CEQ INSTITUTE
COMMITMENT TO EQUITY
Tulane University

A guide to naming Harmonized Microdata variables

Maynor Cabrera, Samantha Greenspun and Sandra Martínez

Version July 3, 2017

I. Introduction

A key goal of the CEQ Institute is to establish the CEQ Data Center, a rich publicly available information system. The Data Center will consist of: the CEQ Standard Indicators Database, the CEQ Master Workbooks collection (MWB), and the Harmonized Microdata Bank—the core of the information system. The entire system will enable researchers, international organizations, and policymakers, among others, to monitor progress in fiscal redistributive efforts to achieve equity goals. Specifically, the Harmonized Microdata Bank will enable researchers to conduct cross-country analyses using each country´s harmonized household survey data, as opposed to, for example, using deciles or income categories, which will boost the richness and statistical power of the analysis.

The harmonized household survey data is a country´s household survey at the individual level. In order to create the harmonized household survey data, CEQ Teams will do the following. Using the final dataset that resulted from the completed CEQ Assessment, each CEQ Team will follow the methodology in Lustig (2017) as well as the instructions in this guide to walk them through transforming the original CEQ Assessment dataset into the harmonized household survey dataset.

Given the importance of the Harmonized Microdata for the establishment of the CEQ Data Center, ensuring the quality and comparability of datasets is crucial. Therefore, the objective of this document (which is to be used with the Lustig (2017) methodology) is to provide guidelines on how to structure the Harmonized Microdata. The next section, Section II, provides the household, household members, CEQ Income concepts, fiscal interventions, and target and beneficiary population, among other variables that must be included in the Harmonized Microdata. It also explains how to name, label and organize these variables. The final section, Section, III provides an applied example as well as useful Stata code.

The CEQ Team appreciates your contribution to this project.

Please do not hesitate to contact Maynor Cabrera at maynor.cabrera@ceqinstitute.org and Sandra Martinez at sandra.martinez@ceqinstitute.org if you have any questions or comments.

II. Constructing the Harmonized Microdata

Some important notes before starting to prepare your data set:

1. The harmonized microdata needs to be:
    a. at **the individual level** (where each observation is an individual rather than a household). This is necessary because it is not feasible to estimating coverage, target population, education, and infrastructure indicators when using the dataset is at the household level.
    b. in **per capita** terms. However, you might also want to add variables in per adult equivalent terms if applicable.
    c. in **annual basis**.
2. **All variables must be in one file**. You can use the "notes" feature in Stata if you want to add more information into the dataset about respective variables. To do this, use the Stata command "`notes`"[1].
3. All the variables and the values of categorical variables must be labeled. Labeling variables not only allows other users to understand the content of the variable, e.g. which type of fiscal intervention is included and the definition of the values of categorical variables, but it also facilitates the comprehension of Section E of the MWBs[2].
4. It is important to use "double" format for fiscal interventions and CEQ income definitions to prevent substantial discrepancies in the MWB due to rounding errors[3].

Below you will find instructions on how to create, name, label, and organize variables for the harmonized microdata according to the following variable groups types:

[1] For a description of this command visit http://www.stata.com/help.cgi?notes
[2] Section E of the MWB is automatically filled in by the CEQ Stata Package.
[3] You can do this by putting the command `set type double` at the beginning of each file.

A. Household Variables

The CEQ Country Teams must include household variables in the Harmonized Microdata. The following table shows which variables must be included and how to name and label them.

Table 1: Harmonized Microdata household variables and labels

Variable	Label	Contents	Definition
hhid	Household identifier	Code	Unique identifier per household
hsize	Household size	[1, 2,…N]	Number of household members, which excludes external members of the household like boarders, live-in domestic servants, and (if applicable) their families. These external household members should not be included in any income calculations.
adulteq	Household size in adult equivalent scale	[1,1.08, 1.25…N]	If applicable, include a measure of adult equivalence used in the study. If you do this, you must explain how it was calculated. For examples on how to report scales of adult equivalence see https://www.oecd.org/eco/growth/OECD-Note-EquivalenceScales.pdf
urban	Living in urban area	[0=rural, 1=urban]	Indicator of residency in urban versus rural areas. The classification should follow the country-specific guidelines.
inf_water	Access to piped water	[0=no, 1=yes]	Please consider the following categories defined by the WHO/UNICEF Joint Monitoring Programme: • Piped water into dwelling, also called a household connection, is defined as a water service pipe connected with in-house plumbing to one or more taps (e.g. in the kitchen and bathroom). • Piped water to yard/plot, also called a yard connection, is defined as a piped water connection to a tap placed in the yard or plot outside the house.
inf_elect	Access to electricity	[0=no, 1=yes]	Connection of dwelling to electricity supply network.
inf_roof	Roofing material	[0=no, 1=yes]	This variable measures if the roof was built with quality materials instead of precarious materials (for example, cane, palm leave or zinc). Authors must specify country-specific criteria for quality versus precarious materials.
inf_walls	Quality walls	[0=no, 1=yes]	This variable measures if the wall was built with quality materials instead of precarious materials (for example, waste materials, palm leave or zinc/tin). Authors must specify country-specific criteria for quality versus precarious materials.
inf_sewage	Quality sanitation	[0=no, 1=yes]	Please consider the following definition of categories described by the WHO/UNICEF Joint Monitoring Programme: • Flush toilet uses a cistern or holding tank for flushing water, and a water seal (which is a U-shaped pipe below the seat or squatting pan) that prevents the passage of flies and odors. • Piped sewer system is a system of sewer pipes, also called sewerage, that is designed to collect human excrement (feces and urine) and wastewater and subsequently removes them from the household. • Septic tank is an excrement collection device consisting of a watertight settling tank, which is normally located underground, away from the house or toilet.
inf_floor	Quality floor	[0=no, 1=yes]	This variable measures if the house has a floor made of quality materials or a floor made of materials that are considered precarious (for example, dirt). Authors must specify country-specific criteria for quality versus precarious materials.
inf_roads	Access to paved roads	[0=no, 1=yes]	This variable measures whether the household has access to paved roads or not.

B. Household Member Variables

The CEQ Country Teams must include household member variables in the Harmonized Microdata. The following table shows which variables must be included and how to name and label them.

Table 2: Harmonized Microdata household members variables' names and labels

Variable	Label	Contents	Definition
memb_no	Correlative number of member	[1..N]	Correlative number of member of household
age	Age of the household member	Age in years	Age in years
gender	Gender [0Male, 1Female]	[0=Male, 1=Female]	Gender of the individual
relation	Relationship to household head	[1=Head, 2=Spouse/partner, 3=Son/daughter (of the head of the household and/or of the partner of the head of the household) 4=Parents Mother/father in law 5=Other relatives 6=Other non relative]	Classification of household members according to their relation to the head
at_school	Currently attending school	[0=No, 1=Yes]	This variable defines whether or not the individual is currently attending school.
level_school	Level of schooling that is attending	Please add the minimum and maximum theoretical age of each level using the format [x-y]. For example: 1=Preschool [4-5] 2= Primary [6-13] etc.. [0=Not attending, 1=Preschool [x-y] 2=Primary [x-y] 3=Lower secondary [x-y] 4=Upper secondary [x-y] 5=Secondary (total) [x-y] 6= Post-secondary [x-y] non-tertiary 7=Tertiary (Bachelor's or equivalent) [x-y] 8=Master, Doctoral or equivalent [x-y] 9=Other (non specified before)	It is important to use the same levels as in the per capita imputations for monetized benefits of education. However, if budgetary information allows you to estimate educational benefits per level, like those specified in the Contents column, you should use ISCED categories if it is possible. See Unesco mappings here http://uis.unesco.org/en/isced-mappings. Please note that you will not necessarily have all the categories. Please leave the categories that do not apply blank. If you have lower and upper secondary, please include them separately and as a total. Post-secondary non-tertiary education provides learning experiences building on secondary education, preparing for labour market entry as well as tertiary education. It aims at the individual acquisition of knowledge, skills and competencies lower than the level of complexity characteristic of tertiary education. If you have vocational education, this might correspond to secondary or post-secondary education, please be sure to categorize it where it corresponds. If you use Category 9, please give the definition of this level. Please do not include adult education in primary and/or secondary education; this should be treated as a separate category.
type_school	Type of school that it is attending (public, private, semi-private)	[1=Public, 2=Private, 3=Semi-private (subsidized by Government), 4=Other	This variable indicates what type of school the individual is currently attending.

C. CEQ Income Concepts

The CEQ Country Teams must include the CEQ Income concepts variables in the Harmonized Microdata. The team must adopt the following structure for naming the CEQ core income concepts variables:

xx_zz

Where:

- [xx] corresponds to the name of the core income concept – see Annex 1 for a Diagram of the CEQ Core Income Concepts[4]:

Letters	Core income concept
[ym]	Market Income:
[yp]	Market Income plus pensions
[yn]	Net Market Income
[yt]	Taxable Income
[yg]	Gross Income
[yd]	Disposable Income
[yc]	Consumable Income
[yf]	Final Income

- [zz] corresponds to the scale of the variable:

Letters	Scale
[pc]	per capita
[pa]	per adult equivalent
[hh]	household

Please note that the per capita scale is mandatory since it is comparable internationally.

[4] According to new version of CEQ Handbook, to estimate Pensions as Deferred Income Scenario (PDI) and Pensions as Government Transfers (PGT), you only need to estimate one set of CEQ income concepts. According to Higgins (2017), "Producing results for a CEQ Assessment only requires producing one set of section E of the MWB, since this single set of results for section E encapsulates both treatments of pensions. Specifically, in the PDI case, pre-fiscal income is market income plus pensions and in the PGT case, pre-fiscal income is market income. The user would thus focus on results using market income plus pensions as pre-fiscal income when looking at results treating pensions as deferred income, and using market income as pre-fiscal income when looking at results treating pensions as government transfers. Section D, on the other hand, automatically creates separate sets of results for the two scenarios so that results are easier to interpret."

The CEQ Country Teams must follow the examples of the following table for naming and labeling the CEQ core income concepts variables.

Table 3: Harmonized Microdata CEQ income concepts variables' names and labels

Variable	Label	Content	Definition
ym_pc	Market Income (per capita)	Per capita estimated market income in LCU (annual basis)	See Chapter 5 of Handbook (Higgins and Lustig, 2017) for definitions.
yp_pc	Market Income plus pensions (per capita)	Per capita estimated market income plus pensions in LCU (annual basis)	See Chapter 5 of Handbook (Higgins and Lustig, 2017) for definitions.
yg_pc	Gross Income (per capita)	Per capita estimated gross income in LCU (annual basis)	See Chapter 5 of Handbook (Higgins and Lustig, 2017) for definitions.
yt_pc	Taxable Income (per capita)	Per capita estimated taxable income in LCU (annual basis)	See Chapter 5 of Handbook (Higgins and Lustig, 2017) for definitions.
yd_pc	Disposable Income (per capita)	Per capita estimated disposable income in LCU (annual basis)	See Chapter 5 of Handbook (Higgins and Lustig, 2017) for definitions.
yc_pc	Consumable Income (per capita)	Per capita estimated consumable income in LCU (annual basis)	See Chapter 5 of Handbook (Higgins and Lustig, 2017) for definitions.
yf_pc	Final Income (per capita)	Per capita estimated final income in LCU (annual basis)	See Chapter 5 of Handbook (Higgins and Lustig, 2017) for definitions.

D. Fiscal Interventions

D.1 Naming fiscal interventions variables

The CEQ Country Teams must include the fiscal interventions variables in the Harmonized Microdata. The number of the fiscal intervention variables that you will have depends on the scope of the analysis of your specific country.

The team must adopt the following structure for naming the fiscal intervention variables:

`xxx_yyyy_zz`

Where

- [xxx] corresponds to the type of fiscal intervention:

Letters	Fiscal Intervention category
[dtx]	Direct Tax
[itx]	Indirect Tax
[dtr]	Direct Transfers
[pen]	Contributory Pensions
[sub]	Indirect Subsidy
[con]	Social contributions
[hlt]	In Kind Transfers Education
[edu]	In Kind Transfers Health
[oth]	Other in Kind Transfers such as housing
[pen]	Contributory pensions

- [yyyy] corresponds to the name of the fiscal intervention. Please create an acronym or abbreviation for the program to which you are referring. This acronym or abbreviation must be up to 4 letters. The following table shows two examples of acronyms using four letters:

Letters	Name of the program
[pitx]	Personal income tax
[opor]	Oportunidades

- [zz] corresponds to the scale of the variable:

Letters	Scale
[pc]	per capita
[pa]	per adult equivalent
[hh]	household
[in]	individual

These letters specify if the values are per capita [pc] or per adult equivalent [pa]. Although you have defined per capita or per adult equivalent benefits, it is also important to specify when you have variables at the individual [in] or household [hh] level. If, in addition to the per capita variable, the dataset also contains variables that capture benefits that go to an individual, the structure for naming the variable is [xxx]_[yyyy]_[in]. If the benefit goes to household, the structure for naming the variable is [xxx]_[yyyy]_[hh].

D.2 Labeling fiscal interventions variables

It is important that you use appropriate labels for each fiscal intervention. The CEQ Country Teams must use the following structure:

"[kkk][Name of the program][Scale of the variable]"

Where

- [kkk] corresponds to the type of direct transfer:

Letters	Type of Direct Transfer
[cct]	Conditional Cash Transfer
[ncc]	Non-Conditional Cash Transfer
[nct]	Near Cash Transfer
[ncp]	Non Contributory Pension

The following table shows examples on how to follow the above-mentioned structure of the variable names and their labels.

Table 4: Harmonized Microdata fiscal interventions' variables names and labels

Variable	Label	Contents	Definition
Direct Taxes			
dtx_pitx_pc	Personal Income Tax (per capita)	Total amount paid by the household for personal income tax in per capita terms. The value is in LCU and on an annual basis.	Household per capita value in LCU (annual basis) of the personal income tax: [name or regime of income tax that you are allocating]
dtx_pat_pa	Payroll Tax (per adult equivalent)	Total amount paid by the household for payroll tax in adult equivalent terms. The value is in LCU and on an annual basis.	Household per adult equivalent value in LCU (annual basis) of the personal income tax: [name or regime of income tax that you are allocating]
dtx_pat_hh	Payroll Tax (household)	Total amount paid by the household for payroll tax. The value is in LCU and on an annual basis.	Household value in LCU (annual basis) of the personal income tax: [name or regime of income tax that you are allocating]
…add other variables as needed			
Contributions to social security			
con_ssip_pc	Contributions to the Social Security Institute (SSI) (per capita)	Total amount paid by the household for social contributions to the social security institute in per capita terms. The value is in LCU and on an annual basis.	Household per capita value in LCU (annual basis) of [name or regime of social contributions that you are allocating].
con_hein_pc	Contributions to the public health care system (per capita)	Total amount paid by the household for social contributions to the health system in per capita terms. The value is in LCU and on an annual basis.	Household per capita value in LCU (annual basis) of [name or regime of social contributions that you are allocating].
…add other variables as needed			
Direct transfers			
dtr_ccft_pc	CCT [short name of the program] (per capita)	Total amount received by the household for the CCT program in per capita terms. The value is in LCU and on an annual basis.	Household per capita value in LCU (annual basis) of [name of the CCT program that you are allocating]
dtr_ncpe_pc	NCP [short name of the program] (per capita)	Total amount received by the household for the noncontributory pensions in per capita terms. The value is in LCU and on an annual basis.	Household per capita value in LCU (annual basis) of [name of the Non-Contributive Pension program that you are allocating]
…add other variables as needed			
Indirect subsidies			
sub_gaso_pc	Subsidy to Gas Oil (per capita)	Total amount received by the household for the subsidy to gas oil in per capita terms. The value is in LCU and on an annual basis.	Household per capita value in LCU (annual basis) of [name of subsidy program that you are allocating].
…add other variables as needed			
Indirect taxes			
itx_vat_pc	Value-Added Tax (per capita)	Total amount paid by the household for the value-added tax in per capita terms. The value is in LCU and on an annual basis.	Household per capita value in LCU (annual basis) of [complete name of indirect tax that you are allocating].

itx_etob_pc	Tobacco Excises (per capita)	Total amount paid by the household for the tobacco excises tax in per capita terms. The value is in LCU and on an annual basis.	Household per capita value in LCU (annual basis) of [complete name of indirect tax that you are allocating].
…add other variables as needed			
In-Kind Health			
hlt_prim_pc	In-Kind Health Benefits: Primary Level (per capita)	Total amount received by the household for primary health services in per capita terms. The value is in LCU and on an annual basis.	Household per capita value in LCU (annual basis) of [level that you are allocating].
hlt_hosp_pc	In-Kind Health Benefits: Hospitals (per capita)	Total amount received by the household for hospital services in per capita terms. The value is in LCU and on an annual basis.	Household per capita value in LCU (annual basis) of [level that you are including].
…add other variables as needed			
In-Kind Education			
edu_prim_pc	In-Kind Education Benefits: Primary Level (per capita)	Total amount received by the household for primary education in per capita terms. The value is in LCU and on an annual basis.	Household per capita value in LCU (annual basis) of [level that you are allocating].
edu_tert_pc	In-Kind Education Benefits: Tertiary Level (per capita)	Total amount received by the household for tertiary education in per capita terms. The value is in LCU and on an annaul basis.	Household per capita value in LCU (annual basis) of [level that you are allocating].

Examples on labels for fiscal interventions:

dtx_pits_pc "Personal Income Tax Simplified Regime (per capita)"

itx_etob_pc "Excises on Tobacco (per capita)"

dtr_bofa_pc "CCT Bolsa Familia (per capita)"

dtr_unbe_pc "Unemployment Benefits (per capita)"

dtr_bpc_pc "NCP Beneficio de Prestacao Continuada (per capita)"

edu_prim_pc "Primary level of education benefits (per capita)"

hlt_outp_pc "Outpatient Health Services (per capita)"

E. Target and beneficiary population, and Tax Payers

Before generating the target and beneficiary population variables, the CEQ Country Team must define who is the target populations per each fiscal intervention and know how to treat benefit recipients when it is not clear who the direct beneficiary is.

This guide does not explain how to identify the population who receive benefits and pay taxes because this process takes place when fiscal interventions are being allocated. You can find a detailed explanation of how to allocate the different fiscal interventions in Higgins and Lustig (2017).

This section first gives some shed light on how to define target population and how to treat benefit recipients when it is not clear who the direct beneficiary is, then follows with an applied example of target and beneficiary population, and finally explains how to create and label the corresponding variables.

E1. Defining Target Population

Defining the Target population can be sometimes a difficult task because program rules are not always clear. In this sense, it is important to note that CEQ Country Teams must take the stablished targeting criteria for defining the target population of each fiscal intervention. If the targeting rules are not defined, do not assume or guess the definition of the target population. Table 5 gives a general guide on possible criteria for defining the target population of each type of fiscal intervention. Some examples based on Table 5 are: conditional cash transfer can target household that live under the extreme national poverty line and at least one member of the household is a child; noncontributory pensions can target individuals who are over 65 years old and does not have the right to receive a contributory pension; primary education can target children who are between 6 and 13 years old. Please note that for tertiary education, the age range proposed to define target population is from the theoretical entrance age to the theoretical entrance age plus theoretical duration in years of the first stage of tertiary education according to national criteria. For contributory health benefits, the suggested target population is all who are eligible according to national criteria, considering both contributors and dependents. For non-contributory health benefits, the suggested target population is all who are eligible for non-contributory health benefits according to national criteria and not eligible for contributory health benefits; exclude those likely to have private health insurance[5].

Table 5: Target population definitions.

Fiscal Intervention	Target Population
Direct Taxes	Include individuals who have taxable income larger than the minimum legal taxable income.
Direct Transfers	Include individuals or households that meet the program`s eligibility rules (if there is a defined criteria) or proxies to target the poor. Please explain if the targeted beneficiary is the household or individual. If the targeting rules are not defined, don't assume or guess the definition of target population. If you don't have target population for a specific program you still will be able to analyze the coverage of the program that will be available in MWB E18. Some examples of targeting criteria are: • Age • Attendance to public school • Children in the household • Educational level (of household members or household head) • Ethnical group • Geographical • Gender • National socioeconomic groups • Not being part of social security system • Proxy-mean test • Use of public facilities (health or public pharmacies) • Vulnerable population (orphan, widows, etc.)

[5] This paragraph is based on Higgins (2017).

Pensions	
Old age Pensions (Contributory and Noncontributory)	Include individuals in retirement age (according to national criteria) • For example, in some countries, this could be o 65 years for male o 60 years for female
Noncontributory Pensions (Social or minimum pension)	Include individuals who meet the program`s eligibility rules. • For example, in some countries, this could be o 65 years for male o 60 years for female
Indirect Taxes	The concept of a "target population" is not applicable
Indirect Subsidies	The concept of a "target population" is not applicable
Education	
Pre-school, Primary, Secondary	Include individuals who are within the theoretical ages by educational level. • Educational levels used for defining the theoretical ages for target population must be consistent with the educational levels used to impute the benefits. • The benefits could have been allocated using: • National criteria: in this case use the entrance age and theoretical duration in years stablished in the country under analysis. • International Standard Classification of Education (ISCED): in this case use the entrance age and theoretical duration in years consistent with this classification. Information is available in Unesco Mappings, (http://uis.unesco.org/en/isced-mappings). Please keep in mind that the definition of the educational levels use for the allocation of the benefits, for the coverage and target population must be consistent.
Tertiary	Include individuals who are within the theoretical ages of the first stage of tertiary education. • The definition of tertiary education used for defining the target population must be consistent with the definition of tertiary education used to impute the benefits of tertiary education. • The benefits could have been allocated using: • National criteria: in this case use the theoretical entrance age and the theoretical entrance + theoretical duration (in years) according to national criteria. • International Standard Classification of Education (ISCED): in this case use the entrance age and theoretical duration in years consistent with this classification. Information is available in Unesco Mappings, (http://uis.unesco.org/en/isced-mappings). Please keep in mind that the definition of the educational levels use for the allocation of the benefits, for the coverage and target population must be consistent.
Health	
Contributory	All population that is eligible for contributory health system or programs according to national criteria. Consider the contributors and dependents covered by health plan (e.g., wife and children under 5 years). If the eligibility criteria for contributory health system is not well specified in the country, don't include target population.
Non-contributory	All population that is eligible for non-contributory health and is not eligible for contributory health system according to national criteria. Exclude those likely to have private and public health insurance.

	If the country does not have specified criteria for targeting, please use all population that is not covered by the public or privatized contributory health system. If information is available include programs like vaccinations, pre-natal visits, regular checkups for infants, childbirth attention in hospital or specialized health center. To include this programs you must have information on coverage and per-capita expenditure by each type of program. Please keep in mind that the definition used for target population must be consistent with the allocation of health benefits.
Housing	According to program eligibility rules, otherwise don't include target population.

E2. Identifying Benefit Recipients

Sometimes surveys do not specify who the direct beneficiary of a specific social program. This can happen for example when the survey only asks whether anyone in the household receives benefits from a program or when the entire household instead of a specific member of the household is who receive the benefit. In these cases, the CEQ Country Team do the following:

a) mark one member of the household as the direct beneficiary when it is not clear who is the direct beneficiary.
b) select the household head as the direct beneficiary[6] when the entire household is the recipient.

E.3 Applied examples of targeted and recipients beneficiaries

Example 1: Primary Education benefits

The following table shows an example of a Primary Education benefits. The first household has two children that belong to target population, according to their age, but only one is enrolled in a public school of Primary educational level. Second and third household does not have any children eligible to primary education. To identify target population the variable *"edu_prim_ti"* has the value of 1 for two children of the first household. It can also be noticed, through the variable "*edu_prim_ri*" that only one child of this household is enrolled in a public school of Primary level. Finally, the table shows through the variable "*edu_prim_pc*" that the first household receives a per capita benefit of 100 and the second and third household do not receive any benefits for primary education.

hhid	Age	sex	relation	edu_prim_pc	edu_prim_ti	edu_prim_ri	edu_prim_th	edu_prim_rh
1	35	male	Head	100	0	0	.	.
1	33	female	Spouse	100	0	0	.	.
1	8	male	Son	100	1	0	.	.
1	9	female	Daughter	100	1	1	.	.
2	45	male	Head	0	0	0	.	.
2	49	female	Sister	0	0	0	.	.
3	22	male	Head	0	0	0	.	.
3	19	female	Spouse	0	0	0	.	.
3	3	female	Son	0	0	0	.	.

[6] In this case, it will be posible to estimate coverage of households and coverage of direct and indirect beneficiaries can be estimated. According to Higgins (2017); "For programs where the target population is defined at the household level, the head of the targeted household should be marked as the target person in the data set and the direct beneficiary results should be ignored; in other words, for programs defined at the household level, only use the household beneficiary and direct and indirect beneficiary results".

Example 2: CCT targeted to household

The following table shows an example of a CCT program, where the direct beneficiary of the first household –variable hhid with value 1—is the entire household. This can be noticed through the "*dtr_cctx_rh*" variable, because it has a 1 for the head of the household—the rh ending should be used when the direct beneficiary is the entire household. It can also be noticed, through the variable "*dtr_cctx_rh*" that the household meets the requirements of the program for being eligible. This is because that variable has a value of 1 for the head of the household and the ending "th" of the variable name reflects that the target beneficiary is the household. Finally, the table shows through the variable "*dtr_cctx_pc*" that the first household receives a per capita benefit of 75.

hhid	Age	sex	relation	dtr_cctx_pc	dtr_cctx_ti	dtr_cctx_ri	dtr_cctx_th	dtr_cctx_rh
1	35	male	Head	75	.	.	1	1
1	33	female	Spouse	75	.	.	0	0
1	8	male	Son	75	.	.	0	0
1	9	female	Daughter	75	.	.	0	0
2	45	male	Head	0	.	.	0	0
2	49	female	Sister	0	.	.	0	0
3	22	male	Head	0	.	.	1	1
3	19	female	Spouse	0	.	.	0	0
3	3	female	Son	0	.	.	0	0

E.4 Generating Variables and its labels for Target and Beneficiary Population

The CEQ Country Team needs to create three types of variables in regard to the target and beneficiary population: one variable for recipients of benefits, one variable for the target population according to eligibility criteria (see Table 5 for target population definitions), and another one for whether the individual's age corresponds to the target age cohort for each level of education. The number of variables for the three overarching target and beneficiary population type variables depends on the scope of the analysis of each country.

The CEQ Country Teams must use the following structure to name the target and beneficiary variables:

`xxx_yyyy_ww`

Where

- [xxx] represents the type of intervention:

Letters	Fiscal Intervention category
[dtx]	Direct Tax
[itx]	Indirect Tax
[dtr]	Direct Transfers
[pen]	Pensions
[sub]	Indirect Subsidy
[con]	Social contributions
[hlt]	In Kind Transfers Education
[edu]	In Kind Transfers Health
[oth]	Other in Kind Transfers such as housing
[pen]	Contributory pensions

- [yyyy] corresponds to the name of the fiscal intervention. Please create an acronym or abbreviation for the program to which you are referring. This acronym or abbreviation must be up to 4 letters.

Letters	Name of the program
[pitx]	Personal income tax
[opor]	Oportunidades

- [ww] define targets and recipients, and their corresponding level.

Letters	Targets, recipients and its level
[ti]	individual to whom the benefit (tax) was targeted
[th]	household to whom the benefit (tax) was targeted
[ri]	individual who received (paid) the benefit (tax)
[rh]	household who received (paid) the benefit (tax)

The following table shows examples on how to follow the above-mentioned structure of the variable names and their labels.

Table 6: Harmonized Microdata targeting and coverage variables' name and labeling

Variable	Label	Contents	Definition
Direct transfers			
dtr_ccft_th	Target of CCT Family Transfer (FT) program	[0=No, 1=Yes]	The name of the CCT program that you are allocating.
dtr_ncpe_ri	Recipient of Non-Contributory pensions	[0=No, 1=Yes]	The name of the Non-Contributory Pension program that you are allocating.
…add other variables as needed			
Indirect subsidies			
sub_gaso_rh	Recipient of Subsidy to Gas Oil	[0=No, 1=Yes]	Name of subsidy program that you are allocating.
…add other variables as needed			
In-Kind Education			
edu_prim_ri	Recipient of primary education	[0=No, 1=yes]	Attends (level that you are allocating).
edu_prim_ti	Target of primary education	[0=No, 1=yes]	Within the age range to attend (level that you are allocating).

F. Other: survey and poverty variables

The following table shows the name of the survey and poverty variables and their corresponding labels, which must be used.

Variable	Label	Contents	Definition
Weight	Sampling weight	Value of sampling weight	
Psu	Primary sampling Unit	Value of primary sampling unit	
Strata	Sampling stratum	Value of Strata	
…add other variables as needed			
Poverty Lines			
pline_ext	National Extreme poverty line	Value of Extreme Poverty Line in LCU (yearly basis)	
pline_mod	National Moderate poverty line	Value of Moderate Poverty Line in LCU (yearly basis)	
…add other variables as needed			

III. Applied example

Application to the Panama case:

Variable name	Label	Definition
hhid	"Household Identifier"	Unique identifier per household
Hsize	"Household size"	Number of household members
Urban	"Living in urban area"	[0=rural, 1=urban]
Direct Tax		
dtx_pits_pc	"Personal Income Tax Simplified Regime (per capita)"	Household per capita value in LCU (annual basis) of the personal income tax regime: persona natural
dtx_pits_hh	"Personal Income Tax Simplified Regime (household)"	Household value in LCU (annual basis) of the personal income tax regime: persona natural
dtx_pits_ri	"Payer Personal Income Tax Simplified Regime "	Payer of personal income tax regime: persona natural
dtx_pits_ti	"Target Personal Income Tax Simplified Regime"	Target of personal income tax regime: persona natural
Direct transfer		
dtr_ccto_pc	"CCT Oportunidades (per capita)"	Household per capita value in LCU (annual basis) of Red de Oportunidades
dtr_ccto_hh	"CCT Oportunidades (household)"	Household value in LCU (annual basis) of Red de Oportunidades
dtr_ccto_rh	"Recipient of oportunidades"	Beneficiary of Red de Oportunidades
dtr_ccto_th	"Target of oportunidades"	Target of Red de Oportunidades
dtr_buni_pc	"Scholarship Beca Universal (per capita)"	Household per capita value in LCU (annual basis) of Beca Universal
dtr_buni_hh	"Scholarship Beca Universal (household)"	Household value in LCU (annual basis) of Beca Universal
dtr_buni_in	"Scholarship Beca Universal (individual)"	Individual value in LCU (annual basis) of Beca Universal
dtr_buni_ri	"Recipient of Beca Universal"	Beneficiary of Beca Universal
dtr_buni_ti	"Target of Beca Universal"	Target of Beca Universal
Subsidies		
sub_elec_pc	"Subsidy to electricity (per capita)"	Household per capita value in LCU (annual basis) of subsidy to electricity
sub_elec_hh	"Subsidy to electricity (household)"	Household value in LCU (annual basis) of subsidy to electricity
sub_elec_rh	"Recipient of subsidy to electricity"	Beneficiary of subsidy to electricity
sub_elec_th	"Target of subsidy to electricity"	Target of subsidy to electricity
In kind education		
edu_lws_in	"In-Kind education benefits: Lower Secondary (individual)"	Individual value in LCU (annual basis) of lower secondary education
edu_lws_hh	"In-Kind education benefits: Lower Secondary (household)"	Household value in LCU (annual basis) of lower secondary education
edu_lws_pc	"In-Kind education benefits: Lower Secondary (per capita)"	Household per capita value in LCU (annual basis) of lower secondary education
edu_lws_ri	"Recipient of lower secondary education"	Attends public school in lower secondary level
edu_lws_ti	"Target of lower secondary education"	Within the age range to attend lower secondary level
Health		
hlt_prim_in	"In-Kind Health Benefits: Primary Level (individual)"	Individual value in LCU (annual basis) of primary health care
hlt_prim_hh	"In-Kind Health Benefits: Primary Level (household)"	Household value in LCU (annual basis) of primary health care

hlt_prim_pc	"In-Kind Health Benefits: Primary Level (per capita)"	Household per capita value in LCU (annual basis) of primary health care
hlt_prim_ri	"Recipient of primary health care"	Beneficiary of primary health care
hlt_prim_ti	"Target of primary health care"	Target of primary health care
Income concepts		
yp_pc	Market Income plus pensions (per capita)	Per capita estimated market income in LCU (annual basis)
yp_hh	Market Income (household)	Total household estimated market income plus pensions in LCU (annual basis)
yn_pc	Net Market Income (per capita)	Per capita estimated net market income in LCU (annual basis)
yn_hh	Net Market Income (household)	Total household estimated net market income in LCU (annual basis)
yg_pc	Gross Income (per capita)	Per capita estimated gross income in LCU (annual basis)
yg_hh	Gross Income (household)	Total household estimated gross income in LCU (annual basis)
yt_pc	Taxable Income (per capita)	Per capita estimated taxable income in LCU (annual basis)
yt_hh	Taxable Income (household)	Total household estimated taxable income in LCU (annual basis)
yd_pc	Disposable Income (per capita)	Per capita estimated disposable income in LCU (annual basis)
yd_hh	Disposable Income (household)	Total household estimated disposable income in LCU (annual basis)
yc_pc	Consumable Income (per capita)	Per capita estimated consumable income in LCU (annual basis)
yc_hh	Consumable Income (household)	Total household estimated consumable income in LCU (annual basis)
yf_pc	Final Income (per capita)	Per capita estimated final income in LCU (annual basis)
yf_hh	Final Income (household)	Total household estimated final income in LCU (annual basis)

Example Stata Code.

```
Notes:

oldvar stands for old variable, which refers to the name of the variable in your
original data set

label def yesno 0 "No" 1 "Yes"  //This will be used in different labels

*1.- Survey and household variables

* a.-Survey variables

rename oldvar weight

rename oldvar psu

rename oldvar strata

rename oldvar pline_ext


label var weight "Sampling weight"

label var psu "Primary sampling unit"

label var strata "Sampling stratum"

label var pline_ext "National Extreme poverty line"

label var pline_mod "National Moderate poverty line"


*b.- household variables

rename oldvar hhid

rename oldvar hsize

rename oldvar adulteq

rename oldvar urban

rename oldvar inf_water

rename oldvar inf_elect

rename oldvar inf_roof

rename oldvar inf_walls

rename oldvar inf_sewage

rename oldvar inf_floor

rename oldvar inf_roads


label var hhid "Household identifier"

label var hsize "Household size"

label var adulteq "House size in adult equivalent scale"

label var urban "Living in urban area|0=rural 1=urban"

label def 0 "Rural" 1 "Urban"
```

```
label var inf_water "Access to piped water"

label val inf_water yesno

label var inf_elect "Access to electricity"

label val inf_elect yesno

label var inf_roof "Quality roofing [list materials]"

label val inf_roof yesno

label var inf_walls "Quality walls [list materials]"

label val inf_walls yesno

label var inf_sewage "Quality sanitation"

label val inf_sewage yesno

label var inf_floor "Quality floor [list materials]"

label val inf_floor yesno

label var inf_roads "Access to paved roads"

label val inf_roads yesno


*2.-Household member variables

rename oldvar age

rename oldvar gender

rename oldvar relation [you may need to recode this variable]

rename oldvar at_school

rename oldvar level_school

rename oldvar type_school


label var age "Age of the household member"

label var gender "Gender [0Male,1Female]"

label def sex 0 "Male" 1 "Female"

label val gender sex

label var relation "Relationship to household head"

label def headrelation 1 "Head" 2 "Spouse/partner" 3 "Daughter/son" 4 "Parents/in
law" 5 "Other relatives" 6 "Other non relatives"

label val relation headrelation

label var at_school "attending to a school"

label val at_school yesno

label var level_school "Level of schooling that is attending"

label def level 1 "Preschool" 2 "Primary" 3 "Lower secondary" 4 "Upper secondary
" 5 "Secondary total" 6 "Post secondary non-tertiary" 7 "Tertiary (Bachelor`s)" 8
"Master, doctoral, equi" 9 "Other"
```

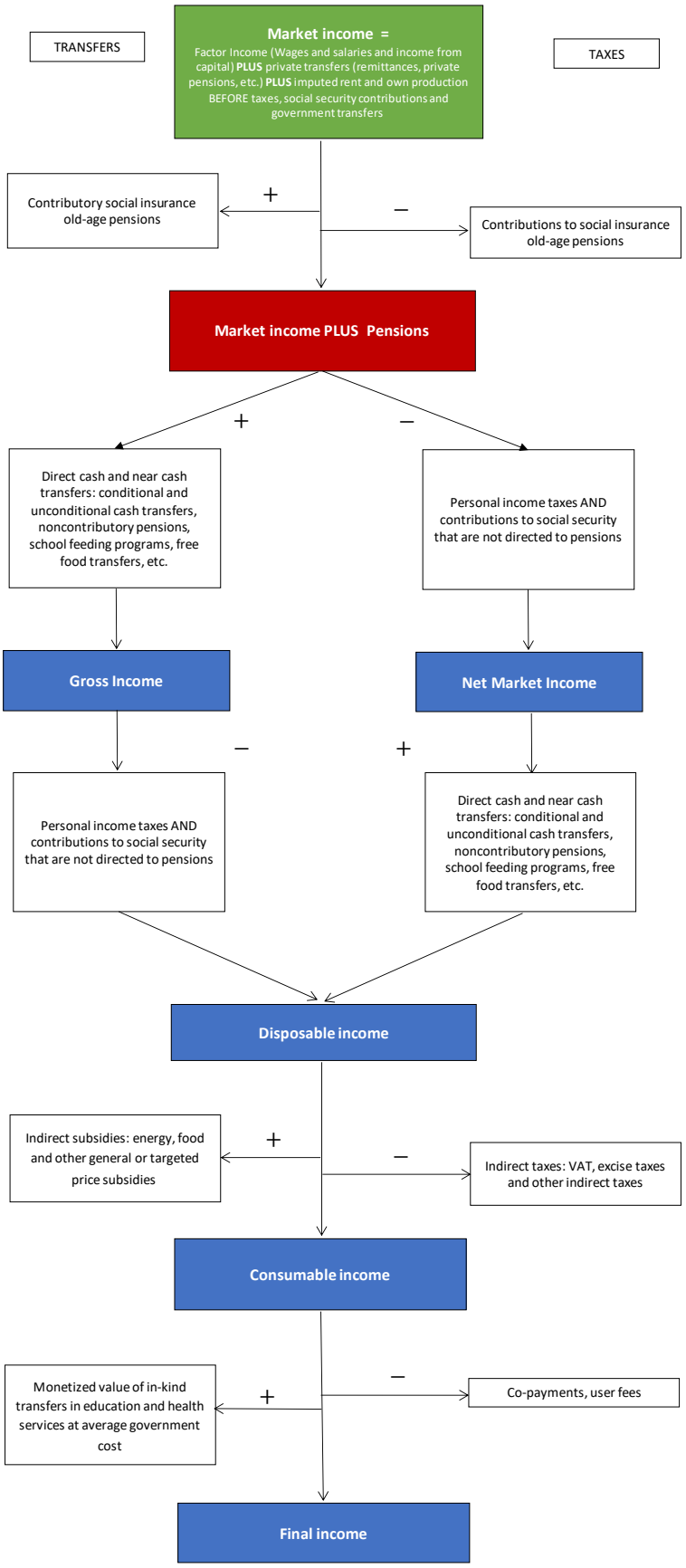
```
label val level_school level

label var type_school "Type of school that is attending (Public, private,semi-
priv,other)"

label def type 1 "Public" 2 "Private" 3 "semi-privater" 4 "Other"

label val type_school type


*3.- CEQ Income Concepts

rename oldvar  ym_pc

rename oldvar  yp_pc

rename oldvar  yn_pc

rename oldvar  yg_pc

rename oldvar  yt_pc

rename oldvar  yd_pc

rename oldvar  yc_pc

rename oldvar  yf_pc


label var yp_pc "Market Income (per capita)"

label var yp_pc "Market Income plus pensions  (per capita)"

label var yn_pc "Net Market Income  (per capita)"

label var yg_pc "Gross Income  (per capita)"

label var yt_pc "Taxable Income  (per capita)"

label var yd_pc "Disposable Income  (per capita)"

label var yc_pc "Consumable Income  (per capita)"

label var yf_pc "Final Income  (per capita)"


*In case you have at household level in addition to per capita

rename oldvar  yp_hh

rename oldvar  yn_hh

rename oldvar  yg_hh

rename oldvar  yt_hh

rename oldvar  yd_hh

rename oldvar  yc_hh

rename oldvar  yf_hh


label var ym_hh "Market Income (household)"
```

```
label var yp_hh "Market Income plus pensions  (household)"

label var yn_hh "Net Market Income  (household)"

label var yg_hh "Gross Income  (household)"

label var yt_hh "Taxable Income  (household)"

label var yd_hh "Disposable Income (household)"

label var yc_hh "Consumable Income (household)"

label var yf_hh "Final Income (household)"


*Fiscal Interventions

*Direct taxes

rename oldvar  dtx_pit_pc // remember you can change "pit" for the letters that
best describe the name of the fisc. intervention

label var dtx_pit_pc "Personal Income Tax [please write the name of the regime]
(per capita)"

*Indirect taxes

rename oldvar  itx_vat_pc // remember you can change "vat" for the letters that
best describe the name of the fisc. intervention

rename oldvar  itx_etob_pc // remember you can change "etob" for the letters that
best describe the name of the fisc. intervention

label var itx_vat_pc  "Value added tax (per capita)"

label var itx_etob_pc "Excises on Tobacco (per capita)"


* Direct transfers

rename oldvar dtr_xxxx_pc // remember you can change "cctf" for the letters that
best describe the name of the fisc. intervention

rename oldvar dtr_bofa_pc // remember you can change "cctf" for the letters that
best describe the name of the fisc. intervention

rename oldvar dtr_ccto_pc // remember you can change "ccto" for the letters that
best describe the name of the fisc. intervention

rename oldvar dtr_ncbs_pc // remember you can change "ncbs" for the letters that
best describe the name of the fisc. intervention

label var dtr_xxxx_pc "[type of transfer] [name of transfer] (per capita)" //
General model

label var dtr_bofa_pc "CCT Bolsa Familia (per capita)"

label var dtr_ccto_pc "CCT Oportunidades (per capita)"

label var dtr_ncbs_pc "NCP Pension basica solidaria (per capita)"

* Subsidies

rename oldvar  sub_gaso_pc // remember you can change "gaso" for the letters that
best describe the name of the fisc. intervention

rename oldvar  sub_elec_pc // remember you can change "elec" for the letters that
best describe the name of the fisc. intervention
```

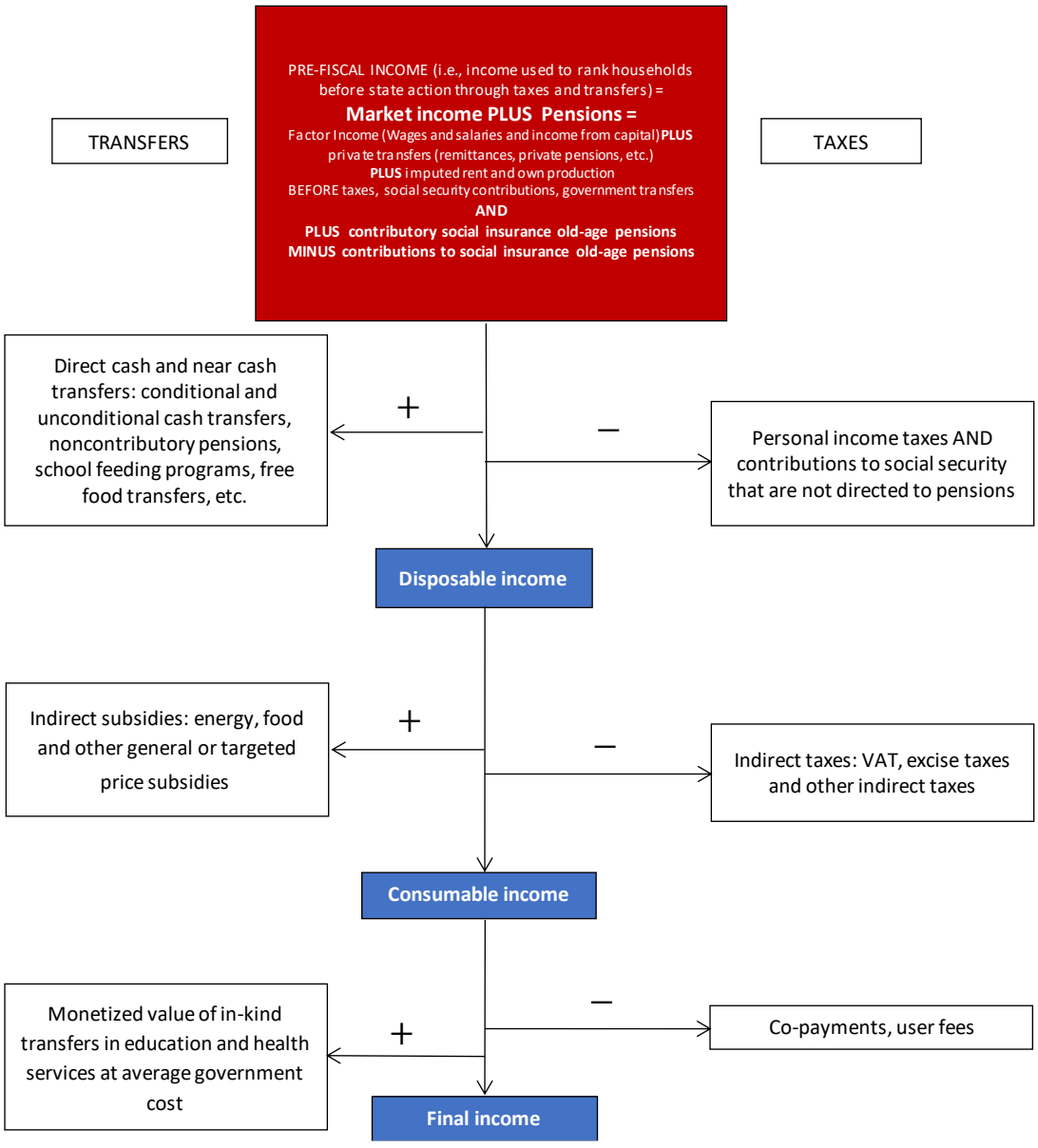
```
label var sub_gaso_pc "Subsidy to Gas Oil (per capita)"

label var sub_elec_pc "Subsidy to electricity (per capita)"


*Contributions to social security

rename oldvar  con_chlt_pc // remember you can change "chlt" for the letters that best describe the name of the fisc. intervention

rename oldvar  con_chlt_pc // remember you can change "chlt" for the letters that best describe the name of the fisc. intervention

label var con_hein_pc "Contribution to the public health care system (per capita)"

label var con_ssip_pc "Contribution to the Social Security Instute (SSI)(per capita)"


*Health

rename oldvar hlt_prim_pc

rename oldvar hlt_syt_pc

label var hlt_prim_pc "In-Kind Health Benefits: Primary Level (per capita)"

label var hlt_prim_pc "In-Kind Health Benefits: Secondary and tertiary Level (per capita)"

*Education

rename oldvar edu_pesc_pc

rename oldvar edu_prim_pc

rename oldvar edu_lwsc_pc

rename oldvar edu_upsc_pc

rename oldvar edu_tosc_pc

rename oldvar edu_psnt_pc

rename oldvar edu_tert_pc

rename oldvar edu_pter_pc

rename oldvar edu_diff_pc


label var edu_pesc_pc "In-Kind education benefits: Pre-school (per capita)"

label var edu_prim_pc "In-Kind education benefits: Primary (per capita)"

label var edu_lwsc_pc "In-Kind education benefits: Lower Secondary (per capita)"

label var edu_upsc_pc "In-Kind education benefits: Upper Secondary (per capita)"

label var edu_tosc_pc "In-Kind education benefits: Total Secondary (per capita)"

label var edu_psnt_pc "In-Kind education benefits: Post Secondary non tertiary (per capita)"
```

```
label var edu_tert_pc "In-Kind education benefits: Tertiary (per capita)"

label var edu_pter_pc "In-Kind education benefits: Post tertiary (per capita)"

label var edu_diff_pc "In-Kind education benefits: Diferencial (per capita)" //
this is an example for category 9 = other
```

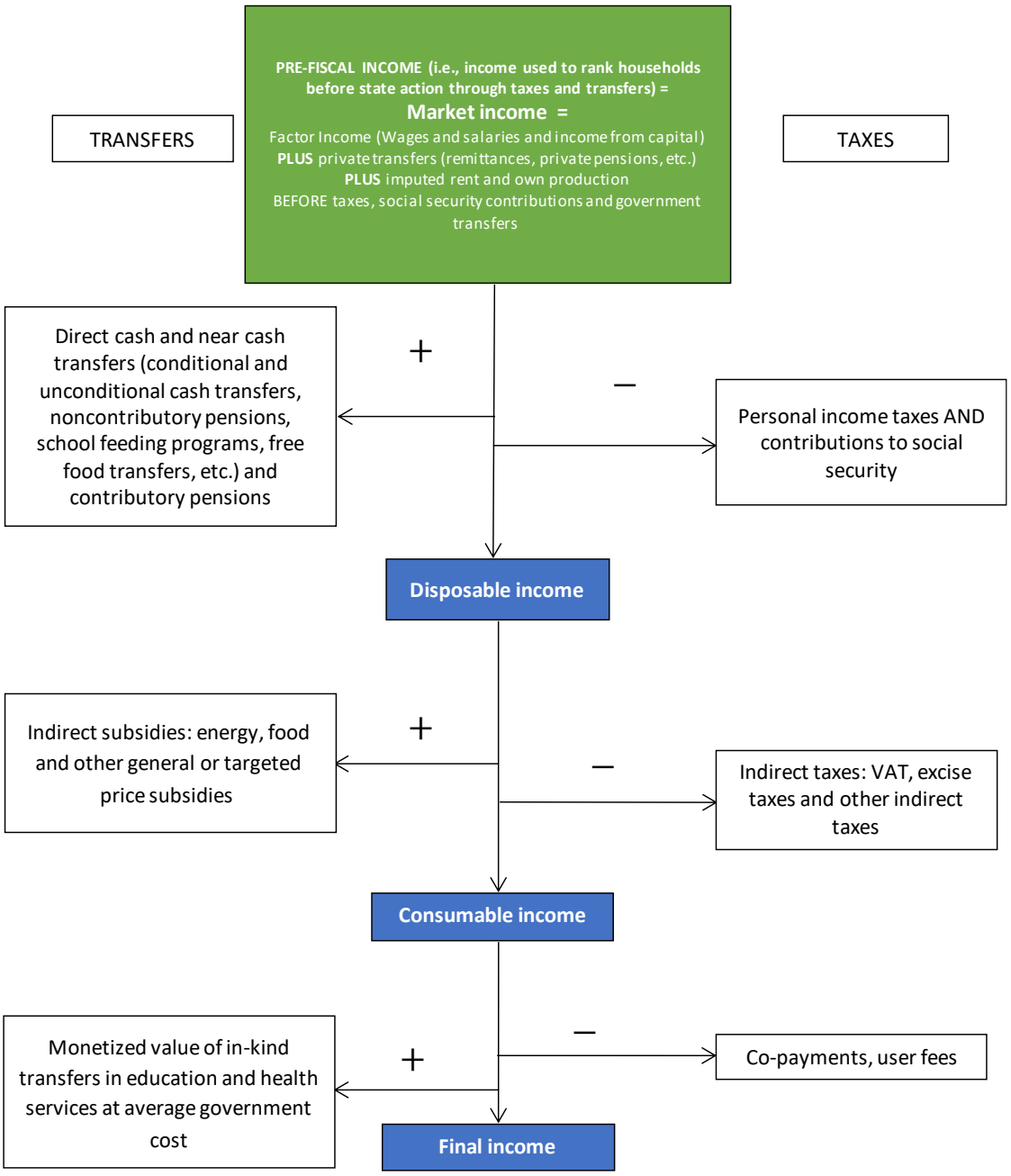
Annex 1. CEQ Core Income Definitions.



Contributory Pensions as Deferred Income



Contributory Pensions as Government Transfer



References

Higgins, Sean and Nora Lustig. 2017. "Allocating Taxes and Transfers, Constructing Income Concepts, and Completing Sections A, B, and C of CEQ Master Workbook," Chapter 5 in *Commitment to Equity Handbook. Estimating the Impact of Fiscal Policy on Inequality and Poverty*, edited by Nora Lustig (Brookings Institution Press and CEQ Institute, Tulane University).

Higgins, Sean. 2017. "Producing Indicators and Results, and Completing Sections D and E of CEQ Master Workbook Using the CEQ Stata Package," Chapter 7 in *Commitment to Equity Handbook. Estimating the Impact of Fiscal Policy on Inequality and Poverty*, edited by Nora Lustig (Brookings Institution Press and CEQ Institute, Tulane University).

Lustig, Nora, editor. Forthcoming. *Commitment to Equity Handbook. Estimating the Impact of Fiscal Policy on Inequality and Poverty* (Brookings Institution Press and CEQ Institute, Tulane University). Online version October 31, 2016 is available by clicking here.

LIS,http://www.lisdatacenter.org/data-access/web-tabulator/list-of-variables/

UNESCO, http://uis.unesco.org/en/isced-mappings