Household surveys in practice

Workshop on Measuring Sustainable Agriculture, Food Security and Poverty Alleviation for enhancing Accountability in the Post 2015 Development Agenda.

24-28 November 2014, Bogor Indonesia

Alick Nyasulu
United Nations Statistical Institute for Asia and the Pacific (SIAP)

Outline

1. Household survey design
   - Sampling
   - daily diary vs. recall
   - different recall periods
   - different income/consumption modules
   - Non-response bias
2. Common problems on data processing
3. Common mistakes when calculating poverty measures

Household survey data: sampling

Malawi 1997 and 2004 household survey

Survey year	Mean income per person	Gini index (%)	Headcount (%)
1997/98	399.2	50.3	65.9
2004	483.3	39.0	42.7

Household survey data: sampling

Malawi 1997 and 2004 household survey

more than 4000 households drop from 1997 sample

Survey year	# of obs.	Mean income per person	Gini index (%)	Headcount (%)
1997/98	6586	399.2	50.3	65.9
2004	11280	483.3	39.0	42.7

Household survey data: sampling

- Different sample size/frame will cause comparison problems.
- Vietnam 2010 survey vs. previous rounds
- India National Sample Survey NSS: “thick” and “thin” rounds (30-40% sample size of “thick”)
- Indonesia and other countries
- China 2013 national household survey
  - census frame vs. legal resident registration
  - how to compare with previous rural/urban surveys?

Household survey data: daily diary vs. recall

Example of China SW poverty monitoring survey 1995-1996

1995 survey: one time recall method

1996 survey: daily diary method

1995 mean income per capita: 854.56 Yuan

1996 mean income per capita: 992.74 Yuan

Is there 16% increase in per capita income in one year?

Household survey data: different recall periods

Example of India NSS 55th round

	Recall period	
	all previous rounds	NSS 55
Edu. Medical, clothing, durable goods	last 365 days	last 365 days
food	last 30 days	last 7 days
others	last 30 days	last 30 days

Household survey data: different income/consumption modules

Example of Honduras 1997 and 1999 surveys

	income module 1 Headcount(%)	income module 2 Headcount(%)
1997	24.1	12.0
1999	26.3	10.7

Household survey data: different income/consumption modules

Example of Honduras:

	income module 1 Headcount(%)	income module 2 Headcount(%)
1997	24.1	12.0
1999	26.3	10.7
2003	n.a	13.8

Different income modules?

Summary Report | Refresh | Copy this table only

Country	Year	Data type	Pov.line (PPP$/mo)	Mean ($)	Headcount (%)	Pov. gap (%)	Squared pov. gap	Watts index	Gini index	MLD index	Pop (
Honduras	2009	i	38.00	202.62	17.92	9.35	6.86	0.32	56.95	0.82	
Honduras	2008	i	38.00	208.87	21.36	11.75	8.86	0.42	61.33	1.00	
Honduras	2007	i	38.00	186.77	16.26	6.45	3.65	0.12	56.16	0.61	
Honduras	2006	i	38.00	168.05	22.85	10.84	6.93	0.21	57.56	0.69	
Honduras	2005	i	38.00	155.24	26.44	13.25	8.75	0.27	59.65	0.75	
Honduras	2004	i	38.00	154.48	25.27	11.39	7.07	0.23	58.49	0.71	
Honduras	2003	i	38.00	150.45	26.15	11.97	7.39	0.24	58.71	0.71	
Honduras	2002	i	38.00	144.44	28.16	14.72	10.19	0.37	58.89	0.82	
Honduras	2001	i	38.00	171.57	17.98	8.12	5.05	0.16	54.38	0.60	
Honduras	1999	i	38.00	132.38	25.43	12.52	8.42	0.30	55.36	0.69	
Honduras	1998	i	38.00	141.73	25.52	12.76	8.59	0.30	57.43	0.74	
Honduras	1997	i	38.00	136.61	20.61	7.44	3.76	0.11	52.73	0.51	

Household survey data: different income/consumption modules

Example of Ethiopia 2000 surveys:

	Sample size	Mean exp./p	Headcount (%)	Gini (%)
Welfare Monitoring survey 2000	25016	46.0	81.3	49.0
HH income & exp. survey 2000	16672	92.5	21.9	30.0

Household survey data: different income/consumption modules

Example of Ethiopia 2000 surveys:

Reason: different consumption modules

	Sample size	Mean exp./p	Headcount (%)	Gini (%)
Welfare Monitoring survey 2000	25016	46.0	81.3	49.0
HH income & exp. survey 2000	16672	92.5	21.9	30.0

Nonresponse bias in measuring poverty and inequality

- High nonresponse rates of 10-30% are now common
- LSMS: 0-26% nonresponse (Scott and Steele, 2002)
- UK surveys: 15-30%
- US: 10-20%
- Concerns that the problem might be increasing

Nonresponse bias in measuring poverty and inequality

Compliance is unlikely to be random:

- Rich people have:
  - higher opportunity cost of time
  - more to hide (tax reasons)
  - more likely to be away from home?
  - multiple earners
- Poorest might also not comply:
  - alienated from society?

Common problems on data processing

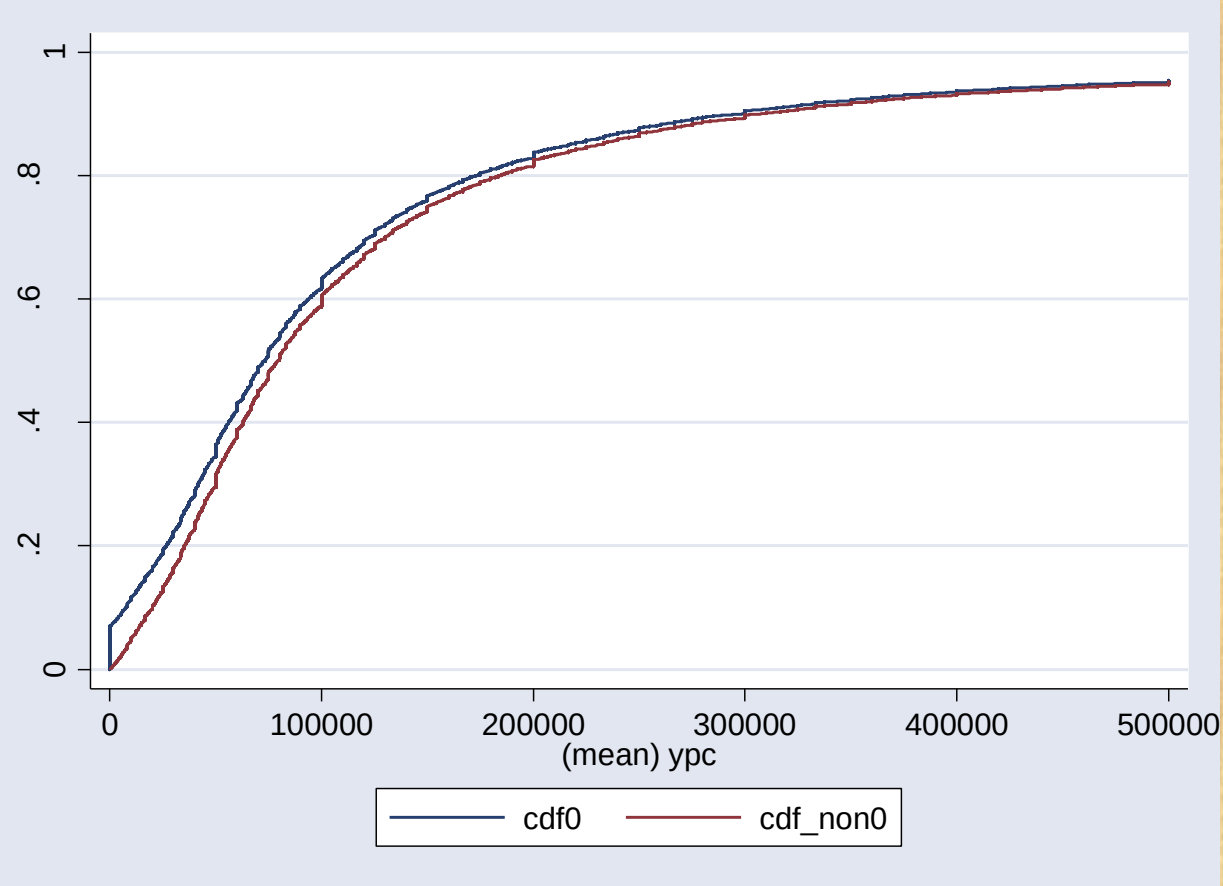
1. Income/consumption aggregates
2. Valuing income in kind
3. Missing value
4. Outliers

Missing, zero and outliers

- Never mix missing value and zero;
- Examples from Latin America Caribbean labor force surveys
- Outliers: check carefully and always keep original records
  - Income by sources
  - Sub components of consumptions

Missing and outliers

Examples from Colombia 2000 survey – 7% are zero income



Summary Report

Refresh | Copy this table only

Country	Year	Data type	Pov.line (PPP$/mo)	Mean ($)	Headcount (%)	Pov. gap (%)	Squared pov. gap	Watts index	Gini index	MLD index	Population (mil.)	Survey year	D c
Colombia	2010	i	38.00	280.97	8.16	3.78	2.62	0.16	55.91	0.70	46.29	2010	
Colombia	2009	i	38.00	264.41	9.67	4.73	3.34	0.22	56.67	0.76	0.00	2009	
Colombia	2008	i	38.00	251.62	11.32	5.77	4.12	0.27	57.23	0.81	45.01	2008	
Colombia	2007	i	38.00	287.43	8.84	4.03	2.67	0.15	58.88	0.75	0.00	2007	
Colombia	2006	i	38.00	258.21	11.04	5.41	3.77	0.23	58.66	0.81	0.00	2006	
Colombia	2005	i	38.00	207.25	12.71	6.33	4.54	0.30	56.12	0.80	43.04	2005	
Colombia	2004	i	38.00	174.25	18.98	10.13	7.50	0.52	58.29	1.02	0.00	2004	
Colombia	2003	i	38.00	166.94	19.61	10.90	8.35	0.60	57.86	1.07	0.00	2003	
Colombia	2002	i	38.00	182.93	20.26	11.17	8.53	0.62	60.68	1.16	41.08	2002	
Colombia	2001	i	38.00	172.64	19.18	10.86	8.40	0.60	58.01	1.09	0.00	2001	
Colombia	2000	i	38.00	172.64	17.85	11.46	9.71	0.86	58.68	1.36	0.00	2000	
Colombia	1999	i	38.00	203.61	16.18	9.47	7.42	0.55	58.74	1.08	39.10	1999	
Colombia	1996	i	38.00	222.27	13.00	7.62	6.06	0.42	56.94	0.93	37.11	1996	
Colombia	1992	i	38.00	239.20	6.29	4.01	3.42	0.27	51.45	0.70	0.00	1992	

- Thank You