
Published by

Lustig, Nora.
Project MUSE. muse.jhu.edu/book/61437.

For additional information about this book
https://muse.jhu.edu/book/61437
Chapter 9

ANALYZING THE IMPACT OF FISCAL POLICY ON ETHNO-RACIAL INEQUALITY

Rodrigo Aranda and Adam Ratzlaff

As shown in previous chapters, the Commitment to Equity (CEQ) analysis provides researchers with a comprehensive and comparable set of indicators to determine the impacts of fiscal intervention on poverty and inequality. However, inequality may take many different forms and be based on biases that are beyond the control of individuals. Race, gender, location, and parental characteristics can have important implications for the economic and social outcomes of individuals. In Latin America, ethno-racial inequalities are particularly prevalent; indigenous peoples and African descendants are faced with higher rates of poverty, lower average incomes, and lower access to services. In an effort to determine if government fiscal interventions are exacerbating or reducing ethno-racial inequalities in Latin America, the Inter-American Development Bank (IDB) has partnered with the CEQ Institute to finance the adoption of the CEQ analysis to explore the impacts of fiscal policies on ethno-racial inequality in the Latin America and Caribbean region (LAC).

This chapter and the corresponding component of the CEQ Master Workbook, section F, “Results by Ethnicity and Race,” were prepared as part of the Inter-American Development Bank’s technical cooperation “Improving Race and Ethnicity Data Instruments for Policy Analysis and Formulation” (RG-T1906), led by Judith Morrison, Senior Advisor, Gender and Diversity Division (SCL/GDI). Through this technical cooperation, funding was made available for the Inter-American Development Bank–Commitment to Equity Incidence of Taxes and Social Spending by Ethnicity and Race Study for Bolivia, Brazil, Guatemala, and Uruguay.


2Molinas Vega and others (2012).

3de Ferranti and others (2004); Hall and Patrinos (2006); Ñopo (2012).
A necessary first step in calculating the impact of fiscal policy on reducing ethno-racial inequality is to determine the appropriate indicators for measuring ethno-racial inequalities and what measures should be used to determine the impact of fiscal policy on these indicators. To do this, we will utilize the measures discussed by Lustig.\textsuperscript{4} To measure levels of inequality across ethno-racial lines, four different measures will be utilized:

1. *Income gaps*, in terms of the mean incomes or share of income held by different ethno-racial populations, provide for absolute and relative sizes of the ethno-racial inequality at the aggregate level.

2. *Contribution to overall inequality* can be determined using a decomposable measure of inequality such as the Theil coefficient. The benefit to the Theil coefficient is that it can be decomposed to determine the level of national inequality due to inter- and intra-ethno-racial group inequalities. This is particularly important as it provides us with a better understanding of the dynamics not only between ethno-racial groups, but also within these populations. It is important to note here that policies may reduce inequality between groups while exacerbating inequalities within specific populations.

3. *Inequality of opportunity* is a concept popularized by Roemer and further applied in Ferreira and Robalino and Molinas Vega and others to determine the extent to which characteristics or circumstances outside of an individual’s control (for example, not due to personal effort or preference) affect his or her economic and social outcomes.\textsuperscript{5} These circumstances frequently include characteristics such as gender, location (urban/rural), levels of parental education, and race or ethnicity. In a society that is ethno-racially equal or colorblind, one would expect to see no inequality of opportunity due to ethno-racial differences. Here, inequality of opportunity can be used to assess the extent to which fiscal policy equalizes opportunities and reduces inequality. More details on how this is calculated are provided below.

4. *Poverty* headcounts, gaps, and severity measures can be utilized to provide a better understanding of differences in the well-being of different ethno-racial populations with a particular emphasis on what is happening at the bottom of the income distribution. Having data on the different levels and magnitudes of poverty is particularly important in showing what types of policies are benefiting the most disadvantaged segments of the population.

All of the measures indicated above can be calculated using the different income concepts utilized in the CEQ analysis, thereby allowing us to determine the fiscal impact of specific sets of policies on ethno-racial inequality.

\textsuperscript{4}Lustig (2017).

\textsuperscript{5}Roemer (1998); Ferreira and Robalino (2010); Molinas Vega and others (2012).
In order to determine the effectiveness of programs at reducing ethno-racial inequality, two different measures will be utilized to determine if the impact of specific programs or sets of programs help to reduce ethno-racial inequality:

1. *Progressivity* will be determined by calculating the share of benefits going to different ethno-racial groups relative to their respective shares of the population or their respective share of income. A program is deemed to be relatively progressive if the share of benefits received is greater than the disadvantaged group’s share of income (for example, making incomes more equitable) and is considered absolutely progressive if the share of benefits received is greater than their share of the total population.\(^6\)

2. *Pro-disadvantaged group:* While examining progressivity provides a way of measuring if fiscal policy reduces ethno-racial inequality, a targeted poverty reduction policy may appear to be progressive due to the number of individuals of a particular ethno-racial group who are in poverty. Fiscal policy is designated as pro-disadvantaged group if the impact of direct taxes and transfers produces a greater likelihood for members of the disadvantaged group to escape poverty than for advantaged populations.

It is important to note that for a policy to be pro-disadvantaged group, it must violate horizontal equity, or the premise that individuals of equal income should be treated equally. By treating the poor of a particular ethno-racial group differently, a policy violates this criterion.

Section F of the *CEQ Master Workbook (MWB)* (available online in part IV of this Handbook; CEQ Institute [2018]), “Results by Ethnicity and Race,” allows users to produce all of the results necessary to conduct an analysis of the impacts of fiscal policy across ethno-racial lines in one easy-to-use workbook with accompanying Stata ado-file. This workbook presents a compendium of the CEQ main results in a manner that allows for easy interpretation across ethno-racial lines. This chapter describes the different indicators and sheets presented in section F of the *CEQ Master Workbook* (see table 9-1) and details on how to use the *ceqrace.ado* Stata command to produce these results for each sheet of the workbook.

The *ceqrace.ado* Stata command is designed to automatically fill in the values for nineteen of the twenty-six Excel sheets listed in table 9-1. The remaining Excel sheets must be filled in manually. The program allows users to estimate the results for each of these sheets separately in Stata and export them to the Excel workbook. It is also designed to be flexible such that it can match the different data and statistical requirements of each country.

\(^6\)The group that has lower per capita incomes is considered the disadvantaged group in this exercise. In all of the four countries analyzed, the disadvantaged group refers to the indigenous or African descendant population.
In order to utilize the `ceqrace.ado`, it is necessary to have the basic software requirements of Stata 13.0 (or a more recent version) and Microsoft Excel (.xls or .xlsx format). As for data requirements, the program works on Stata datasets with data at the individual level and includes the main variables used in the CEQ framework such as income concepts, taxes, transfers, as well as sociodemographic characteristics of individuals. While the other sections of the CEQ analysis are designed to utilize either individual or household level data, for the analysis by race and ethnicity only individual-level datasets can be utilized due to the need to identify individuals by race.

| Table 9-1 |
| Sheets Presented in Section F: Results by Race and Ethnicity |

**1. Background information**

- F1. Key assumptions*
- F2. Ethno-racial definitions*
- F3. Ethno-racial populations
- F4. Linked information*

**2. Results**

- F5. Population composition
- F6. Income distribution
- F7. Summary poverty rates
- F8. Summary poverty gap rates
- F9. Summary poverty gap squared rates
- F10. Summary inequality indicators
- F11. Mean incomes
- F12. Incidence by decile
- F13. Incidence by income group
- F14. Cross-race incidence**
- F15. Horizontal equity**
- F16. Fiscal profile
- F17. Coverage rates (total population)
- F18. Coverage rates (target population)
- F19. Leakages**
- F20. Mobility matrices
- F21. Education (totals)
- F22. Education (rates)**
- F23. Infrastructure access
- F24. Theil disaggregation
- F25. Inequality of opportunity
- F26. Significance

*Note: Sheets with an * must be filled in manually. Sheets with an ** are filled in automatically using the results from a different sheet. The remaining sheets can be completed using the `ceqrace.ado` command in Stata.
or ethnicity. While some indicators in the analysis are generated at the household level—for example, using the ethno-racial identity of the head of the household for identification purposes—it is preferable to utilize the self-identification method for all individuals in the household.

The main syntax for the command is:

```
ceqrace using filename [weight] [if] [in] [, table(name) options]
```

For each of the different Excel sheets the command asks for the Excel filename, the number of the table, weights, and ethno-racial group identifiers. The race or ethnic group identifiers must be dichotomous variables and should be arranged such that:

- `race1`: White/non-ethnic population
- `race2`: Indigenous population
- `race3`: African descendant population
- `race4`: Other Races/Ethnicities
- `race5`: Non-responses

The program requires that at least two different groups have been defined as dichotomous variables. The remaining options for running the analysis are specific to the sheet and will be discussed in detail below. For a summary of variables, their format, and options to be used with this command, see table 9-2.

It is also important to note that this workbook is preset to produce results using the regional income group definitions as well as country-specific poverty results. Where the country-specific poverty lines are used, authors will input the value of the national extreme and moderate poverty lines in their Stata command. The `ceqrace.ado` command is preset to use the typical poverty lines of US$1.25 purchasing power parity (PPP) per capita per day, US$2.50 PPP per capita per day (extreme poverty), and US$4 PPP per capita per day (moderate poverty), as well as income groupings for the vulnerable (with incomes between US$4 and US$10 PPP per capita per day), for the middle class (with incomes between US$10 and US$50 PPP per capita per day), and for all individuals with per capita per day incomes above US$50 PPP. All of these income groupings utilize the 2005 PPP conversion rate.

Although section F of the CEQ Master Workbook and the `ceqrace.ado` are preset to utilize these income lines, users may opt to change poverty lines to fit their research needs using the option `cut()`. However, it is important to note that this will not change the labels presented in the Excel file. Thus, if users choose to use different income groups, they need to manually adjust these labels; not doing so may cause confusion for end users of the workbook.

---

7 As discussed above, the default options are `cut1(1.25), cut2(2.50), cut3(4.00), cut4(10.00),` and `cut5(50.00)`. 
<table>
<thead>
<tr>
<th>Concept</th>
<th>Option in <code>ceqrace.ado</code></th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethno-racial groups</td>
<td><code>race1(varname)</code></td>
<td>All variables have to be dummies and identify the ethnicity-race of each individual. <code>race1</code> is for Indigenous population, <code>race2</code> is for White/non-ethnic population, <code>race3</code> is for African descendant population, <code>race4</code> other races, and <code>race5</code> for non-responses.</td>
</tr>
<tr>
<td></td>
<td><code>race2(varname)</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>race3(varname)</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>race4(varname)</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>race5(varname)</code></td>
<td></td>
</tr>
<tr>
<td>Income concepts</td>
<td><code>original(varname)</code></td>
<td>These variables must have the per capita income concepts in local currency units.</td>
</tr>
<tr>
<td></td>
<td><code>market(varname)</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>mpluspensions(varname)</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>netmarket(varname)</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>gross(varname)</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>taxable(varname)</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>disposable(varname)</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>consumable(varname)</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>final (varname)</code></td>
<td></td>
</tr>
<tr>
<td>Tax and transfer concepts</td>
<td><code>dtax(varname)</code></td>
<td>These variables must have the tax or transfer concepts in per capita local currency units.</td>
</tr>
<tr>
<td></td>
<td><code>contrib(varname)</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>conypensions(varname)</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>contpensions(varname)</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>noncontrib(varname)</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>flagcct(varname)</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>otransfers(varname)</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>isubsidies(varname)</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>itax(varname)</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>ikeduc(varname)</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>ikhealth(varname)</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>hurban(varname)</code></td>
<td></td>
</tr>
<tr>
<td>PPP conversion options</td>
<td><code>ppp(real)</code></td>
<td>These options accept only numbers; only one of the daily, monthly, and yearly options can be used.</td>
</tr>
<tr>
<td></td>
<td><code>cpibase(real)</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>cpisurvey(real)</code></td>
<td></td>
</tr>
<tr>
<td>Poverty lines and income group cut-offs</td>
<td><code>nextreme(string)</code></td>
<td><code>nextreme( )</code> and <code>nmoderate( )</code> accept numerical values as well as a variable; the values have to be at the same time and currency unit as the income variables.</td>
</tr>
<tr>
<td></td>
<td><code>nmoderate(string)</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>cut1(real)</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>cut2(real)</code></td>
<td></td>
</tr>
<tr>
<td></td>
<td><code>cut3(real)</code></td>
<td></td>
</tr>
<tr>
<td>Concept</td>
<td>Option in <code>ceqrace.ado</code></td>
<td>Description</td>
</tr>
<tr>
<td>------------------</td>
<td>-------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>cut4(real)</td>
<td>The <code>cut( )</code> options allow the user to use different thresholds for poverty in daily PPP. If this option is not used, the program automatically uses the income group cut-offs.</td>
</tr>
<tr>
<td></td>
<td>cut5(real)</td>
<td></td>
</tr>
<tr>
<td>Coverage</td>
<td>cct(varname)</td>
<td>These variables have to be in monetary units at the individual/household level (depending on who receives the benefit) in the same units as income concepts and tax and transfers.</td>
</tr>
<tr>
<td></td>
<td>noncontrib(varname)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>pensions(varname)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>unemploy(varname)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>foodtransfers(varname)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>otransfers(varname)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>health(varname)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>pensions(varname)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>scholarships(varname)</td>
<td></td>
</tr>
<tr>
<td>Target population</td>
<td>tarcct(varname)</td>
<td>These variables have to be dummies that identify the target population for each concept.</td>
</tr>
<tr>
<td></td>
<td>tarncp(varname)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>tarpen(varname)</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>age(varname)</td>
<td>Age variable has to be the age of the individual. The rest of the variables have to be dummies that identify whether each individual satisfies each condition.</td>
</tr>
<tr>
<td></td>
<td>edpre(varname)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>redpre(varname)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>edpri(varname)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>redpri(varname)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>edsec(varname)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>redsec(varname)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>edter(varname)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>redter(varname)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>edpublic(varname)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>edprivate(varname)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>attend(varname)</td>
<td></td>
</tr>
<tr>
<td>Infrastructure</td>
<td>water(varname)</td>
<td>These variables have to be dummies that identify whether the individual lives in a household with access to each specific concept.</td>
</tr>
<tr>
<td>access</td>
<td>electricity(varname)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>walls(varname)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>floors(varname)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>roof(varname)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>sewage(varname)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>roads(varname)</td>
<td></td>
</tr>
<tr>
<td>Household</td>
<td>hhead(varname)</td>
<td>The dataset to use has to be at the individual level; <code>hhead</code> is the dummy variable that identifies who is the household head for each household, and <code>hhid</code> is the variable that uniquely identifies each household in the data.</td>
</tr>
<tr>
<td></td>
<td>hhid(varname)</td>
<td></td>
</tr>
</tbody>
</table>

(continued)
Table 9-2 (continued)

<table>
<thead>
<tr>
<th>Concept</th>
<th>Option in ceqrace.ado</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circumstance</td>
<td>gender(varname)</td>
<td>The gender variable has to take the value of 1 for the gender the user chooses to use.</td>
</tr>
<tr>
<td></td>
<td>urban(varname)</td>
<td>urban is a dummy variable that identifies individuals living in an urban context.</td>
</tr>
<tr>
<td></td>
<td>edpar(varname)</td>
<td>edpar is parents' years of education.</td>
</tr>
<tr>
<td>Survey information</td>
<td>hsize(varname)</td>
<td>hsize is a variable with the number of members of each household.</td>
</tr>
<tr>
<td></td>
<td>psu(varname)</td>
<td>psu and strata are variables that identify primary sampling units and strata, respectively.</td>
</tr>
<tr>
<td></td>
<td>strata(varname)</td>
<td></td>
</tr>
</tbody>
</table>

1 Background Information

The first part of section F of the CEQ Master Workbook requires authors to fill in much of the background information necessary to conduct the general CEQ analysis, information on the different ethno-racial populations that are being analyzed, and some of the relevant background information for analyzing the results of the study. Many of these sheets will need to be filled in manually (without the aid of the ceqrace.ado command).

1.1 Sheet F1. Key Assumptions

Sheet F1 presents the key assumptions utilized in the CEQ analysis. This sheet is highly important for end users of the data as it is critical to have this information available for interpreting the results of the study. While this sheet will need to be filled in manually by authors, it includes similar information as that presented in the “Key Assumptions” sheet featured in section C, “Methodology of the CEQ Master Workbook” (sheet C2). Nonetheless, it is important that authors complete sheet F1 as well, as it allows users to conduct much of the CEQ analysis by race and ethnicity using only the one section of the CEQ Master Workbook and ensures that results are interpreted correctly and accurately.

1.2 Sheet F2. Ethno-Racial Definitions

While some countries clearly define ethno-racial categories that should be utilized for the CEQ fiscal incidence analysis, the definitions vary by country and by survey. Most
Latin American countries have transitioned to using self-identification as the primary method for determining the ethno-racial categorization of individuals or households, although some countries in the region continue to use maternal language as the determinant of ethno-racial group. Additionally, some populations may have multiple identities depending on the context in which they are being considered. Thus, defining how each study examines ethno-racial populations may be an important factor for providing policy recommendations specific to different segments of society. Further questions on how race and ethnicity should be imputed to individuals who are not asked to self-identify, as well as on how to impute race or ethnicity to the household level, are important and can have profound effects on the results of the analysis. In order to ensure that results are comparable to other studies as well as to verify that the definitions used are understandable to a broader audience, authors should clarify how the different ethno-racial populations are defined for the purpose of their study.

Additionally, this sheet includes information not only on the survey being used, but also on national census results. Differences between the definitions of different ethno-racial categories, the manner in which the question on ethno-racial identity was asked, or how the sample was constructed in the census as compared to the survey being utilized may lead to findings that contradict what would be expected based on census results. Having information on how ethno-racial populations are defined in these two datasets allows users to see if there are differences and if so what these differences may be.

This sheet must be filled in manually by the authors.

1.3 Sheet F3. Ethno-Racial Populations

Sheet F3 expands upon the information presented in the previous sheet by looking at the size of each ethno-racial population and comparing it to census figures. This allows researchers to have a better understanding of the representativeness of the survey being used (when compared to census results) and allows them to express whether they believe that the trends that are seen across ethno-racial lines are truly representative (both in magnitude as well as direction) of national results. Knowing the differences between national surveys and censuses is particularly important given that the sample design in some countries may not take race and ethnicity into account. While most of this sheet is completed using the ceqrace.ado Stata command, data from national censuses must be completed manually by authors.

To fill in this sheet using the ceqrace.ado Stata command, race dummy variables and weights are required. Below is an example of how this can be run:

```
ceqrace [pw = weight] using CEQ_Ethno_Racial_MWB.xlsx, race1(indig) race2(white) race3(afrd) race4(orace) race5(nonrace) table(f3)
```
Where CEQ_Ethno_Racial_MWB.xlsx is the name of the Excel file being used and the race variables are all dichotomous, it is important to note that the table option must include the number of the sheet preceded by an “f” in order to automatically fill in the Excel file.

1.4 Sheet F4. Linked Information

The linked information contained in this sheet provides some additional background information on the different policies that are considered as part of the CEQ analysis. It also allows authors to quickly fill in much of the background information that is necessary to complete tables throughout this workbook. Data that should be filled in by the authors includes information on calculating the conversion rates from local currency units (LCU) to US dollars in 2005 and 2011 purchasing power parity (PPP), information on the national poverty lines used in the country, additional information on the programs that are being analyzed as part of the _CEQ Assessment_, and information on the country’s education system. Data for generating the conversion factors between LCU and 2005 or 2011 PPP can be found in the World Bank’s World Development Indicators.\(^8\) This information will be used to convert LCU into PPP on several sheets throughout section F of the workbook and to convert national poverty lines in LCU into PPP numbers. Official names of the different programs that are being aggregated or used in this section of the analysis should also be provided so that end users are better able to understand the different elements that are being considered as part of the analysis. For education information, it is important that users input the targeted age ranges for different educational levels as this information has important implications for calculating educational enrollment rates (see sheets F21 and F22).

Authors must complete this sheet manually.

2 Results

Part II of section F of the _CEQ Master Workbook_ presents the results of the _CEQ Assessment_ necessary to conduct the analysis across ethno-racial lines in a user-friendly format. This section includes many of the tables and figures that researchers may want to consider when comparing the impact of fiscal policy across ethno-racial lines.

2.1 Sheet F5. Population Composition

An important element in assessing ethno-racial inequality is understanding how the population is distributed across socioeconomic and ethno-racial lines. This sheet presents the population distribution and magnitude disaggregated by decile and income

\(^8\)World Bank (2017).
group across ethno-racial lines for original\textsuperscript{9} and Disposable Income concepts. It is important to note that, although national results will be the same, the ethno-racial results by decile will differ from those presented in section D ("Summary of Results") and section E ("Output Tables") of the CEQ Master Workbook because this worksheet defines deciles nationally and then disaggregates by ethno-racial category rather than presenting the deciles within each ethno-racial group. In other words, the results presented here will express the share of the different population segments in each decile rather than presenting the characteristics of the different ethno-racial groups by decile.

To fill in this sheet using the ceqrace.ado Stata command, it is necessary to have variables generated for original income, Disposable Income, household identifier, consumer price index, purchasing power parity variables, and dummy variables for each ethno-racial category. The syntax should follow:

\begin{verbatim}
ceqrace [pw = weight] using CEQ_Ethno_Racial_MWB.xlsx, race1(indig) race2(white) race3(afrd) race4(orace) race5(nonrace) table(f5) o(y_m) d(y_d) hhid(hhid) ppp(7.65) cpibase(78.661) cpisurvey(105.196) year
\end{verbatim}

Where CEQ_Ethno_Racial_MWB.xlsx is the name of the Excel file being used, the race variables are dichotomous, original(varname) specifies the original income variable in local currency units,\textsuperscript{10} disposable(varname) is Disposable Income, hhid(varname) is the variable that uniquely identifies the household, ppp( ) is the purchasing power parity (PPP) conversion factor (local currency units [LCU] per international dollar, consumption-based) from the year of PPP (usually either 2005 or 2011), cpibase( ) is the consumer price index (CPI) of the base year (year of PPP, usually 2005 or 2011), cpisurvey( ) is the CPI for the year of the household survey, and finally, year indicates that income variables are defined in annual terms (although it is preferable to use annualized data, daily and monthly can also be specified if the author chooses).

\subsection*{2.2 Sheet F6. Income Distribution}

Sheet F6 builds upon the data in sheet F5 by presenting the distribution of income by ethno-racial group as well as nationally. Results are given using both decile and

\textsuperscript{9}Original income might vary depending on whether one is running an analysis using pensions as deferred income (PDI) or pensions as government transfers (PGT) so Market Income or Market Income plus Pensions variables have to be used for this option depending on the scenario.

\textsuperscript{10}Original income might vary depending on whether one is running an analysis using PDI or PGT, so Market Income or Market Income plus Pensions variables have to be used for this option depending on the scenario.
income groups for original and Disposable Income. As with sheet F5, these decile results will differ from the disaggregation presented in sections D and E the CEQ Master Workbook due to the manner in which deciles are defined in this section of the workbook.\(^{11}\)

To fill in this sheet using the `ceqrace.ado` Stata command, it is necessary to have variables generated for original income, Disposable Income, household identifier, consumer price index, and purchasing power parity variables, and dummy variables for each ethno-racial category. The syntax should follow:

```
ceqrace [pw = weight] using CEQ_Ethno_Racial_MWB.xlsx, race1(indig) race2(white) race3(afrd) race4(orace) race5(nonrace) table(f6) o(y[m]) d(y[d]) hhid(hhid) ppp(7.65) cpibase(78.661) cpisurvey(105.196) year
```

Where `CEQ_Ethno_Racial_MWB.xlsx` is the name of the Excel file being used, the race variables are dichotomous, `original(varname)` specifies the original income variable in local currency units,\(^{12}\) `disposable(varname)` is Disposable Income, `hhid()` is the variable that uniquely identifies the household, `ppp()` is the purchasing power parity (PPP) conversion factor (LCU per international dollar, consumption-based) from the year of PPP (usually either 2005 or 2011), `cpibase()` is the consumer price index (CPI) of the base year (year of PPP, usually 2005 or 2011), `cpisurvey()` is the CPI for the year of the household survey, and finally, `year` indicates that income variables are defined in annual terms (daily and monthly can also be used).

### 2.3 Sheet F7. Summary Poverty Rates

Poverty headcount rates are key to determining levels of social exclusion and inequality across ethno-racial lines. Sheet F7 presents poverty headcount rates by race and ethnicity as well as nationally for each of the different core income concepts and generates tables that can be used to demonstrate the impacts of fiscal policy on poverty across ethno-racial lines.

To fill in this sheet using the `ceqrace.ado` Stata command, race dummy variables, weights, `market(varname)`, `mpluspensions(varname)`, `netmarket(varname)`, `gross(varname)`, `taxable(varname)`, `disposable(varname)`, `consumable(varname)`, `nextreme(string)`, and `nmoderate(string)` options are required and the following syntax should be used:

---

\(^{11}\)For description of why decile results may differ, please refer to the discussion of sheet F5.

\(^{12}\)Original income might vary depending on whether one is running an analysis using PDI or PGT, so Market Income or Market Income plus Pensions variables have to be used for this option depending on the scenario.
FISCAL POLICY AND ETHNO-RACIAL INEQUALITY

ceqrace [pw = weight] using CEQ_Ethno_Racial_MWB.xlsx, race1(indig) race2(white) race3(afrd) race4(orace) race5(nonrace) table(f7) m(y_m) mplusp(y_mp) n(y_nm) g(y_g) taxab(y_taxab) d(y_d) c(y_c) ppp(7.65) cpibase(78.661) cpisurvey(105.196) year next(137) nmod(350)

Where CEQ_Ethno_Racial_MWB.xlsx is the name of the Excel file, the race variables are dummies, m(varname) specifies the Market Income variable in local currency units, mplusp(varname) is Market Income plus Pensions, n(varname) is Net Market Income, g(varname) is Gross Income, taxab(varname) is Taxable Income, d(varname) is Disposable Income, and c(varname) is Consumable Income. ppp( ) is the PPP conversion factor (LCU per international dollar, consumption-based) from the year of PPP (usually either 2005 or 2011), cpibase( ) is CPI of the base year (year of PPP, usually 2005 or 2011), cpisurvey( ) is the CPI for the year of the household survey, and finally, year indicates that the income variables are annual. next( ) and nmod( ) set the national extreme and moderate poverty lines which should be in LCU and the same periodicity as the income variables.

2.4 Sheet F8. Summary Poverty Gap Rates

This sheet mirrors the results presented on sheet F7, but utilizing poverty gap rates rather than the poverty headcount. In addition to tables and figures presenting the poverty gap results, this sheet also automatically calculates the budget that would be required to completely eliminate poverty assuming that programs were perfectly targeted at each of the core income concepts.

To fill in this sheet using the ceqrace.ado Stata command, race dummy variables, weights, market(varname), mpluspensions(varname), netmarket(varname), gross(varname), taxable(varname), disposable(varname), consumable(varname), final(varname), poverty line options, nextreme(string), and nmoderate(string) are required, and the following syntax should be used:

ceqrace [pw = weight] using CEQ_Ethno_Racial_MWB.xlsx, race1(indig) race2(white) race3(afrd) race4(orace) race5(nonrace) table(f8) m(y_m) mplusp(y_mp) n(y_nm) g(y_g) taxab(y_taxab) d(y_d) c(y_c) f(y_f) ppp(7.65) cpibase(78.661) cpisurvey(105.196) year next(137) nmod(350)

Where CEQ_Ethno_Racial_MWB.xlsx is the name of the Excel file, the race variables are dummies, m(varname) specifies the Market Income variable in local currency units, mplusp(varname) is Market Income plus Pensions, n(varname) is Net Market Income, g(varname) is Gross Income, taxab(varname) is Taxable Income, d(varname) is Disposable Income, c(varname) is Consumable Income, and f(varname) is Final Income.
ppp( ) is the PPP conversion factor (LCU per international dollar, consumption-based) from the year of PPP (usually either 2005 or 2011), \( \text{cpibase}( ) \) is CPI of the base year (year of PPP, usually 2005 or 2011), \( \text{cpisurvey}( ) \) is the CPI for the year of the household survey, and finally, \( \text{year} \) indicates that the income variables are annual. \( \text{next}( ) \) and \( \text{nmod}( ) \) are the national extreme and moderate poverty lines which should be in LCU and the same periodicity as the income variables.

2.5 Sheet F9. Summary Poverty Gap Squared Rates

Sheet F9 completes the Foster, Greer, and Thorbecke (1984) family of poverty measures by presenting results on poverty severity (poverty gap squared) across ethno-racial lines for each of the core income concepts. Like the previous two sheets, sheet F9 presents the results alongside easy-to-use figures for regional and national extreme and moderate poverty lines.

To fill in this sheet using the ceqrace.ado Stata command, race dummy variables, weights, \( \text{market}(\text{name}) \), \( \text{mpluspensions}(\text{name}) \), \( \text{netmarket}(\text{name}) \), \( \text{gross}(\text{name}) \), \( \text{taxable}(\text{name}) \), \( \text{disposable}(\text{name}) \), \( \text{consumable}(\text{name}) \), \( \text{final}(\text{name}) \), \( \text{nextreme}(\text{string}) \), and \( \text{nmoderate}(\text{string}) \) are required and the following syntax should be used:

\[
\text{ceqrace \{pw = \text{weight}\} using CEQ\_Ethno\_Racial\_MWB.xlsx, racel(indig) race2(white) race3(afrd) race4(orace) race5(nonrace) table(f9) m(y\_m) mplusp(y\_mp) n(y\_nm) g(y\_g) taxab(y\_taxab) d(y\_d) c(y\_c) f(y\_f) ppp(7.65) cpibase(78.661) cpisurvey(105.196) year next(137) nmod(350)}
\]

Where CEQ_Ethno_Racial_MWB.xlsx is the name of the Excel file, the race variables are dummies, \( m(\text{name}) \) specifies the Market Income variable in local currency units, \( \text{mplusp}(\text{name}) \) is Market Income plus Pensions, \( n(\text{name}) \) is Net Market Income, \( g(\text{name}) \) is Gross Income, \( \text{taxab}(\text{name}) \) is Taxable Income, \( \text{d}(\text{name}) \) is Disposable Income, \( \text{c}(\text{name}) \) is Consumable Income, and \( f(\text{name}) \) is Final Income. \( \text{ppp}( ) \) is the PPP conversion factor (LCU per international dollar, consumption-based) from the year of PPP (usually either 2005 or 2011), \( \text{cpibase}( ) \) is CPI of the base year (year of PPP, usually 2005 or 2011), \( \text{cpisurvey}( ) \) is the CPI for the year of the household survey, and finally, \( \text{year} \) indicates that the income variables are annual. \( \text{next}( ) \) and \( \text{nmod}( ) \) are the national extreme and moderate poverty lines which should be in LCU and the same periodicity as the income variables.

2.6 Sheet F10. Summary Inequality Indicators

Many different measures are used to calculate income inequality in a given society. This sheet features three of these measures: the Gini coefficient, the Theil coefficient,
and the 90/10 index. While the national results presented on this page may be more important than those disaggregated by ethno-racial group since they capture inter- and intra-group inequality rather than just intra-group inequality, it is important to examine these results both at the national level and disaggregated as some policies may decrease inequality nationally while exacerbating inequalities within particular ethno-racial groups. Similarly, programs may increase inequality nationally while decreasing intra-group inequalities. Like sheets F7, F8, and F9, sheet F10 presents results with easy-to-use tables and figures.

To fill in this sheet using the `ceqrace.ado` Stata command, race dummy variables, weights, `market(varname)`, `mpluspensions(varname)`, `netmarket(varname)`, `gross(varname)`, `taxable(varname)`, `disposable(varname)`, `consumable(varname)`, and `final(varname)` are required, and the following syntax should be used:

```
ceqrace [pw = weight] using CEQ_Ethno_Racial_MWB.xlsx, race1(indig) race2(white) race3(afrd) race4(orace) race5(nonrace) table(f10) m(y_m) mplusp(y_mp) n(y_nm) g(y_g) taxab(y_taxab) d(y_d) c(y_c) f(y_f)
```

Where `CEQ_Ethno_Racial_MWB.xlsx` is the name of the Excel file, the race variables are dummies, `m(varname)` specifies the Market Income variable in local currency units, `mplusp(varname)` is Market Income plus Pensions, `n(varname)` is Net Market Income, `g(varname)` is Gross Income, `taxab(varname)` is Taxable Income, `d(varname)` is Disposable Income, `c(varname)` is Consumable Income, and `f(varname)` is Final Income.

### 2.7 Sheet F11. Mean Incomes

In examining inequalities across ethno-racial lines, it is also important to consider gaps in mean incomes held by individuals of different ethno-racial groups. Sheet F11 presents the mean incomes experienced by each ethno-racial population at each of the different income concepts. Results are presented both in 2005 PPP dollars as well as in local currency units. As with the preceding sheets, results are presented as easy-to-use figures and tables.

To fill in this sheet using the `ceqrace.ado` Stata command, race dummy variables, weights, `market(varname)`, `mpluspensions(varname)`, `netmarket(varname)`, `gross(varname)`, `taxable(varname)`, `disposable(varname)`, `consumable(varname)`, and `final(varname)` are required, and the following syntax should be used:

```
ceqrace [pw = weight] using CEQ_Ethno_Racial_MWB.xlsx, race1(indig) race2(white) race3(afrd) table(f11) m(y_m) mplusp(y_mp) n(y_nm) g(y_g) taxab(y_taxab) d(y_d) c(y_c) f(y_f)
```
2.8 Sheet F12. Incidence by Decile

When conducting the CEQ fiscal incidence analysis, one of the most important elements is determining the incidence of different fiscal interventions on household income. When analyzing the effects of fiscal policy across ethno-racial lines the same holds true. Sheet F12 presents the magnitude of interventions in each decile, disaggregated by ethno-racial group as well as nationally, measured in local currency units. Results are also presented as a share of original income for each population. While the results shown on this sheet should be the same as those on sheet D4 of the CEQ Master Workbook and in section E for the national level, when disaggregated by ethno-racial group, results will be different from those shown for particular groups’ respective sections D and E of the CEQ Master Workbook as deciles are defined differently.

To fill in this sheet using the ceqrace.ado Stata command, race dummy variables, weights, hhid(varname), original(varname), market(varname), mpluspensions(varname), netmarket(varname), gross(varname), taxable(varname), disposable(varname), consumable(varname), final(varname), dtax(varname), contributions(varname), contpensions(varname), contypensions(varname), noncontributory(varname), flagcct(varname), otransfers(varname), isubsidies(varname), itax(varname), ikeducation(varname), ikhealth(varname), and hurban(varname) are required, and the following syntax should be used:

```stata
ceqrace [pw = weight] using CEQ_Ethno_Racial_MWB.xlsx, race1(indig) race2(white) race3(afrd) race4(orace) race5(nonrace) table(f12) hhid(hhid) o(y_m) m(y_m) contp(contp) conyp(conyp) mplusp(y_mp) dtax(dtax) n(y_nm) nonc(nonc) flagcct(fcct) otran(otran) g(y_g) taxab(y_taxab) d(y_d) isub(isub) itax(itax) c(y_c) ike(ik_e) ikh(ik_h) hu(hu) f(y_f)
```

Where CEQ_Ethno_Racial_MWB.xlsx is the name of the Excel file, the race variables are dummies, hhid() is the variable that uniquely identifies the household, o(varname) specifies the original income variable in local currency units, m(varname) is Market Income, contp(varname) are contributions to pensions, conyp(varname) are contributory pensions, mplusp(varname) is Market Income plus Pensions, dtax(varname) are direct taxes, n(varname) is Net Market Income, nonc(varname) are noncontributory pensions, flagcct(varname) is the Flagship Conditional Cash Transfer Program (CCT), otran(varname) are other direct transfers, g(varname) is Gross Income, taxab(varname) is Taxable Income, d(varname) is Disposable Income, isub(varname) and itax(varname) are indirect subsidies and taxes, respectively, c(varname) is Consumable Income, ike(varname), ikh(varname), and hu(varname) are in-kind education, in-kind health, and in-kind housing and urban benefits respectively, and f(varname) is Final Income.
2.9 Sheet F13. Incidence by Income Group

While Sheet F12 presents the incidence results of the analysis by decile, sheet F13 complements this by conducting the same analysis by income group. This allows researchers to utilize populations that have the same income or to examine the impact of policies on particular income groups within the different ethno-racial groups. These results will be the same as those presented on sheet D4 of the CEQ Master Workbook for each respective ethno-racial group.

To fill in this sheet using the `ceqrace.ado` Stata command, race dummy variables, weights, original(varname), market(varname), mpluspensions(varname), gross(varname), taxable(varname), disposable(varname), consumable(varname), final(varname), dtax(varname), contributions(varname), contpensions(varname), contypensions(varname), noncontributory(varname), flagcct(varname), otransfers(varname), isubsidies(varname), itax(varname), ikeducation(varname), ikhealth(varname), hurban(varname), and poverty line options are required, and the following syntax should be used:

```
ceqrace [pw = weight] using CEQ_Ethno_Racial_MWB.xlsx, race1(indig) race2(white) race3(afrd) race4(orace) race5(nonrace) table(f13) o(y) m(y) contp(contp) contyp(contyp) mplusp(y_mp) dtax(dtax) n(y_nm) nonc(nonc) flagcct(fcct) otran(otran) g(y_g) taxab(y_taxab) d(y_d) isub(isub) itax(itax) c(y_c) ike(ik_e) ikh(ik_h) hu(hu) f(y_f) ppp(7.65) cpibase(78.661) cpisurvey(105.196) year
```

Where CEQ_Ethno_Racial_MWB.xlsx is the name of the Excel file, the race variables are dummies, o(varname) specifies the original income variable in local currency units, m(varname) is Market Income, contp(varname) are contributions to pensions, contyp(varname) are contributory pensions, mplusp(varname) is Market Income plus Pensions, dtax(varname) are direct taxes, n(varname) is Net Market Income, nonc(varname) are noncontributory pensions, flagcct(varname) is the CCT, otran(varname) are other direct transfers, g(varname) is Gross Income, taxab(varname) is Taxable Income, d(varname) is Disposable Income, isub(varname) and itax(varname) are indirect subsidies and taxes respectively, c(varname) is Consumable Income, ike(varname), ikh(varname), and hu(varname) are in-kind education, in-kind health, and housing and urban benefits, respectively, and f(varname) is Final Income.

2.10 Sheet F14. Cross-Race Incidence

While sheets F12 and F13 present the results of the fiscal incidence analysis across ethno-racial lines, these results may be difficult to read. Sheet F14 utilizes the analysis
presented on Sheet F12 to show the findings of the incidence analysis by ethno-racial group in an easy-to-read table. The results reveal the share of benefits (or payments) received (paid out) by each ethno-racial group as a share of total benefits (or payments). When this is compared to the population (row 8) or income (rows 9, 15, and 18 for Market, Disposable, and Consumable, respectively) shares, the progressivity of different policy interventions can be examined. As discussed above, policies are considered to be regressive when the share of benefits (taxes) being received (paid) by the disadvantaged population is less (more) than its share of national income, relatively progressive when the share of benefits (taxes) being received (paid) by the disadvantaged population is more (less) than its share of national income, and absolutely progressive when the share of benefits being received by the disadvantaged population is more than its share of the population.

This sheet is filled in automatically using the results calculated from sheet F12.

2.11 Sheet F15. Horizontal Equity

The impact of the fiscal policies targeted to the poor may appear to be ethno-racially progressive due to greater poverty rates among the disadvantaged population(s). This can lead to questions about whether the program benefits the poor of a particular group more or less than other segments of the population. Sheet F15 examines the incidence of different policy interventions among the poor of each ethno-racial group relative to its population and income shares. This allows us to examine whether policies are disproportionately benefiting the poor of a particular ethno-racial group or whether certain policies appear to be ethno-racially progressive or regressive due to differences in the socioeconomic status of the different populations. If the share of benefits going to a particular population is equal to its share of the poor, policies are considered to be colorblind; that is, they do not violate horizontal equity by benefiting the poor of particular populations more than others.

This sheet is filled in automatically using the results calculated on sheet F13.

2.12 Sheet F16. Fiscal Profile

In addition to looking at the share of benefits going to each ethno-racial population, it is important to see the impact on incomes within each of these populations. Looking at the fiscal profile sheet allows us to see these changes in mean income, in terms of local currency units and as a share of the different income concepts. In addition to looking at the impacts on mean income among individuals of each race or ethnicity, this sheet looks at the differences that occur in households headed by members of different races or ethnicities. This allows us to see if there are differences between interracial households and single-race households.

To fill in this sheet using the cegrace.ado Stata command, race dummy variables, weights, original(varname), disposable(varname), consumable(varname),
FISCAL POLICY AND ETHNO-RACIAL INEQUALITY

(varname), age(varname), pensions(varname), hhe(varname), hhid(varname), and poverty line options are required, and the following syntax should be used:

c eqrace [pw = weight] using CEQ_Ethno_Racial_MWB.xlsx, race1(indig) race2(white) race3(afrd) race4(orace) race5(nonrace) table(f16) o(y_m) d(y_d) c(y_c)f(y_f) pens(pensions) hhe(hhe_id) hhid(hh_id) ppp(7.65) cpibase(78.661) cpisurvey(105.196) year

Where CEQ_Ethno_Racial_MWB.xlsx is the name of the Excel file, the race variables are dummies, and o(varname) specifies the original income variable in local currency units. Original income is used in order to assert whether the analysis that is being run uses pensions as deferred income (PDI) or pensions as government transfers (PGT). d(varname) is Disposable Income, c(varname) is Consumable Income, f(varname) is Final Income, pens(varname) are pensions, hhe(varname) is a dummy variable that identifies the household head, hhid(varname) is the household identifier, ppp( ) is the PPP conversion factor (LCU per international dollar, consumption-based) from the year of PPP (usually either 2005 or 2011), cpibase( ) is CPI of the base year (year of PPP, usually 2005 or 2011), cpisurvey( ) is the CPI for the year of the household survey, and finally, year indicates that the income variables are annual.

2.13 Sheet F17. Coverage Rates (Total Population)

In addition to looking at the impacts of fiscal policy between ethno-racial groups on the aggregate, it is important to look at what share of each ethno-racial population is receiving benefits from the different fiscal interventions. The coverage rates of the different populations allow researchers to have a better understanding of the targeting of programs, in addition to seeing their impact on incomes and poverty. Sheet F17 looks at the coverage rates of the total population, regardless of whether all individuals making up the population are the desired targets of particular fiscal interventions. These results are disaggregated by ethno-racial group as well as by income group.

There are multiple ways that one can calculate coverage rates. For the purpose of the CEQ analysis, coverage rates of direct beneficiaries, indirect beneficiaries, and households may all be interesting and can be calculated for each of these distinct populations. In order to understand the differences between the different coverage rates, it is necessary to understand what populations are being considered as part of each group.

1. Direct beneficiaries are those who report being recipients of a particular intervention. In cases where benefits are directed at the household, direct beneficiaries will be imputed to the head of the household or to all members of the household depending on the targeting method being utilized. In some cases, households may have more than one direct beneficiary.
2. **Beneficiary households** are households in which at least one direct beneficiary resides.

3. **Direct and indirect beneficiaries** are all individuals who reside within a beneficiary household.

To calculate the coverage rates using these different methods requires dividing the number of beneficiaries by the total population in the case of direct and indirect beneficiaries and by the total number of households in the case of beneficiary households. For additional information on how to calculate each of the different coverage rates, please refer to chapter 8 of this Handbook.\(^{13}\)

To fill in this sheet using the `ceqrace.ado` Stata command, race dummy variables, weights, original\((\text{varname})\), cct\((\text{varname})\), noncontrib\((\text{varname})\), unemploy\((\text{varname})\), foodtransf\((\text{varname})\), otransfers\((\text{varname})\), health\((\text{varname})\), pensions\((\text{varname})\), hhe\((\text{varname})\), hhid\((\text{varname})\), and poverty lines are required, and the following syntax should be used:

```
ceqrace [pw = weight] using CEQ_Ethno_Racial_MWB.xlsx, rac1(indig) race2(white) race3(afrd) race4(orace) race5(nonrace) table(f17) o(y_m) cct(cct) nonc(nonc) unem(unemployment) foodt(f_tran) otran(o_tran) hea(health) pen(pensions) hhe(hhe_id) hhid(hh_id) ppp(7.65) cpibase(78.661) cpisurvey(105.196) year
```

Where `CEQ_Ethno_Racial_MWB.xlsx` is the name of the Excel file, the race variables are dummies, \(o(\text{varname})\) specifies the original income variable in local currency units, \(cct(\text{varname})\) are conditional cash transfers, \(nonc(\text{varname})\) are noncontributory pensions, \(unem(\text{varname})\) are unemployment benefits, \(foodt(\text{varname})\) are food transfers, \(otran(\text{varname})\) are other direct transfers, \(hea(\text{varname})\) are health transfers, \(pen(\text{varname})\) are pensions, \(hhe(\text{varname})\) is a dummy variable that identifies the household head, \(hhid(\text{varname})\) is the household identifier, \(ppp(\ )\) is the PPP conversion factor (LCU per international dollar, consumption-based) from the year of PPP (usually either 2005 or 2011), \(cpibase(\ )\) is the CPI of the base year (year of PPP, usually 2005 or 2011), \(cpisurvey(\ )\) is the CPI for the year of the household survey, and finally, \(year\) indicates that the income variables are annual.

### 2.14 Sheet F18. Coverage Rates (Target Population)

Building upon the results of sheet F17, this sheet examines the coverage rates among the population that is the desired target of specific interventions. The target population is likely to differ by intervention. For example, pensions may be targeted to individuals over a particular age, while some social cash transfers may be targeted to

---

\(^{13}\)Higgins (2018).
heads of households with children within a particular age range. These targeted coverage rates are calculated using the same three population definitions given above. However, the coverage rates presented on this sheet do not include recipients who are not part of the desired population. In the case of households, the denominator includes all households where at least one individual with the desired characteristics resides, while for direct and indirect beneficiaries, the denominator includes all individuals who reside in a household where at least one individual with the desired characteristics resides.

To fill in this sheet using the ceqrace.ado Stata command, race dummy variables, weights, original(varname), CCT(varname), noncontrib(varname), pensions(varname), hhe(varname), hhid(varname), tarcct(varname), tarcnp(varname), tarpen(varname), and poverty line options are required and the following syntax should be used:

```stata
ceqrace [pw = weight] using CEQ_Ethno_Racial_MWB.xlsx, race1(indig) race2(white) race3(afrd) race4(orace) race5(nonrace) table(f18) o(y_m) cct(cct) nonc(nonc) pen(pensions) hhe(hhe_id) hhid(hh_id) tarcnp(tncp) tarcct(tcct) tarpen(tpen) ppp(7.65) cpibase(78.661) cpisurvey(105.196) year
```

Where CEQ_Ethno_Racial_MWB.xlsx is the name of the Excel file, the race variables are dummies, o(varname) specifies the original income variable in local currency units, cct(varname) are conditional cash transfers, nonc(varname) are noncontributory pensions, pen(varname) are pensions, hhe(varname) is a dummy variable that identifies the household head, hhid(varname) is the household identifier, tarcnp(varname) is a dummy variable that identifies noncontributory pensions target population, tarcct(varname) is a dummy variable that identifies CCT’s target population, tarpen(varname) is a dummy variable that identifies pensions target population. ppp( ) is the PPP conversion factor (LCU per international dollar, consumption-based) from the year of PPP (usually either 2005 or 2011), cpibase( ) is the CPI of the base year (year of PPP, usually 2005 or 2011), cpisurvey( ) is the CPI for the year of the household survey, and finally, year indicates that the income variables are annual.

### 2.15 Sheet F19. Leakages

Programs are often likely to direct some benefits to a segment of the population that does not meet the desired targeting characteristics. Using the results of the two different coverage sheets (F17 and F18), this sheet seeks to explain if the leakages from these programs benefit a particular ethno-racial group more than another. These are calculated by taking the total size of benefits and subtracting the amount of benefits that are received by the target population. Results are calculated in both 2005 PPP values.
and local currency, as well as in terms of a percentage of total spending on a particular intervention.

This sheet is filled in automatically using the results presented on sheets F17 and F18.

### 2.16 Sheet F20. Mobility Matrices

In order to determine if a program is “pro-disadvantaged group,” it is necessary to determine if the impact of fiscal policies leads to a higher probability of escaping poverty for the disadvantaged population than for the advantaged population. To calculate this, this workbook utilizes the mobility matrices discussed in Lustig and Higgins (2013). These matrices look at the population that is in or out of poverty at two different income concepts. This sheet presents these mobility matrices for each of the different ethno-racial populations and calculates the probability of an individual living in poverty at Market Income escaping poverty through fiscal interventions. Probabilities of escaping poverty are calculated from Consumable, Disposable, and Final Income, all with respect to Market Income for each of the different ethno-racial populations using the regional poverty lines of $2.50 PPP per capita per day and $4 PPP per capita per day. The results represented in the mobility matrices on this sheet should match those found on sheet D10 for each of the ethno-racial groups.

To fill in this sheet using the `ceqrace.ado` Stata command, race dummy variables, weights, original(`varname`), disposable(`varname`), consumable(`varname`), final(`varname`), and poverty lines are required, and the following syntax should be used:

```
ceqrace [pw = weight] using CEQ_Ethno_Racial_MWB.xlsx, racel(indig) race2(white) race3(afrd) race4(orace) race5(nonrace) table(f20) o(y_m) d(y_d) c(y_c) f(y_f) ppp(7.65) cpibase(78.661) cpisurvey(105.196) year
```

Where `CEQ_Ethno_Racial_MWB.xlsx` is the name of the Excel file, the race variables are dummies, `o(`varname`)` specifies the original income variable in local currency units, `mplusp(`varname`)` is Market Income plus Pensions, `n(`varname`)` is Net Market Income, `g(`varname`)` is Gross Income, `taxab(`varname`)` is Taxable Income, `d(`varname`)` is Disposable Income, `c(`varname`)` is Consumable Income, and `f(`varname`)` is Final Income, `ppp()` is the PPP conversion factor (LCU per international dollar, consumption-based) from the year of PPP (usually either 2005 or 2011), `cpibase()` is the CPI of the base year (year of PPP, usually 2005 or 2011), `cpisurvey()` is the CPI for the year of the household survey, and finally, `year` indicates that the income variables are annual.
2.17 Sheet F21. Education (Totals)

One area that is commonly cited as a source of ethno-racial inequality is educational outcomes. This is also where government provision of services is an important tool in closing ethno-racial inequalities. Thus, looking at the differences in educational attainment and enrollment can be crucial to explaining ethno-racial inequalities. This sheet looks at the size of different ethno-racial populations that are attending public and private educational institutions in order to see the impacts of government services at closing inequalities in access to education.

To fill in this sheet using the ceqrace.ado Stata command, race dummy variables, weights, original(varname), edpre(varname), edpri(varname), edsec(varname), edter(varname), redpre(varname), redpri(varname), redsec(varname), redter(varname), edpublic(varname), edprivate(varname), and attend(varname) are required, and the following syntax should be used:

```
ceqrace [pw = weight] using CEQ_Ethno_Racial_MWB.xlsx, race1(indig) race2(white) race3(afrd) race4(orace) race5(nonrace) table(f21) oy(m) edpre(ed_pre) edpri(ed_pri) edsec(ed_sec) edter(ed_ter) attend(attendschool) redpre(red_pre) redpri(red_pri) redsec(red_sec) redter(red_ter) hhe(id_hhead) hhid(id_hhid) edpriv(private) edpub(public)
```

Where CEQ_Ethno_Racial_MWB.xlsx is the name of the Excel file, the race variables are dummies, o(varname) specifies the original income variable in local currency units; edpre(varname), edpri(varname), edsec(varname), edter(varname) are preschool, primary, secondary, and tertiary level of education dummies respectively; redpre(varname), redpri(varname), redsec(varname), redter(varname) are preschool, primary, secondary, and tertiary age ranges dummies, respectively; attend(varname) is a dummy that defines whether the individual attends school; hhe(varname) is a dummy variable that identifies the household head, hhid(varname) is the household identifier; edpriv(varname) and edpub(varname) are dummies that identify whether the individual attends a private or public school; ppp( ) is the PPP conversion factor (LCU per international dollar, consumption-based) from the year of PPP (usually either 2005 or 2011); cpibase( ) is the CPI of the base year (year of PPP, usually 2005 or 2011); cpiSurvey( ) is the CPI for the year of the household survey; and finally, year indicates that the income variables are annual.

---

14If an individual is between the age range of each education level, then the variable takes the value of one.
2.18 Sheet F22. Education (Rates)

Utilizing the population numbers presented on sheet F21, this sheet calculates different education rates and presents them in easy-to-use tables. Both gross and net enrollment rates are calculated for each ethno-racial population at each level of education that is available for analysis, ranging from preschool through tertiary education. These rates are further disaggregated by income group at Disposable Income.

This sheet is filled in automatically using the results presented on sheet F21.

2.19 Sheet F23. Infrastructure Access

Another element of ethno-racial inequality comes from the nonmonetary deprivations that may be experienced by different populations. These may include access to key services that are often considered to be connected to economic performance, such as access to potable water or electricity. This sheet presents the different coverage rates experienced by individuals of different ethno-racial populations for access to running water, electricity, sewage, and roads as well as to well-constructed walls, floors, and roofs. These results are calculated using two different methods, one that examines the coverage rate of the population (weighted households) and one that looks at the coverage rate of households. In addition to showing the coverage rates, this sheet presents the distribution of beneficiaries, both by household and population.

To fill in this sheet using the cegrace.ado Stata command, race dummy variables, weights, original(varname), hhid(varname), hhead(varname), water(varname), electricity(varname), walls(varname), floors(varname), roof(varname), sewage(varname), roads(varname), and poverty lines are required. If one of the infrastructure variables is not included in the dataset, the ado-file will leave those observations blank.

Example:

cegrace [pw = weight] using CEQ_Ethno_Racial_MWB.xlsx, race1(indig) race2(white) race3(afrd) race4(orace) race5(nonrace) table(f23) o(y_m) hhid(id_hh) hhe(id_hhead) water(water) electricity(elect) walls(walls) floors(floors) roof(roof) sewage(sewage) roads(roads) ppp(7.65) cpibase(78.661) cpisurvey(105.196) year

Where CEQ_Ethno_Racial_MWB.xlsx is the name of the Excel file, the race variables are dummies, o(varname) specifies the original income variable in local currency units, hhe(varname) is a dummy variable that identifies the household head, hhid(varname) is the household identifier, water(varname), electricity(varname),
walls(varname), floors(varname), roof(varname), sewage(varname), and roads(varname) are all dummies for having running water, electricity, walls, floors, roof, sewage, and roads, respectively. ppp( ) is the PPP conversion factor (LCU per international dollar, consumption-based) from the year of PPP (usually either 2005 or 2011), cpibase( ) is the CPI of the base year (year of PPP, usually 2005 or 2011), cpisurvey( ) is the CPI for the year of the household survey, and finally, year indicates that the income variables are annual.

2.20 Sheet F24. Theil Decomposition

As discussed above, one of the ways that one can determine the effect of fiscal policy and the magnitude of ethno-racial inequality is to use a decomposable inequality indicator to determine what share of inequality is due to differences in income between income groups. This sheet does just that and uses the decomposable Thiel coefficient to determine what share of inequality is due to differences in incomes between groups and what share of inequality is due to intra-group inequalities. These results are calculated for each of the eight core income concepts. These can be compared to see if the share of inequality due to ethno-racial differences declines as a result of fiscal interventions.

To fill in this sheet using the ceqrace.ado Stata command, race dummy variables, weights, market(varname), mpluspensions(varname), netmarket(varname), gross(varname), taxable(varname), disposable(varname), consumable(varname), final(varname), gender(varname), urban(varname), and edpar(varname) are required.

Example:

ceqrace [pw = weight] using CEQ_Ethno_Racial_MWB.xlsx, race1(indig) race2(white) race3(afrd) race4(orace) race5(nonrace) table(f24) m(y_m) mplusp(y_mp) n(y_nm) g(y_g) taxab(y_taxab) d(y_d) c(y_c) f(y_f) gender(sex) urban(rururb) edpar (parentsed)

Where CEQ_Ethno_Racial_MWB.xlsx is the name of the Excel file, the race variables are dummies, m(varname) specifies the Market Income variable in local currency units, mplusp(varname) is Market Income plus Pensions, n(varname) is Net Market Income, g(varname) is Gross Income, taxab(varname) is Taxable Income, d(varname) is Disposable Income, c(varname) is Consumable Income, and f(varname) is Final Income; gender(varname) is a dummy variable specifying the gender of the individual (1 for women and 0 otherwise), urban(varname) is also a dummy specifying whether the individual lives in an urban or a rural area, and edpar(varname) specifies the years of education of the head of the household.
2.21 Sheet F25. Inequality of Opportunity

As discussed above, one of the ways in which one can measure ethno-racial inequality is through inequality of opportunity. This measure seeks to explain if differences in outcomes are due to characteristics or circumstances outside of an individual’s control rather than being due to personal preferences or effort. In this case, the characteristics that are considered part of the analysis are the individuals’ gender, location (urban/rural), and race or ethnicity. The CEQ race and ethnicity analysis looks specifically at how these characteristics affect inequality of income at each of the different income concepts. By considering the mean incomes of the different combinations of individuals with these characteristics, one can calculate to what extent each of the different characteristics describes the differences in mean incomes. By considering the change in the share of inequality of opportunity explained by race and ethnicity, one can determine if fiscal policy reduces the share of inequality of opportunity explained by race and ethnicity.

To fill in this sheet using the ceqrace.ado Stata command, race dummy variables, weights, market(varname), mpluspensions(varname), netmarket(varname), gross(varname), taxable(varname), disposable(varname), consumable(varname), final(varname), gender(varname), urban(varname) are required.

Example:

```
ceqrace [pw = weight] using CEQ_Ethno_Racial_MWB.xlsx, race1(indig) race2(white) race3(afrd) race4(orace) race5(nonrace) table(f25) m(y_m) mplusp(y_mp) n(y_nm) g(y_g) taxab(y_taxab) d(y_d) c(y_c) f(y_f) gender(sex) urban(rururb)
```

Where CEQ_Ethno_Racial_MWB.xlsx is the name of the Excel file, the race variables are dummies, m(varname) specifies the Market Income variable in local currency units, mplusp(varname) is Market Income plus Pensions, n(varname) is Net Market Income, g(varname) is Gross Income, taxab(varname) is Taxable Income, d(varname) is Disposable Income, c(varname) is Consumable Income, and f(varname) is Final Income, gender(varname) is a dummy variable specifying the gender of the individual (1 for women and 0 otherwise), and urban(varname) is a dummy specifying if the individual lives in an urban or a rural area.

2.22 Sheet F26. Significance

In order to determine whether there are in fact differences in the incomes of different ethno-racial populations, it is necessary to verify that these values are statistically significant. To do this, one can calculate p-values comparing the different indicators across ethno-racial lines. This sheet looks at the p-values for the poverty headcounts.
FISCAL POLICY AND ETHNO-RACIAL INEQUALITY

(US$2.50 and US$4 PPP/day), Gini coefficient, and Theil coefficient between each pairwise set of ethno-racial groups to determine if the differences between each ethno-racial group are statistically significant for each indicator. These are calculated for each of the eight core income concepts.

To fill in this sheet using the ceqrace.ado Stata command, race dummy variables, weights, market(varname), mpluspensions(varname), netmarket(varname), gross(varname), taxable(varname), disposable(varname), consumable(varname), final(varname), psu(varname), strata(varname), and poverty line options are required.

Example:

ceqrace [pw = weight] using CEQ_Ethno_Racial_MWB.xlsx, race1(indig) race2(white) race3(afrd) race4(oracle) race5(nonrace) table(f26) m(y_m) mplusp(y_mp) n(y_nm) g(y_g) taxab(y_taxab) d(y_d) c(y_c) f(y_f) psu(upm) strata(strata) ppp(7.65) cpibase(78.66) cpisurvey(105.196) year

Where CEQ_Ethno_Racial_MWB.xlsx is the name of the Excel file, the race variables are dummies, m(varname) specifies the Market Income variable in local currency units, mplusp(varname) is Market Income plus Pensions, n(varname) is Net Market Income, g(varname) is Gross Income, taxab(varname) is Taxable Income, d(varname) is Disposable Income, c(varname) is Consumable Income, and f(varname) is Final Income; gender(varname) is a dummy variable specifying the gender of the individual (1 for women and 0 otherwise), urban(varname) is also a dummy specifying if the individual lives in an urban or a rural area, psu(varname) is the primary sampling unit, strata(varname) is the strata variable, ppp( ) is the PPP conversion factor (LCU per international dollar, consumption-based) from the year of PPP (usually either 2005 or 2011), cpibase( ) is the CPI of the base year (year of PPP, usually 2005 or 2011), cpisurvey( ) is the CPI for the year of the household survey, and finally, year indicates that the income variables are annual.

References

Bank, Latin America and Caribbean Region Office of the Chief Economist, and Human Development Network Social Protection and Labor Unit).


