A platform for batch analysis of national health surveys using Stata

Leonardo Ferreira

International Center for Equity in Health

Introduction

- Activities of the ICEH
- Background
- Stages of the analysis
- Codes in Stata
- The equiplot graph

ICEH overview

- The International Center for Equity in Health (ICEH) produces standardized and disaggregated results from household surveys

- Database with over 300 surveys representing over 100 countries

- Focus in RMNCH in low-and-middle income countries (LMIC)

- Delivers data to international partners such as WHO, UNICEF and World Bank.

Indicators and stratifiers

Indicators

- Contraceptive rate
- Measles vaccine
- Exclusively breastfed children
- Access to improved drinking water

- ..over 50 indicators

Stratifiers

- Wealth
- Sex of the child
- Urban/rural zone
- Maternal education
- Maternal age
- Geographical region

Surveys

- Main data sources:
  - Demographic and Health Surveys (DHS)
  - Multiple Indicator Cluster Survey (MICS).

- Secondary data sources:
  - Reproductive and Health Surveys (RHS)
  - National surveys

The surveys are **more or less** standardized!

SDG target for equity

17.18 Data, monitoring and accountability

"to increase significantly the availability of high-quality, timely and reliable data disaggregated by ***income, gender, age, race, ethnicity, migratory status, disability, geographic location*** *and other characteristics relevant in national contexts."*

Why is the ICEH work important?

- Able to analyse hundreds of surveys for several indicators and stratifiers

- Standardized indicators allow:
  - Comparison among countries
  - Usage of global, regional and sub-national estimates

- Flexibility to quickly analyse new indicators and/or stratifiers on demand

Data examples - contraception

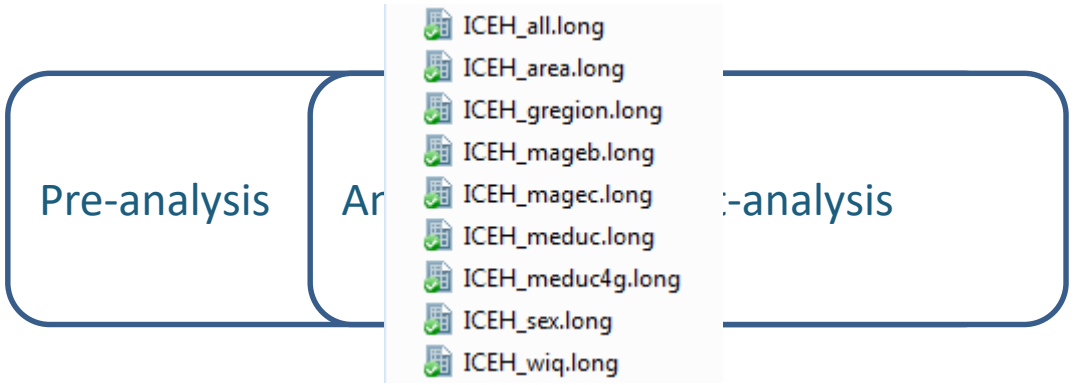
National estimates

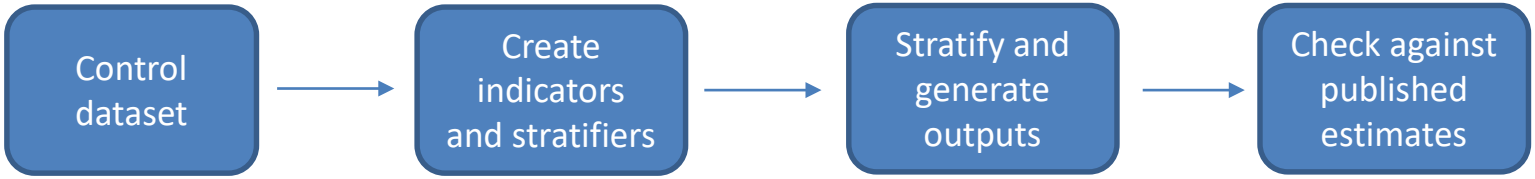
country	year	r
Peru	2000	.6887665
Peru	2006	.7241852
Peru	2012	.7549451

Wealth disaggregated estimates

country	year	level	r
Peru	2000	Q1	.5513875
Peru	2000	Q2	.6734118
Peru	2000	Q3	.7280977
Peru	2000	Q4	.7422365
Peru	2000	Q5	.7368047
Peru	2006	Q1	.6562435
Peru	2006	Q2	.7300255
Peru	2006	Q3	.7463354
Peru	2006	Q4	.7415566
Peru	2006	Q5	.7179778
Peru	2012	Q1	.7279711
Peru	2012	Q2	.7632848
Peru	2012	Q3	.774083
Peru	2012	Q4	.7580208
Peru	2012	Q5	.7456751

Stages of the analysis





Control dataset

country	year_st	sstoolscodes	abarivars	orsvars	medfacvars	sbavars	ancvars
Afghanistan	2010	1 2	a b	ca4a ca4b	a b c d e i j l o	a b c	a b c
Albania	2005	1 2	a	ca2a ca2c	a b d h i j	a b c	a b c
Argentina	2011		a b	ca4a ca4b	a b c d e h i j l o		a b c
Bangladesh	2006	1 2 5	a	ca2a ca2c	a d e h i j o	a b	a b
Bangladesh	2012	1 2 5	a b	ca4a	a b c d e h i j l o	a b c	a b c
Barbados	2012	1 2	a b	ca4a ca4b	a b i j	a b	a b
Belarus	2005	1 2	a	ca2a ca2c	a b c d e h i j l o	a b c	a b c
Belarus	2012	1 2	a b	ca4a ca4b	a b c d e i j h o	a b	a b
Belize	2006	1 2 5	a	ca2a ca2c	a b c d e h i j l o	a b c	a b c
Belize	2011	1 2	a b	ca4a ca4b	a b c d e h i j l o	a b c	a b c
Bhutan	2010	1 2	a b	ca4a ca4b	a b c d e j	a b c	a b c
Bosnia_and_Herzegovina	2006	1 2 5	a	ca2a ca2c	a b c d e h i j l o	a b c	a b c
Bosnia_and_Herzegovina	2011	1 2	a b	ca4a ca4b	a b e h i j l o	a b	a b
Burkina_Faso	2006	1 2	a b c d e f g	ca2a ca2c	a b c e	a b c d	a b c d
Burundi	2005	1 2	a	ca2a ca2c	a b c d e h i j l o	a b c	a b c
CAR	2006	1 2	a	ca2a ca2c	a b c d e h i j l o	a b c	a b c
CAR	2010	1 2	a b	ca4a ca4b	a b c d e h i j l o	a b c	a b c
Cameroon	2006	1 2 5	a	ca2a ca2c	a b c h i j l o	a b c	a b c
Chad	2010	1 2	a b	ca4a ca4b	a b c d e h i j l o	a b c	a b c
Congo_Democratic_Republic	2010	1 2	a b	ca4a ca4b	a b c d e h i j l o	a b d	a b d
Costa_Rica	2011	1 2	a b	ca4a ca4b	a b c e h i j l	a b i	a b i
Cote_dIvoire	2006	1 2 5	a	ca2a ca2c	a b c d e h i j l o	a b c	a b c
Cuba	2006						
Cuba	2010	1 2	a b	ca4a ca4b	a b c d e h		
Cuba	2014	1 2 5	i j	ca4a ca4b	a b c d e	a b c	a b
Djibouti	2006	1 2 5	a	ca2a ca2c	a b c d e h i j l o	a b c	a b c
Gambia	2005	1 2 5	a	ca2a ca2c	a b c d e h i j l o	a b c	a b c
Georgia	2005	1 2 5	a	ca2a	a b c e h i j o	a b	a b
Ghana	2006	11 12	a	ca2a ca2c	a b c d e h i j l o	a b c	a b c
Ghana	2011	1 2	a b	ca4a ca4b	a b c d e i l m o	a b c	a b c
Guinea_Bissau	2006	1 2	a	ca2a ca2c	a b c d e h i j l o	a b c	a b c
Guyana	2006	1 2 5	a	ca2a ca2c	a b c d e f h i j l	a b c d	a b c d

Obtaining control dataset information

```
global ancvars = ancvars[`line']
global sbavars = sbavars[`line']
global homefluidvars=homefluidvars[`line']
global orsvars=orsvars[`line']
global moreflcode=moreflcode[`line']
global medfacvars=medfacvars[`line']
global pnctimesvars=pnctimesvars[`line']
global pncunitvars=pncunitvars[`line']
global abarivars = abarivars[`line']
global sstoolscodes = sstoolscodes[`line']
```

Why do we need the control dataset?

Same categories, different variables

ca13i	str1	%1s	Medicamento: (Antibiótico) pastilla / jarabe
ca13j	str1	%1s	Medicamento: (Antibiótico) Inyección
ca13p	str1	%1s	Medicamento: (Otros medicamentos) Paracetamol/Acetaminofeno
ca13q	str1	%1s	Medicamento: (Otros medicamentos) Aspirina
ca13r	str1	%1s	Medicamento: (Otros medicamentos) Ibuprofeno

ca13a	str1	%1s	Medicamento: Antibiotico pastilla / jarabe
ca13b	str1	%1s	Medicamento: Antibiotico inyeccion
ca13p	str1	%1s	Medicamento: Paracetamol/ Acetaminofeno
ca13q	str1	%1s	Medicamento: Aspirina
ca13r	str1	%1s	Medicamento: Ibuprofeno

Same variables, different categories

WS1 (Fuente principal de agua potable)	code	freq
Acueducto público del IDAAN	1	4596
Acueducto público de la comunidad	2	3400
Acueducto particular	3	113
Pozo brocal protegido	21	98
Pozo artesanal protegido	31	56
Pozo artesanal no protegido	32	160

WS1 (Main source of drinking water)	code	freq
Piped into dwelling	11	4965
Piped into compound, yard or plot	12	47
Piped to neighbour	13	9
Public tap / standpipe	14	36
Tube well, Borehole	21	102
Protected well	31	289

Generating indicators

```
*** MSLV: Measles coverage (Ref. code #15)
gen mslv = .
capture allmissvar $measlesvars
if _rc == 0 & r(allmissvar) == 0 {
  replace mslv = 0 if inrange(agemo,`vacrefperiod') & alive == 1
  egen _mslv = anymatch($measlesvars), v(1/3)
  recode mslv (0=1) if _mslv == 1
}
la val mslv yn
la var mslv "Received measles vaccination"
```

Anymissvar & allmissvar

```
program define allmissvar, rclass
  version 12.0

  syntax anything [if] [in]

  local newlist ""
  foreach i of local anything {
    if (regexm("`i'", "[\*\?]")==1) {
      unab templist: `i'
      local i = "`templist'"
    }
    local newlist "`newlist' `i'"
  }
*** Values to 1, 2 ...
  tokenize `newlist'
  local nvars: word count `newlist'                    // n
  local allmissvar = 0
  forval var = 1 / `nvars' {
    cap quietly summ ``var''
    if r(N)==0 | _rc>0 {
      local ++allmissvar
    }
  }
  return scalar allmissvar = `allmissvar'==`nvars'

 end
```

```
program define anymissvar, rclass
  version 12.0

  syntax varlist(numeric) [if] [in]

*** Values to 1, 2 ...
  tokenize `varlist'
  local nvars: word count `varlist'
  local anymissvar = 0
  forval var = 1 / `nvars' {
    quietly summ ``var''
    if r(N)==0 local anymissvar = 1
  }
  return scalar anymissvar = `anymissvar'

 end
```

Results in matrices

```
OUT_gregion[8,180]
  OUT_area[2,180]
OUT_meduc[3,180]
   OUT_wiq[5,180]
   OUT_sex[2,180]
   OUT_all[1,252]
```

OUT_all[1,252]

	bcgv_r	bcgv_se	bcgv_t	bcgv_p	bcgv_ll	bcgv_ul	bcgv_df	bcgv_crit
1	.65466829	.01893556	34.57348	1.28e-126	.61656713	.69088155	433	1.9654578

	bcgv_sii	bcgv_siise	dptv_r	dptv_se	dptv_t	dptv_p	dptv_ll	dptv_ul
1	.22741043	.06123841	.41048502	.01893592	21.677585	4.228e-71	.37384392	.44814707

	dptv_cix	dptv_cixse	dptv_sii	dptv_siise	poliov_r	poliov_se	poliov_t	poliov_p
1	.10836814	.02457947	.24772844	.05556402	.49367058	.01864347	26.479544	8.409e-93

country	year	bcgv_r	bcgv_se	bcgv_t	bcgv_p	bcgv_ll	bcgv_ul	bcgv_df	bcgv_crit
Afghanistan	2010	.6546683	.0189356	34.57348	0	.6165671	.6908816	433	1.965458
Albania	2005	.9701173	.0163471	59.34509	0	.9141634	.989996	144	1.976575
Argentina	2011	.	.	.	.	.	.	.	.
Bangladesh	2006	.9705105	.0031824	304.9612	0	.9635859	.9761509	1835	1.961258
Bangladesh	2012	.	.	.	.	.	.	.	.
Barbados	2012	.	.	.	.	.	.	.	.
Belarus	2005	.995512	.002598	383.187	0	.9860311	.9985674	261	1.969095
Belarus	2012	.	.	.	.	.	.	.	.
Belize	2006	.9127005	.0239113	38.17022	0	.8518398	.9500277	74	1.992543
Belize	2011	.9795812	.0070386	139.1725	0	.9599018	.989706	154	1.975488

The equiplot

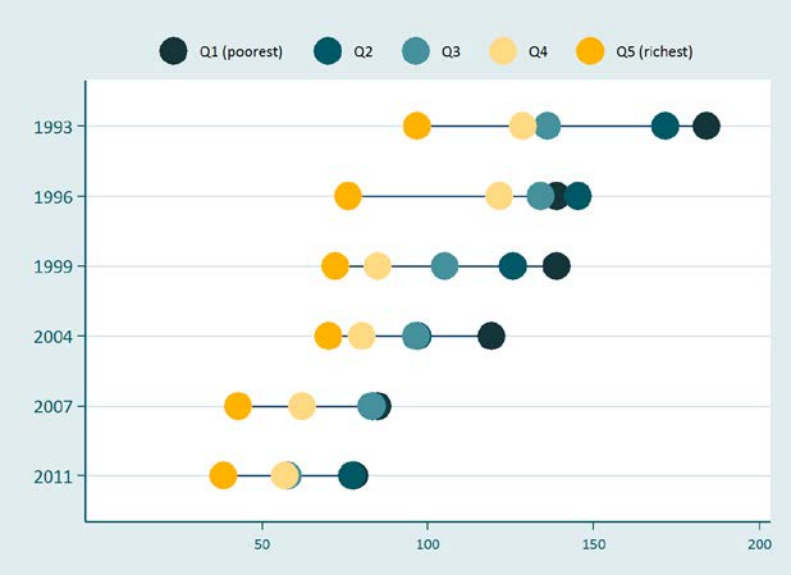
- A graph developed by the ICEH to present inequalities in subgroups

- Can be used to compare countries, interventions, or time trends.

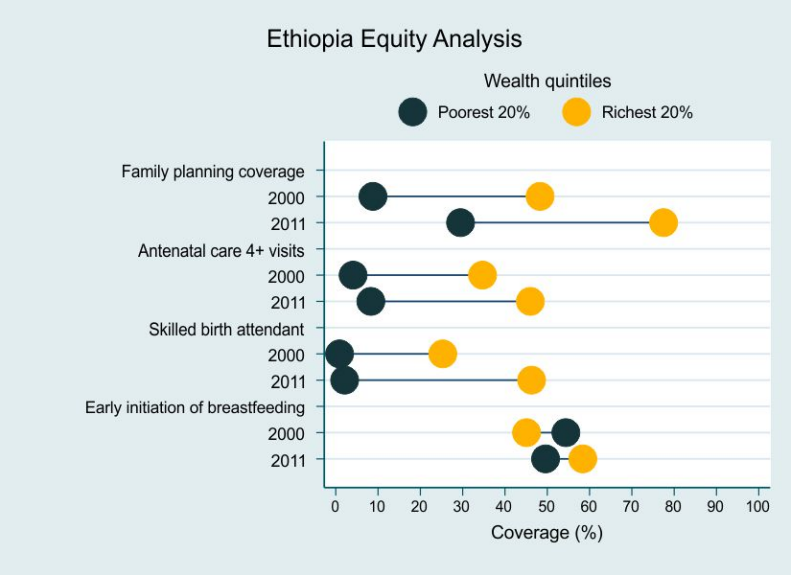
- Freely available at www.equidade.org

The equiplot

U5MR in Bangladesh by wealth quintiles



4 interventions in Ethiopia by wealth



Thank you!

Leonardo Ferreira

lferreira@equidade.org

International Center for Equity in Health

www.equidade.org

Centro Internacional de EQUIDADE em Saúde

UNIVERSIDADE FEDERAL DE PELOTAS RS - BRASIL