys availability in the results window.

`povcalnet cl`: Changes combinatorial query of parameters for one-on-one correspondence of parameters. By default, `povcalnet` creates a combinatorial query of the parameters selected, so that the output contains all the possible combinations between `country()`, `povline()`, `year()`, and `coverage()`. Alternatively, the user may select the subcommand `cl` to parse a one-on-one (i.e., country by country) request. In this case, the first country listed in `country()` will be combined with the first element in options `year()`, `povline()`, and `coverage()`, and similarly with subsequent countries. If only one element is added to parameters `year()`, `povline()`, or `coverage()`, it would be replicated as many times as elements provided in option `countries()`. Option `ppp()` is only allowed in the one-to-one query.⁷

`povcalnet wb`: Downloads World Bank's regional and global aggregation. This function differs from option `aggregate` in two ways: [1] `povcalnet wb` uses a predefined set of countries in each region, whereas option `aggregate` allows users to select their own set of countries for aggregation. [2] The World Bank aggregation (queried by `povcalnet wb`) assumes that the poverty rate for an economy without a household survey is the regional average. This creates differences in two types of estimates: [a] When a region includes countries without a household survey, the number of poor is different between the two methods even when users query the same set of countries. The number of poor according

7. If only one element is added to option `ppp()`, it would be applied to all the countries in `country()`.

to the World Bank method is obtained as the product of the region's headcount index and the total regional population (which includes the population of countries without any household survey). In contrast, when using `aggregate`, the number of poor is the product of the region's headcount index and the total population of the economies included in the aggregation. [b] In computing the poverty estimates for the world, the World Bank aggregation takes the population-weighted average of the regional estimates. Each region is weighted using the total regional population, including the population of countries without any household survey. Countries without a survey are thus implicitly assigned the regional poverty rate. In contrast, when using the option `aggregate`, only countries with a survey are considered (which are then weighted by their population).

`povcalnet test`: Executes the last query in the default internet browser regardless of whether `povcalnet` fails. It is useful for testing the query independently of the performance of the `povcalnet` command. For instance, the poverty headcount of all the survey years available for Colombia is downloaded with the first command line of the code below. The second line executes the same query in the internet browser for testing purposes.

```
povcalnet, countr(COL) clear
povcalnet test
```

3.6 Other options

`ppp(#)`: (seldom used) allows the specification of an alternative PPP exchange rate. The default is the 2011 PPP exchange rate for private consumption, as explained in Atamanov et al. (2018). Altering the PPP with this option only works with one country selected at a time.

`clear`: Replaces data in memory.

3.7 Returned values

`povcalnet` stores the following results in `r()`. Suffix `_#` refers to the number of poverty lines included in `povlines()`:

queries	
r(query_ys_#)	Years queried
r(query_pl_#)	Poverty lines queried
r(query_ct_#)	Countries and coverages queried
r(query_ds_#)	Whether aggregation was used
r(query_#)	Concatenation of the queries above
API parts	
r(server)	Protocol (http://) and server name
r(site_name)	Site names
r(handler)	Action handler
r(base)	Concatenation of server, site name, and handler
additional info	
r(queryfull_#)	Concatenation of base and query_# by "?"
r(np1)	Number of poverty lines
pcn_query	Global with query information in case povcalnet fails. disp "\${pcn_query}" to display

3.8 Installation

The official and stable version of the `povcalnet` command is available at the Boston College Statistical Software Components (SSC) archive, by typing `ssc install povcalnet` in the command window of Stata.⁸

In addition, the development version, which includes the latest updates and features, can be found at the World Bank Github repository, [worldbank/povcalnet](https://github.com/worldbank/povcalnet).⁹ We highly recommend the use of the development version, even though it is not as stable as the version on SSC. This is because the development version is being updated and improved as bugs and issues are discovered. In order to efficiently download this version from Github, we recommend the use of the `github` Stata command developed by E. F. Haghish. The user may execute the following lines in Stata.

```
net install github, from("https://haghish.github.io/github/")
github install worldbank/povcalnet
```

8. There is also a parallel package in R available in CRAN with the same features as the `povcalnet` Stata command (`install.packages("povcalnetR")`)

9. Comments, suggestions, and pull requests are welcome at any time.

4 Empirical applications

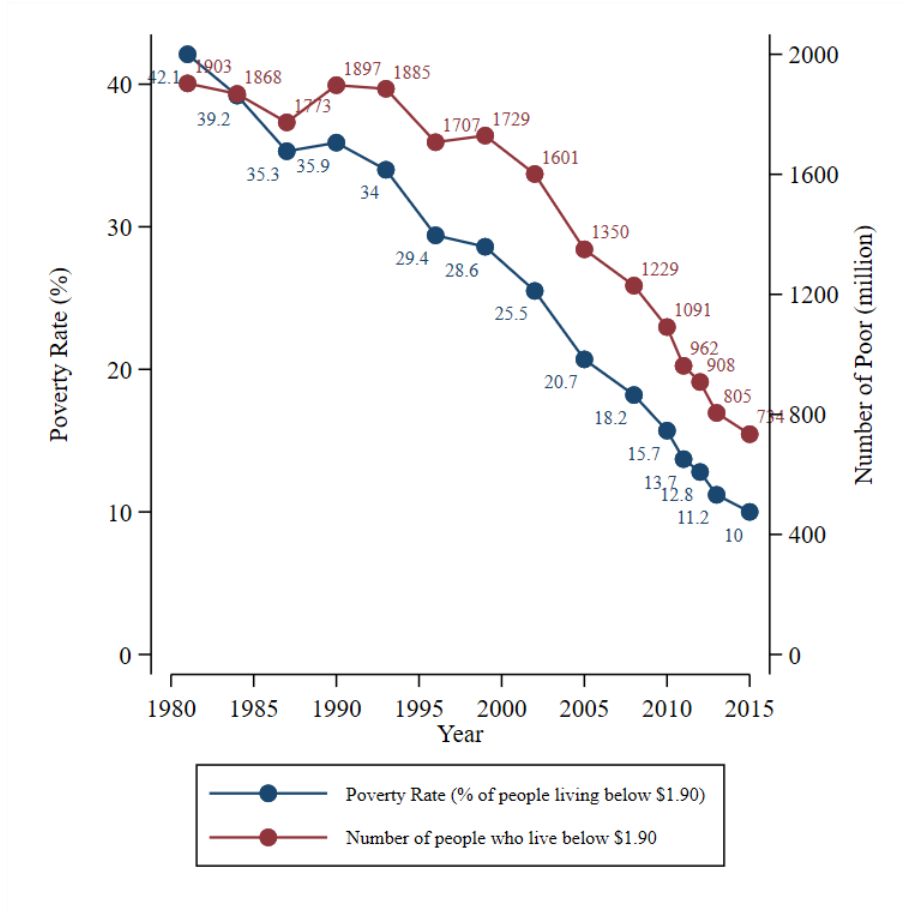
The examples below are intended to present both the scope of the `povcalnet` command and the most recent figures on global poverty as estimated by the World Bank. The Stata code for the figures is provided in the [annex](#).

4.1 Baseline poverty estimates for reference years

The first example shows how to obtain the global estimates of extreme poverty at the international poverty line (IPL) of \$1.9 per day in the reference years. Figure 1 (see the code given in `povcalnet_examples example01`) plots the trend in the global poverty rate and the number of poor between 1990 and 2015. In 2015, the latest available reference year, 10% of the world's population, or 734 million people, lived below the IPL. There is considerable variation in the poverty rate across regions; e.g. 41.4% of the population in Sub-Saharan Africa lives below the IPL, compared with 2.3% in East Asia.

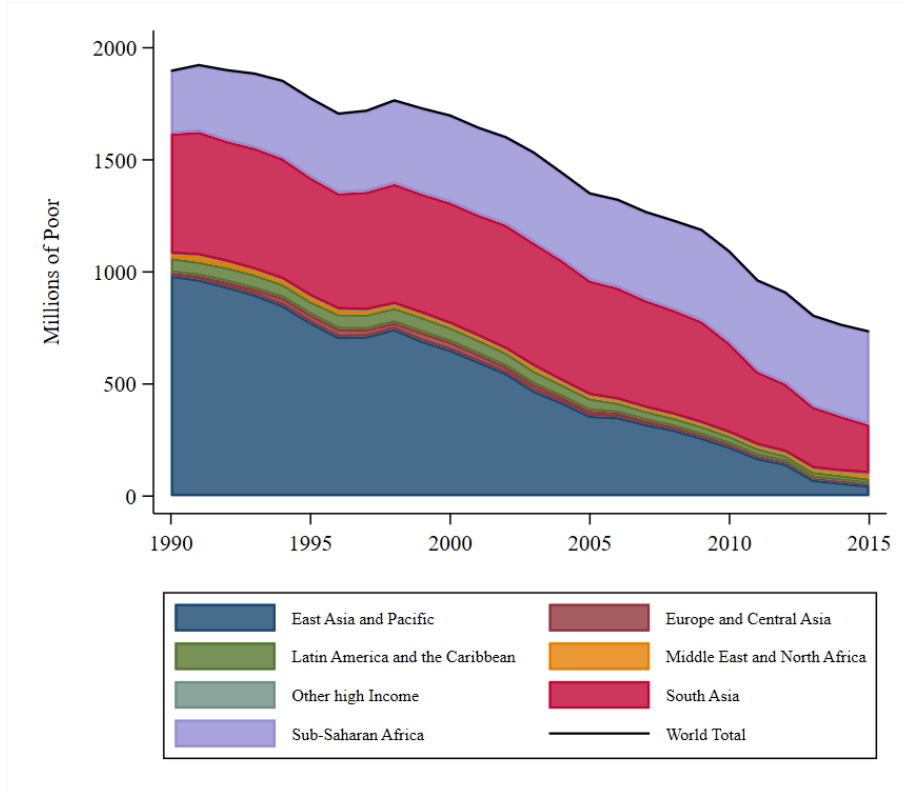
Figure 1 shows considerable progress in reducing the number of people in extreme poverty from 1.9 billion in 1990 to 734 million in 2015. This reduction in the poverty count has been achieved despite an increase in the world population by more than 2 billion people over this period (see discussion in World Bank (2018a)). In terms of the poverty rate, the share of the global population below the IPL dropped from a third in 1990 to a tenth in 2015.

Figure 1: Global Poverty Trends 1990-2015



Similarly, the number of poor people by region can easily be shown as in Figure 2 (see code in `povcalnet_examples/example02`). In 2015, Sub-Saharan Africa accounted for more than half of the global poor, and together with South Asia for more than 85 percent. It is clear that the reduction in global poverty was driven by rapid progress in East Asia. In recent years, the number of poor has also fallen steadily in South Asia. This stands in sharp contrast with Sub-Saharan Africa, where the total number of poor people has actually been increasing over time. As discussed in World Bank (2018a), this shift in the geography of global poverty from high-growth (East Asia and recently South Asia) to low-growth (Sub-Saharan Africa) regions also implies a likely slowdown in the future reduction of global poverty.

Figure 2: Number of Poor People, by Region



4.2 Estimation with user-specified poverty lines

A key feature of PovcalNet is the ability to estimate poverty at any line. For example, `povcalnet, region(lac) year(last) povline(3.2 5.5 15)` loads the latest survey year for all countries in Latin America and the Caribbean, for three different poverty thresholds. With a few additional modifications that are explained in `povcalnet_examples example03`, the distribution of population across three classes—namely, the poor (at two poverty lines), the vulnerable, and the middle class—can be illustrated as in Figure 3. Following World Bank (2018b), the middle class is defined as living on more than \$15 per person per day.

Uruguay has the largest middle class, while Honduras has the smallest. In Colombia and the Dominican Republic, a similar share of the population can be considered middle class, although poverty is considerably higher in Colombia than in the Dominican Republic (27.6% vs. 19.9%, combining the headcount ratios at the two poverty lines). This is consistent with Colombia experiencing greater inequality than the Dominican Republic (for instance, the Gini index in 2017 is 49.7 in Colombia, compared with 45.7 in the Dominican Republic in 2016).

Figure 3: Income Distribution in Latin America and the Caribbean, by Country

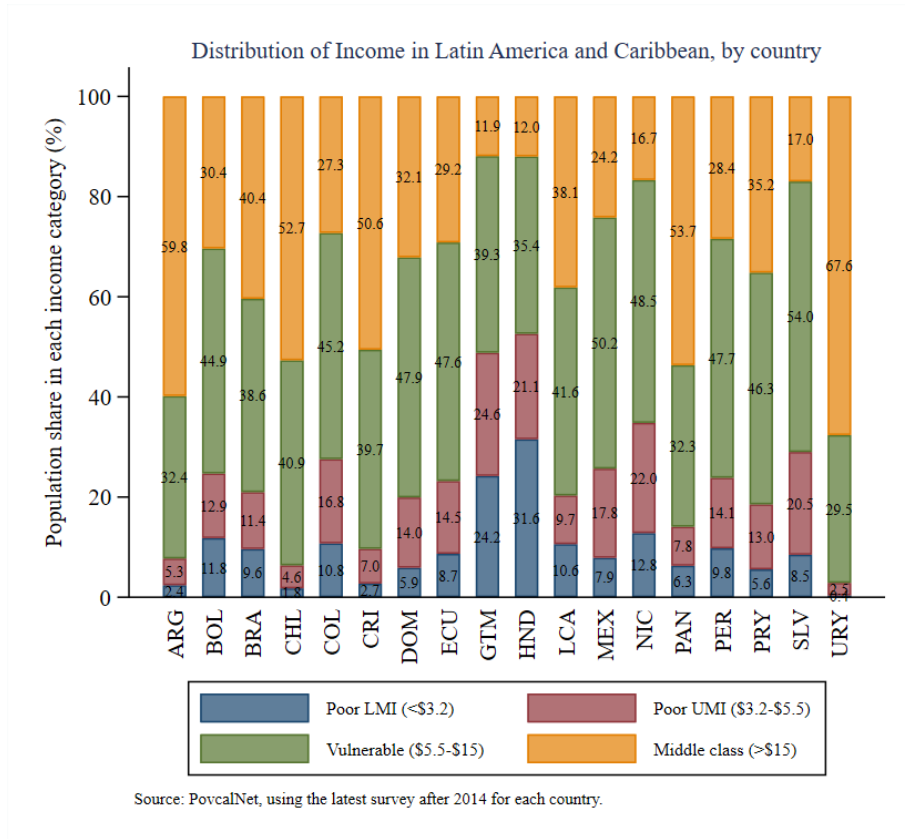


Figure 3 above uses the latest survey year estimates, which are from 2017 in most countries, but from 2014 and 2016 for some countries.¹⁰ To compare across countries in the same reference year, users need to query the reference year estimates instead. The command `povcalnet, region(lac) year(last) povline(3.2 5.5 15) fillgaps` returns the country estimates that have been lined-up to 2015 (the latest reference year).¹¹

4.3 Measures of inequality

PovcalNet reports several measures of inequality, such as the Gini index, the mean log deviation and the income or consumption shares of the decile groups. Inequality measures are only reported in the survey years, since the reference year estimation assumes distribution-neutral growth. Querying

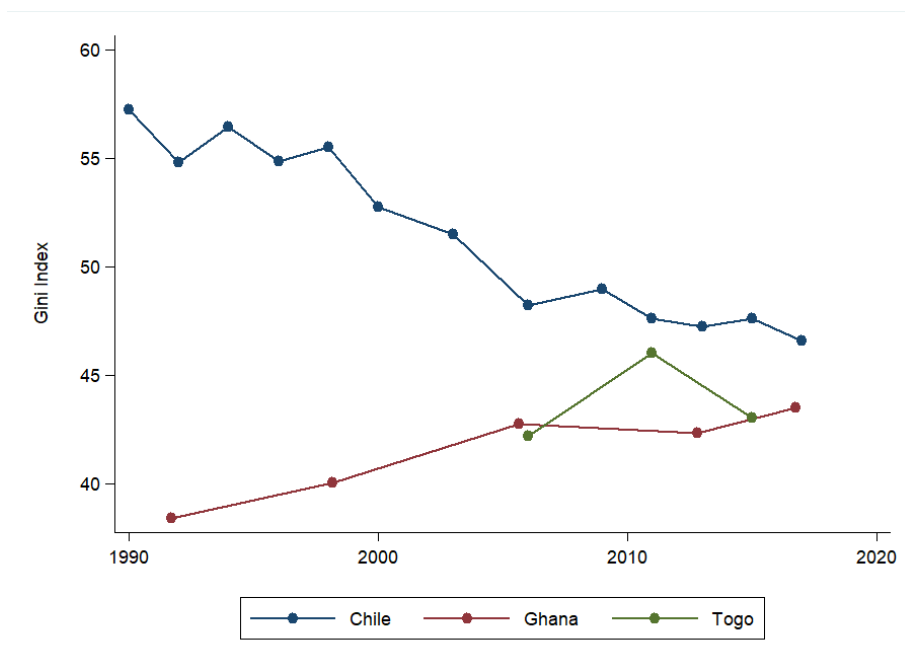
```
povcalnet, country(chl gha tgo) year(all) clear
```

10. Figure 3 includes only countries that had an income survey in 2014 or later. For most countries in the region, the last survey in PovcalNet is from 2017. For the Dominican Republic, Mexico and St. Lucia, the latest surveys are in 2016, and for Guatemala and Nicaragua in 2014.

11. The latest reference year estimate available for Venezuela is for 2013 due to the unavailability of national

loads all available inequality measures for Chile, Ghana and Togo.¹² With some more manipulations (see `povcalnet_examples example04`), it is straightforward to produce Figure 4.

Figure 4: Gini Indices for Chile, Ghana and Togo



In Chile, inequality has been on a downward trend since the 1990s, dropping by more than 10 points between 1990 and 2017.¹³ In contrast, inequality widened in Ghana between the mid-1990s and mid-2000s, with a plateauing in recent years. In Togo, which has a much shorter time series, inequality rose between 2006 and 2011, but turned around thereafter. When comparing the level of the Gini index across the three countries, it is important to bear in mind that Chile uses income, while Ghana and Togo use consumption. Inequality tends to be greater when measured in terms of income than in terms of consumption.¹⁴

PovcalNet also reports the income or consumption shares of the decile groups, which can be used to analyze distributional changes in more detail. In Figure 5, we have drawn (anonymous) growth incidence curves (GIC) for these three countries.¹⁵ The GIC shows the

accounts data after 2014.

12. Chile uses income, while the other two countries use consumption expenditure as their welfare measure.

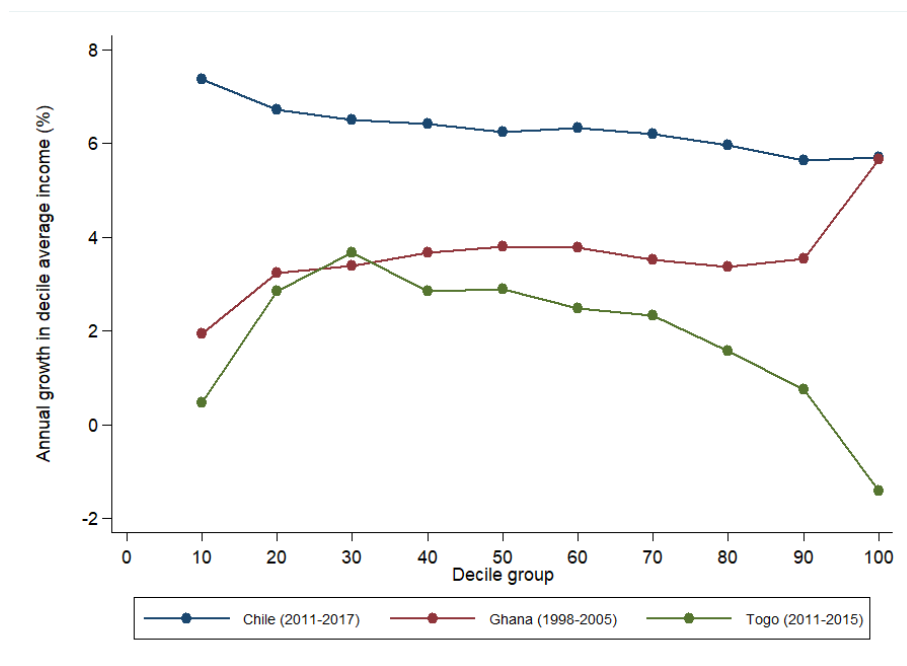
13. For a discussion of the decline in inequality in Latin America, see López-Calva and Lustig (2010)

14. In a group of Latin American countries with both income and consumption surveys, Alvaredo and Gasparini (2015) estimate a multiplicative factor of 0.861 to scale down the income-based Gini indices.

15. Ravallion and Chen (2003) define the GIC in terms of the growth rate in incomes for a percentile (e.g. the 10th percentile). In contrast, we use the mean income of the quantile group (e.g. the bottom 10%) (as in Lakner and Milanovic (2016)). An anonymous GICs ignores the identity of people who make up the quantile groups in the initial and final years. As long as there is some mobility in the distribution, the quantile groups in these two years would consist of different persons.

growth rate of mean income between the initial and final year (varies across countries) for a given quantile group (e.g. the bottom 10%). A downward-sloping GIC is associated with a pro-poor or equalizing pattern of growth.

Figure 5: Income or Consumption Growth for Chile, Ghana and Togo, by Decile Group



Between 2011 and 2017, Chile showed strong growth in income, which was slightly pro-poor. While the bottom 10% grew at 7.4% per year, the top 10% grew at 5.7%. During an earlier period in Ghana (from 1998 to 2005), growth was pro-rich with the top 10% growing at a rate close to 6% and the bottom 10% growing at around 2%.

These patterns in the growth incidence curves are consistent with the trend in the Gini index observed in Figure 4. In Chile, the Gini index decreased modestly from 47.6 in 2011 to 46.6 in 2017. Ghana saw an increase of more than 2 points, from 40.1 in 1998 to 42.8 in 2005. The Gini index in Togo fell by more than 2 points from 46.0 in 2011 to 43.1 in 2015.

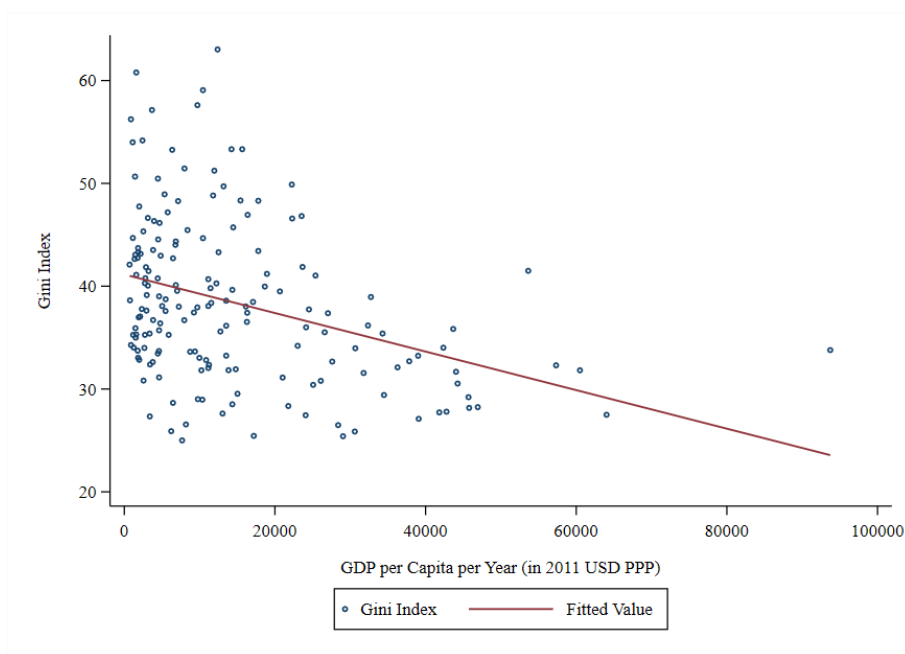
The decile shares can also be used to compute alternative measures of inequality and inclusive growth. For example, the Palma ratio, defined as the ratio of the income share of the top 10% to that of the bottom 40%, is easily calculated (Cobham and Sumner (2013), Palma (2016)). Furthermore, the growth rate of the bottom 40% is also the World Bank’s measure of inclusive growth (referred to as “Shared Prosperity”), and also features in the Sustainable Development Goal 10.1.

4.4 Combining PovcalNet with World Development Indicators

The indicators downloaded using `povcalnet` can be easily merged with any other Stata dataset. Here we illustrate the use with the World Development Indicators (WDI), which can be queried in Stata using the `wbopendata` package.¹⁶

Figure 6 shows the relationship between the Gini index (for the latest available year) and GDP per capita across countries. This replicates Figure 24.1 in Ferreira and Ravallion (2009), and confirms their main two findings: There is a negative (and statistically significant) association between inequality and mean income across countries. Furthermore, the variance of the Gini index declines with mean income. As Ferreira and Ravallion note, this implies that high inequality is primarily found in poor countries, while richer countries experience lower levels of inequality.

Figure 6: Relationship between inequality and GDP



5 Conclusion

This paper has described the Stata command `povcalnet`, which allows users to query the World Bank’s global poverty and inequality measures directly from within Stata. The World

16. Some of the default PovcalNet indicators are also included in the WDI. For example, the poverty headcount ratios in the survey years at \$1.9 (indicator `SI.POV.DDAY`), \$3.2 (`SI.POV.LMIC`) and \$5.5 (`SI.POV.UMIC`), as well as the Gini index (`SI.POV.GINI`). Querying through `povcalnet` offers the flexibility of choosing own poverty lines, includes additional welfare measures (e.g. poverty gaps, mean log deviation), and provides additional details about the survey estimate (e.g. whether it is income or consumption).

Bank's poverty estimates are of strong public interest; for example they form the basis for monitoring countries' progress towards the Sustainable Development Goals. The tool offers considerable flexibility with the ability to estimate poverty at any line, for individual countries, or any group of countries. Accessing PovcalNet directly from within Stata improves the usability of the World Bank's global poverty and inequality numbers, as we have illustrated with several examples in this paper.

References

- Alvaredo, Facundo, and Leo Gasparini. 2015. “Recent Trends in Inequality and Poverty in Developing Countries.” In *Handbook of Income Distribution*, edited by Anthony B. Atkinson and Francois Bourguignon, vol. 2. Elsevier.
- Atamanov, Aziz, Dean M. Jolliffe, Christoph Lakner, and Espen Beer Prydz. 2018. *Purchasing Power Parities Used in Global Poverty Measurement*. Global Poverty Monitoring Technical Note Series 5. The World Bank, September. <https://ideas.repec.org/p/wbk/wbgpmt/5.html>.
- Cobham, Alex, and Andy Sumner. 2013. *Is It All About the Tails? The Palma Measure of Income Inequality*. Center for Global Development Working Paper 343. Center for Global Development.
- Ferreira, Francisco H. G., and Martin Ravallion. 2009. “Poverty and Inequality: The Global Context.” In *The Oxford Handbook of Economic Inequality*, edited by Wiemer Salverda, Brian Nolan, and Timothy M. Smeeding. Oxford University Press.
- Jolliffe, Dean, Peter Lanjouw, Shaohua Chen, Aart Kraay, Christian Meyer, Mario Negre, Espen Beer Prydz, Renos Vakis, and Kyla Wethli. 2015. *A Measured Approach to Ending Poverty and Boosting Shared Prosperity: Concepts, Data, and the Twin Goals*. Policy Research Report. World Bank.
- Lakner, Christoph, and Branko Milanovic. 2016. “Global Income Distribution: From the Fall of the Berlin Wall to the Great Recession.” *World Bank Economic Review* 30 (2): 203–232.
- López-Calva, Luis F., and Nora Lustig, eds. 2010. *Declining Inequality in Latin America: A Decade of Progress?* United Nations Development Programme, Brookings Institution.
- Palma, José Gabriel. 2016. “Measuring Income Inequality: Comments on ‘Do Nations Just Get the Inequality They Deserve? The “Palma Ratio” Reexamined.’” In *Inequality and Growth: Patterns and Policy*, edited by Kaushik Basu and Joseph E. Stiglitz, vol. 2. Palgrave Macmillan.
- Ravallion, Martin, and Shaohua Chen. 2003. “Measuring pro-poor growth.” *Economics Letters* 78 (1): 93–99.
- World Bank. 2016. *Poverty and Shared Prosperity 2016: Taking on Inequality*. World Bank Group.
- . 2018a. *Poverty and Shared Prosperity 2018: Piecing Together the Poverty Puzzle*. World Bank Group.
- . 2018b. *Riding the Wave: An East Asian Miracle for the 21st Century*. World Bank Group.
- Zhao, Qinghua. 2018. “PovcalNet API: Beta 201809.” <http://iresearch.worldbank.org/PovcalNet/docs/PovcalNet%20API.pdf>. (accessed: 01.10.2018).

6 Annex: Economy codes and regional classification used in PovcalNet

6.1 Country codes

Code	Name	Region	Code	Name	Region
ALB	Albania	ECA	LSO	Lesotho	SSA
DZA	Algeria	MNA	LBR	Liberia	SSA
AGO	Angola	SSA	LTU	Lithuania	ECA
ARG	Argentina	LAC	LUX	Luxembourg	OHI
ARM	Armenia	ECA	MDG	Madagascar	SSA
AUS	Australia	OHI	MWI	Malawi	SSA
AUT	Austria	OHI	MYS	Malaysia	EAP
AZE	Azerbaijan	ECA	MDV	Maldives	SAS
BGD	Bangladesh	SAS	MLI	Mali	SSA
BLR	Belarus	ECA	MLT	Malta	OHI
BEL	Belgium	OHI	MRT	Mauritania	SSA
BLZ	Belize	LAC	MUS	Mauritius	SSA
BEN	Benin	SSA	MEX	Mexico	LAC
BTN	Bhutan	SAS	FSM	Micronesia, Federated States of	EAP
BOL	Bolivia	LAC	MDA	Moldova	ECA
BIH	Bosnia and Herzegovina	ECA	MNG	Mongolia	EAP
BWA	Botswana	SSA	MNE	Montenegro	ECA
BRA	Brazil	LAC	MAR	Morocco	MNA
BGR	Bulgaria	ECA	MOZ	Mozambique	SSA
BFA	Burkina Faso	SSA	MMR	Myanmar	EAP
BDI	Burundi	SSA	NAM	Namibia	SSA
CPV	Cabo Verde	SSA	NPL	Nepal	SAS
CMR	Cameroon	SSA	NLD	Netherlands	OHI
CAN	Canada	OHI	NIC	Nicaragua	LAC
CAF	Central African Republic	SSA	NER	Niger	SSA
TCD	Chad	SSA	NGA	Nigeria	SSA
CHL	Chile	LAC	MKD	North Macedonia	ECA
CHN	China	EAP	NOR	Norway	OHI
COL	Colombia	LAC	PAK	Pakistan	SAS
COM	Comoros	SSA	PAN	Panama	LAC
COD	Congo, Democratic Republic of	SSA	PNG	Papua New Guinea	EAP
COG	Congo, Republic of	SSA	PRY	Paraguay	LAC
CRI	Costa Rica	LAC	PER	Peru	LAC
CIV	Cote d'Ivoire	SSA	PHL	Philippines	EAP
HRV	Croatia	ECA	POL	Poland	ECA

Code	Name	Region	Code	Name	Region
CYP	Cyprus	OHI	PRT	Portugal	OHI
CZE	Czech Republic	ECA	ROU	Romania	ECA
DNK	Denmark	OHI	RUS	Russian Federation	ECA
DJI	Djibouti	MNA	RWA	Rwanda	SSA
DOM	Dominican Republic	LAC	WSM	Samoa	EAP
ECU	Ecuador	LAC	STP	Sao Tome and Principe	SSA
EGY	Egypt, Arab Republic of	MNA	SEN	Senegal	SSA
SLV	El Salvador	LAC	SRB	Serbia	ECA
EST	Estonia	ECA	SYC	Seychelles	SSA
SWZ	Eswatini	SSA	SLE	Sierra Leone	SSA
ETH	Ethiopia	SSA	SVK	Slovak Republic	ECA
FJI	Fiji	EAP	SVN	Slovenia	ECA
FIN	Finland	OHI	SLB	Solomon Islands	EAP
FRA	France	OHI	ZAF	South Africa	SSA
GAB	Gabon	SSA	SSD	South Sudan	SSA
GMB	Gambia, The	SSA	ESP	Spain	OHI
GEO	Georgia	ECA	LKA	Sri Lanka	SAS
DEU	Germany	OHI	LCA	St. Lucia	LAC
GHA	Ghana	SSA	SDN	Sudan	SSA
GRC	Greece	OHI	SUR	Suriname	LAC
GTM	Guatemala	LAC	SWE	Sweden	OHI
GIN	Guinea	SSA	CHE	Switzerland	OHI
GNB	Guinea-Bissau	SSA	SYR	Syrian Arab Republic	MNA
GUY	Guyana	LAC	TJK	Tajikistan	ECA
HTI	Haiti	LAC	TZA	Tanzania	SSA
HND	Honduras	LAC	THA	Thailand	EAP
HUN	Hungary	ECA	TLS	Timor-Leste	EAP
ISL	Iceland	OHI	TGO	Togo	SSA
IND	India	SAS	TON	Tonga	EAP
IDN	Indonesia	EAP	TTO	Trinidad and Tobago	LAC
IRN	Iran, Islamic Republic of	MNA	TUN	Tunisia	MNA
IRQ	Iraq	MNA	TUR	Turkey	ECA
IRL	Ireland	OHI	TKM	Turkmenistan	ECA
ISR	Israel	OHI	TUV	Tuvalu	EAP
ITA	Italy	OHI	UGA	Uganda	SSA
JAM	Jamaica	LAC	UKR	Ukraine	ECA
JPN	Japan	OHI	GBR	United Kingdom	OHI
JOR	Jordan	MNA	USA	United States	OHI
KAZ	Kazakhstan	ECA	URY	Uruguay	LAC
KEN	Kenya	SSA	UZB	Uzbekistan	ECA
KIR	Kiribati	EAP	VUT	Vanuatu	EAP
KOR	Korea, Republic of	OHI	VEN	Venezuela, Republica Bolivariana de	LAC
XKX	Kosovo	ECA	VNM	Vietnam	EAP

Code	Name	Region	Code	Name	Region
KGZ	Kyrgyz Republic	ECA	PSE	West Bank and Gaza	MNA
LAO	Lao People's Democratic Republic	EAP	YEM	Yemen, Republic of	MNA
LVA	Latvia	ECA	ZMB	Zambia	SSA
LBN	Lebanon	MNA	ZWE	Zimbabwe	SSA

6.2 Regions

Region	Description
EAP	East Asia and Pacific
ECA	Europe and Central Asia
OHI	Other High Income
LAC	Latin America and the Caribbean
MNA	Middle East and North Africa
SAS	South Asia
SSA	Sub-Saharan Africa

6.3 Reproducible examples

Stata code 1: Global Poverty Trends 1990-2015

```
povcalnet wb, clear
keep if year > 1989
keep if regioncode == "WLD"
gen poorpop = headcount*population
gen hcpercent = round(headcount*100, 0.1)
gen poorpopround = round(poorpop, 1)
tway (sc hcpercent year, yaxis(1) mlab(hcpercent) ///
      mlabpos(7) mlabsize(vsmall) c(1)) ///
      (sc poorpopround year, yaxis(2) mlab(poorpopround) ///
      mlabsize(vsmall) mlabpos(1) c(1)), ///
ytl("Poverty Rate (%) " " ", size(small) axis(1)) ///
ylab(0(10)40, labs(small) nogrid angle(0) axis(1)) ///
ytl("Number of Poor (million)", size(small) axis(2)) ///
ylab(0(400)2000, labs(small) angle(0) axis(2)) ///
xlabel(,labs(small)) xtitle("Year", size(small)) ///
graphregion(c(white)) ysize(5) xsize(5) ///
legend(order(
1 "Poverty Rate (% of people living below $1.90)" ///
2 "Number of people who live below $1.90") si(vsmall) ///
row(2)) scheme(s2color)
```

Stata code 2: Millions of poor by region (reference year)

```
povcalnet wb, clear
keep if year > 1989
```

```

gen poorpop = headcount * population
gen hcpercent = round(headcount*100, 0.1)
gen poorpopround = round(poorpop, 1)
encode region, gen(rid)
levelsof rid, local(regions)
foreach region of local regions {
    local legend = "'legend' 'region' ': label rid 'region'"
}
keep year rid poorpop
reshape wide poorpop,i(year) j(rid)
foreach i of numlist 2(1)7{
    egen poorpopacc`i`=rowtotal(poorpop1 - poorpop`i`)
}
twoway (area poorpop1 year) ///
    (rarea poorpopacc2 poorpop1 year) ///
    (rarea poorpopacc3 poorpopacc2 year) ///
    (rarea poorpopacc4 poorpopacc3 year) ///
    (rarea poorpopacc5 poorpopacc4 year) ///
    (rarea poorpopacc6 poorpopacc5 year) ///
    (rarea poorpopacc7 poorpopacc6 year) ///
    (line poorpopacc7 year, lwidth(midthick) lcolor(gs0)), ///
ytitle("Millions of Poor" " ", size(small)) ///
xtitle(" " " ", size(small)) scheme(s2color) ///
graphregion(c(white)) ysize(7) xsize(8) ///
ylabel(,labs(small) nogrid angle(vertical)) ///
xlabel(,labs(small)) legend(order('legend') si(vsmall))

```

Stata code 3: Distribution of Income in Latin America and Caribbean, by country

```

povcalnet, region(lac) year(last) povline(3.2 5.5 15) clear
keep if datatype==2 & year>=2014 // keep income surveys
keep povertyline countrycode countryname year headcount
replace povertyline = povertyline*100
replace headcount = headcount*100
tostring povertyline, replace format(%12.0f) force
reshape wide headcount,i(year countrycode countryname ) j(povertyline)
string
gen percentage_0 = headcount320
gen percentage_1 = headcount550 - headcount320
gen percentage_2 = headcount1500 - headcount550
gen percentage_3 = 100 - headcount1500
keep countrycode countryname year percentage_*
reshape long percentage_,i(year countrycode countryname ) j(category)
la define category 0 "Poor LMI (< $3.2)" 1 "Poor UMI ($3.2-$5.5)" ///
    2 "Vulnerable ($5.5-$15)" 3 "Middle class (> $15)"
la val category category
la var category ""
local title "Distribution of Income in Latin America and Caribbean, by
country"
local note "Source: PovcalNet, using the latest survey after 2014 for
each country."
local yti "Population share in each income category (%)"
graph bar (mean) percentage, inten(*0.7) o(category) ///
o(countrycode, lab(labsi(small) angle(vertical))) stack asy ///

```

```

blab(bar, pos(center) format(%3.1f) si(tiny))          ///
ti("'title'", si(small)) note("'note'", si(*.7))      ///
graphregion(c(white)) ysize(6) xsize(6.5)             ///
  legend(si(vsmall) r(3)) yti("'yti'", si(small))    ///
ylab(,labs(small) nogrid angle(0)) scheme(s2color)

```

Stata code 4: Gini Indices for Chile, Ghana and Togo

```

program example06
povcalnet, country(chl gha tgo) year(all) clear
  replace gini = gini * 100
  keep if datayear > 1989
  twoway (connected gini datayear if countrycode == "CHL") ///
    (connected gini datayear if countrycode == "GHA")      ///
    (connected gini datayear if countrycode == "TGO"),     ///
  ytitle("Gini Index" " ", size(small))                  ///
  xtitle(" " " ", size(small)) ylabel(,labs(small) nogrid ///
  angle(verticle)) xlabel(,labs(small))                  ///
  graphregion(c(white)) scheme(s2color)                  ///
  legend(order(1 "Chile" 2 "Ghana" 3 "Togo") si(small) row(1))

```

Stata code 5: Welfare Growth for Chile, Ghana and Togo, by decile

```

povcalnet, country(chl gha tgo) year(all) clear
  reshape long decile, i(countrycode datayear) j(dec)
  egen panelid=group(countrycode dec)
  replace datayear=int(datayear)
  xtset panelid datayear
  replace decile=10*decile*mean
  gen g = (((decile/L6.decile)^(1/6))-1)*100
  replace g = (((decile/L7.decile)^(1/7))-1)*100 if countrycode == "GHA"
  replace g = (((decile/L4.decile)^(1/4))-1)*100 if countrycode == "TGO"

  replace dec=10*dec

  twoway (sc g dec if datayear==2017 & countrycode=="CHL", c(1)) ///
    (sc g dec if datayear==2005 & countrycode=="GHA", c(1))      ///
    (sc g dec if datayear==2015 & countrycode=="TGO", c(1)),    ///
  yti("Annual growth in decile average income (%) " " ",      ///
  size(small)) xlabel(0(10)100,labs(small))                  ///
  xtitle("Decile group", size(small)) graphregion(c(white))   ///
  legend(order(1 "Chile (2011-2017)"                          ///
  2 "Ghana(1998-2005)" 3 "Togo (2011-2015)"))                ///
  si(vsmall) row(1)) scheme(s2color)

```

Stata code 6: Relationship between inequality and GDP

```

set checksum off
wbopendata, indicator(NY.GDP.PCAP.PP.KD) long clear
tempfile PerCapitaGDP
save 'PerCapitaGDP', replace

povcalnet, povline(1.9) country(all) year(last) clear
keep countrycode countryname year gini

```



```

drop if gini == -1
* Merge Gini coefficient with per capita GDP
merge m:1 countrycode year using 'PerCapitaGDP', keep(match)
replace gini = gini * 100
drop if ny_gdp_pcap_pp_kd == .
twoway (scatter gini ny_gdp_pcap_pp_kd, mfcolor(%0)          ///
       msize(vsmall)) (lfit gini ny_gdp_pcap_pp_kd),        ///
       ytitle("Gini Index" " ", size(small))                ///
       xtitle(" " "GDP per Capita per Year (in 2011 USD PPP)", ///
       size(small)) graphregion(c(white)) ysize(5) xsize(7) ///
       ylabel(,labs(small) nogrid angle(verticle))          ///
       xlabel(,labs(small)) scheme(s2color)                  ///
       legend(order(1 "Gini Index" 2 "Fitted Value") si(small))

```