Appraising income inequality data bases in LatinAmerica

François Bourguignon

Paris School of Economics

UNU-WIDER, Helsinki, September 2014

Income inequality databases for LAC countries

- **CEPALSTAT**: Statistical Office of the UN Economic Commission for Latin America and the Caribbe
  - Publishes their own inequality measures on the basis of household survey microdata made available to them by member coutries
  - No up-to-date methodology document available (but work in progress)
  - Methodology based on a 1987 paper by Oscar Altimir, with a strong advocacy in favor of ajusting the data for no-reporting or under-reporting
  - Poverty headcount based on Cepalstat poverty lines, themselves relying on updated on national minimum diet cost estimates and Orshansky coefficients
  - Poverty estimates differ from offical national ones: Povcal poverty headcount available on line

Income inequality databases for LAC countries

- **SEDLAC**: Socioeconomic data base for Latin America and the Caribbe, joint venture between CEDLAS at Universidad de la Plata (Argentina) and the World Bank poverty and gender group for Latin America and the Caribbe
  - Publishes their own harmonized inequality measures on the basis of household survey microdata made available to them by MECOVI countries
  - Well-documented fully up-to-date methodology, reasonably close to best practice (and consistent with World Bank's Povcal)
  - Database regularly updated
  - Poverty estimates are those from Povcal – same harmonized data used plus their own estimates with 2.5 and 4 ppp 2005 USD a day poverty lines

Other data bases covering LAC coutries among others

- Primary data bases
  - World Bank Povcal/WYD
  - LIS [Brazil (3), Colombia (3), Mexico(11)]
  - OECD (Mexico, Chile)

- Secondary data base: ATG, WIID, SWIID, UTIP, ..

Questions

1. How close are the inequality (poverty) measures reported by CEPALSTAT and SEDLAC ?

2. Differences in the treatment of missing data, under-reporting and the National Account-Household Survey gap

3. Other methodological issues

1. How close are Cepalstat and Sedlac? Levels of inequality

Figure 1. Gini coefficient (2007-2009 mean) in the Cepalstat and Sedlac data base
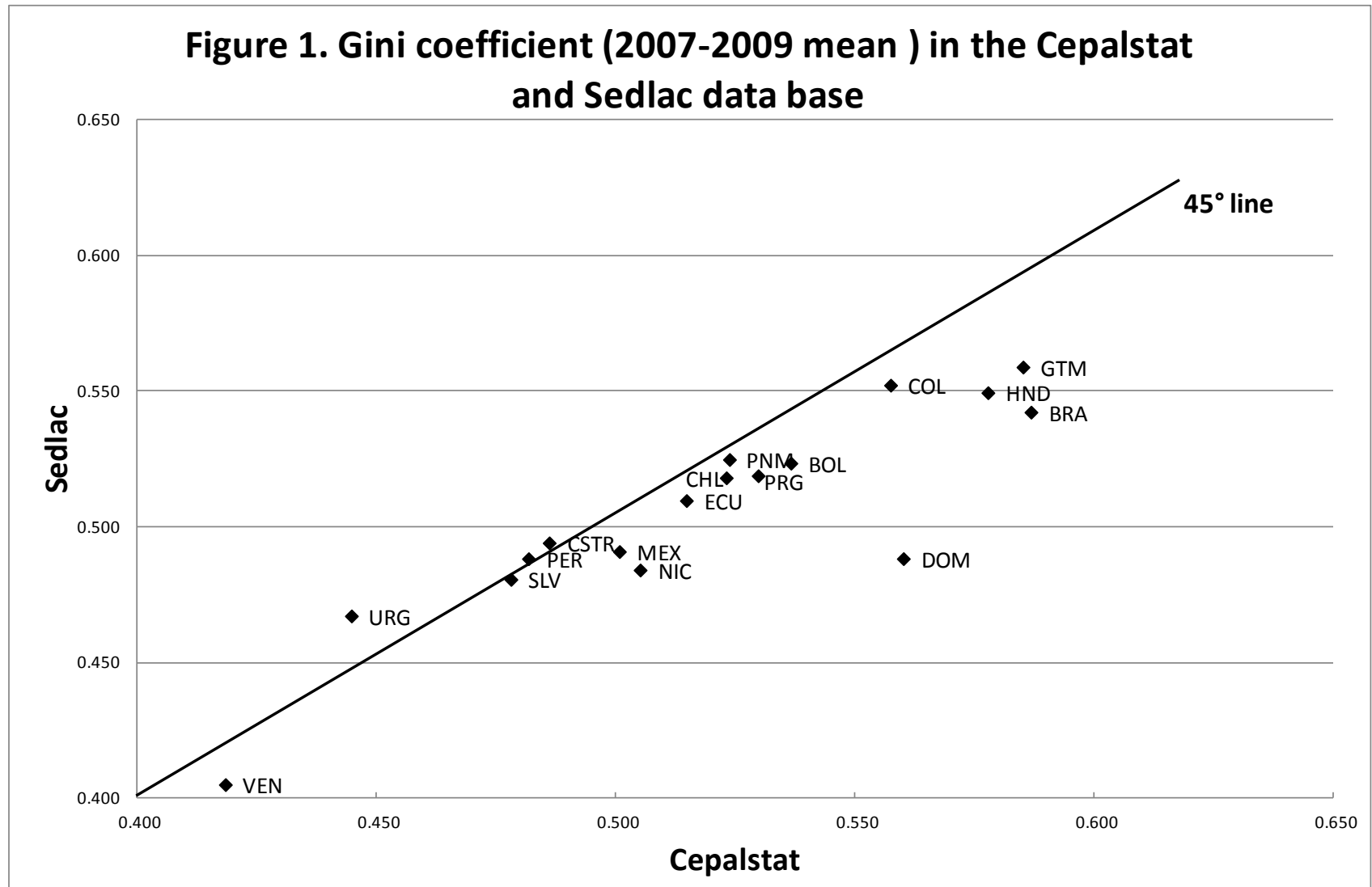
1. How close are Cepalstat and Sedlac? Changes in inequality

Figure 2. Comparing Gini time series from various sources: selected countries
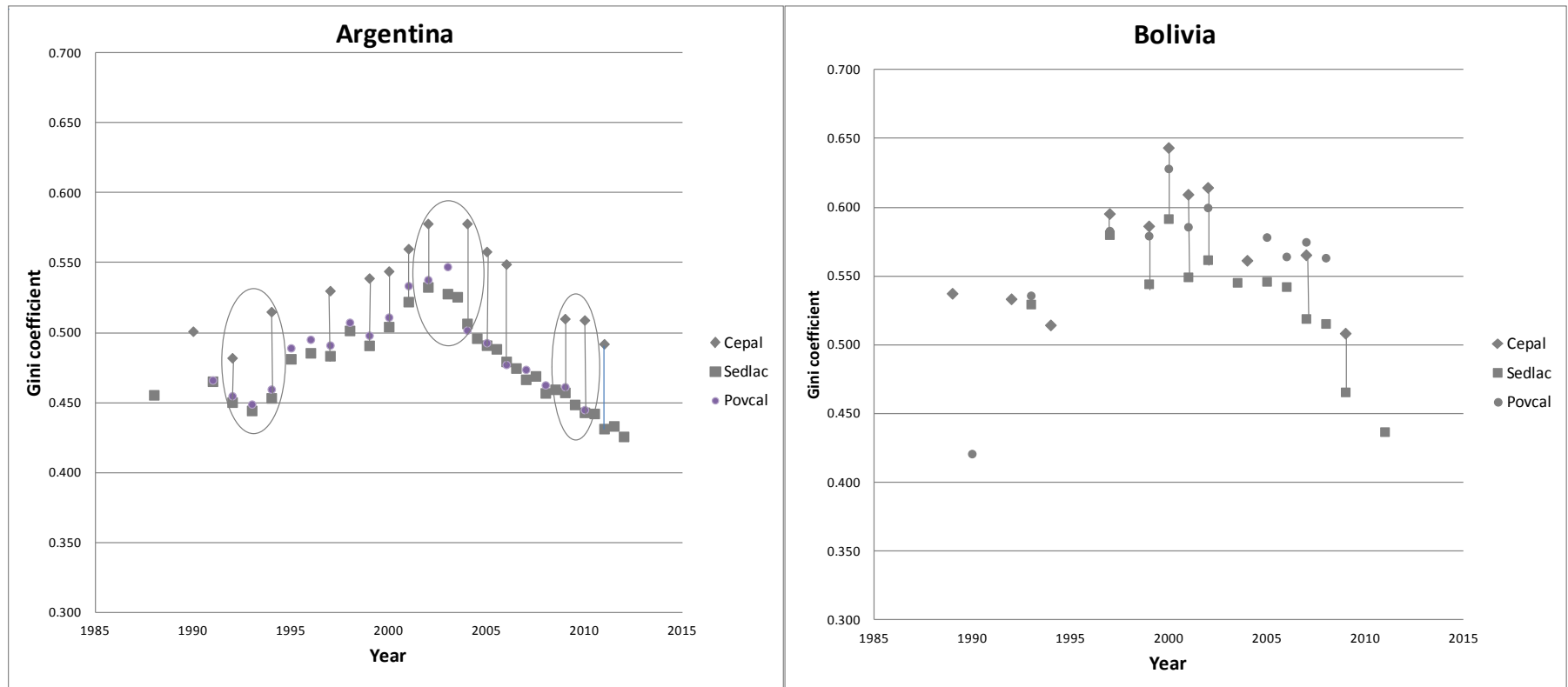
.. How close ... ct'd

Figure 2. (ct'd)
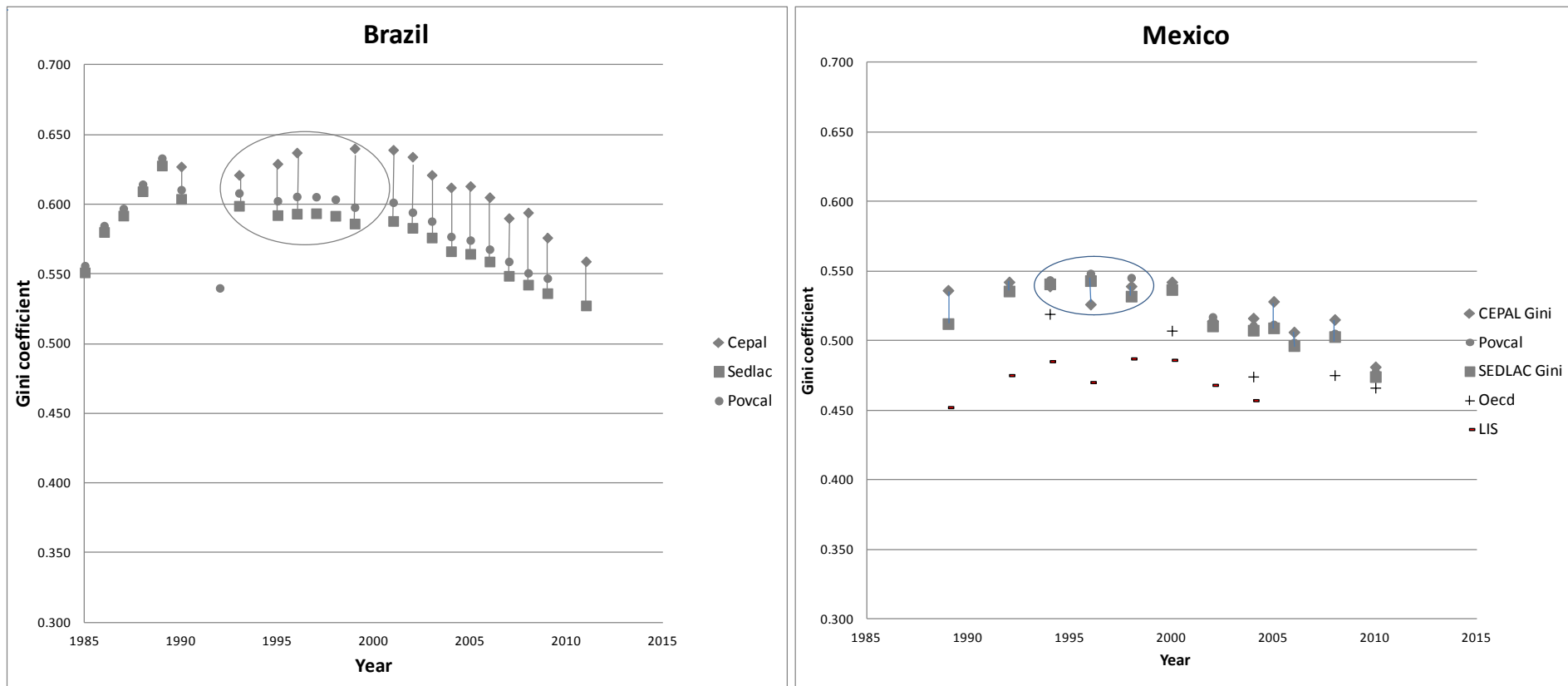
1. How close are Cepalstat and Sedlac? Changes in poverty

Figure 3c. Poverty headcount as reported by CEPALSTAT and World Bank: Costa-Rica, 1980-2012

Figure 3d. Poverty headcount as reported by CEPALSTAT and World Bank: Mexico, 1980-2012
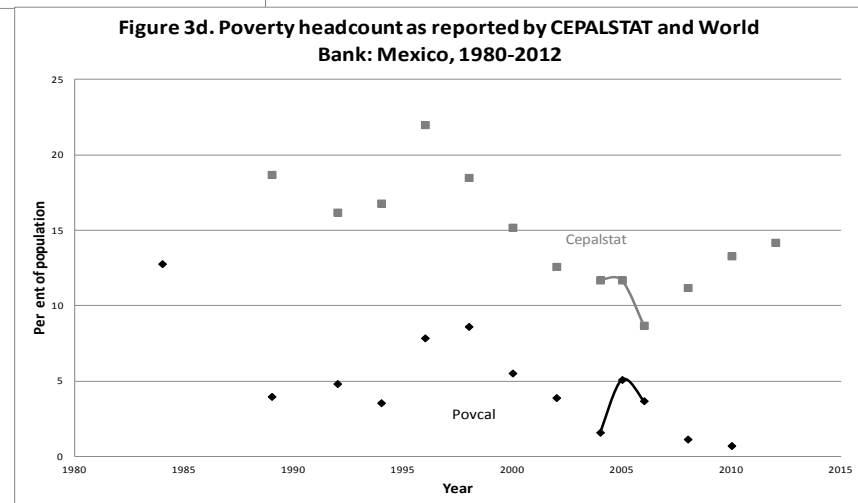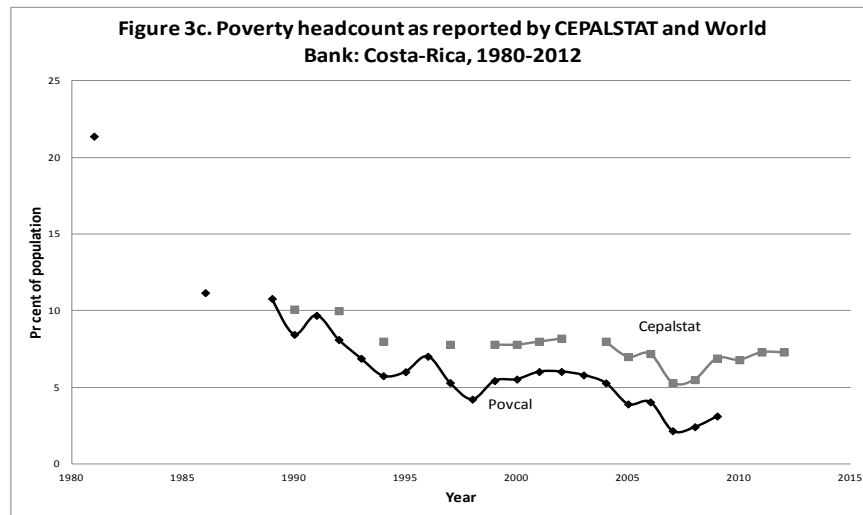
How close ... ct'd

Figure 3a. Poverty headcount as reported by CEPALSTAT and World Bank: Brazil, 1980-2012

Figure 3b. Poverty headcount as reported by CEPALSTAT and World Bank: Colombia, 1980-2012
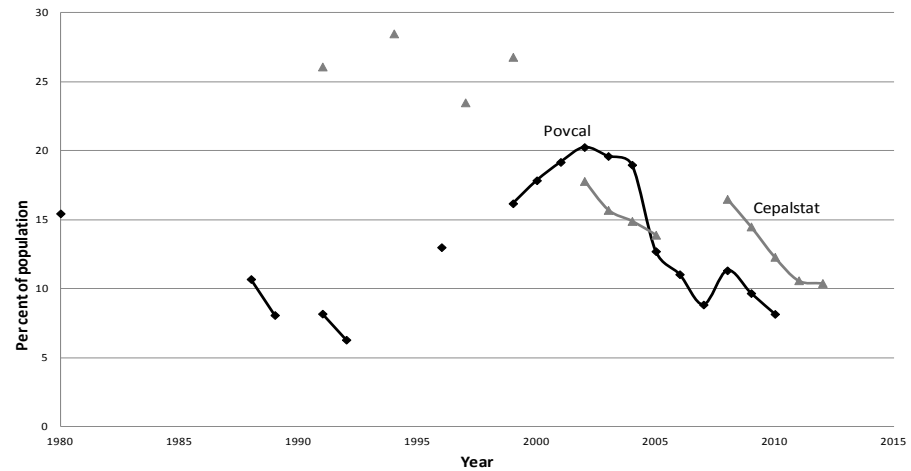
Overall evaluation

- Frequent sizable differences in levels

- Time evolution generally consistent over long periods, but not infrequent divergences

- Sedlac closer to other sources, as well as to independent research work

- Difficult to evaluate updating work because no archive of website at previous dates are available

2. Adjustments for missing data and under-reporting

- Systematic imputation for missing data (matching, hot deck) in Cepalstat
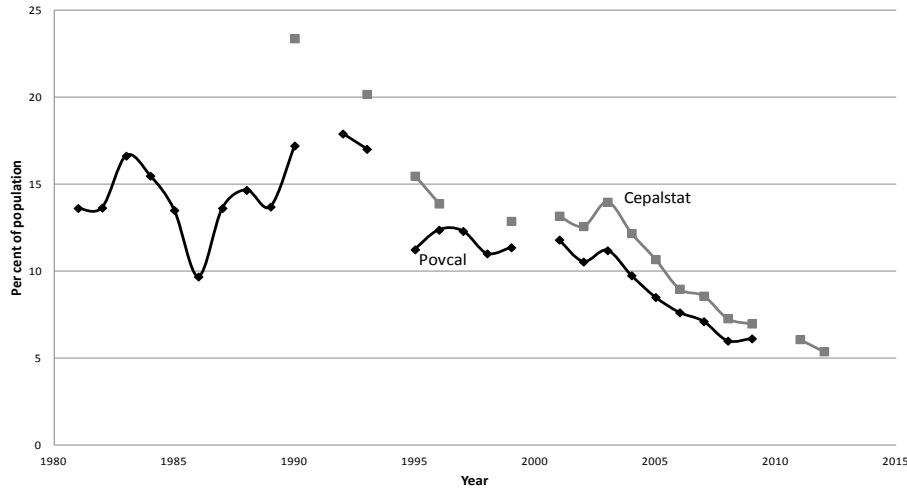
- No imputation in Sedlac, except for imputed rents. Observations with major missing data are dropped (except for poverty).

- Major correction for under-reporting (in comparison with NA) in Cepalstat: probably the main source of discrepancy between the two data bases.
  - All income sources adjusted uniformly by a scale factor equal to NA figure per household/Household Survey mean by household
  - Special treatment for property income (adjusted on the top quintile) and imputed rents

NA/HS discrepancy: case of Chile

Table 1. Chile: National Account to Household Survey adjustment factors by income source

	NA/HS income ratio (all households)							Structure of total income in 2011 (%)	
Year	1996	1998	2000	2003	2006	2009	2011	NA	HS
Wage and salaries	1.00	1.02	0.97	1.01	1.03	1.12	1.11	55.3	60.3
Self-employment	2.07	1.98	1.85	2.00	2.01	2.09	2.09	27.4	15.9
Pensions and benefits	1.42	1.37	1.49	1.16	1.15	1.00	1.00	7.4	9.0
Property	2.74	2.75	2.43	1.84	2.14	1.94	3.51	4.8	1.7
(Top quintile adjustment) [a]	*1.06*	*1.07*	*1.05*	*1.03*	*1.04*	*1.03*	*1.06*		
Imputed rfents	0.46	0.44	0.46	0.44	0.45	0.43	0.48	5.1	13.1
Total	1.21	1.19	1.15	1.16	1.18	1.22	1.21	100.0	100.0
Source: La medicion de los ingresos en la encuesta CASEN, Cepal, various years									
[a] Adjustment factor when the whole discrepancy between NA ad HS is imputed to the top quintile proportionally to total market income									

NA/HS discrepancy: case of Chile

Table 1. Chile: National Account to Household Survey adjustment factors by income source	NA/HS income ratio (all households)							Structure of total income in 2011 (%)	
Year	1996	1998	2000	2003	2006	2009	2011	NA	HS
Wage and salaries	1.00	1.02	0.97	1.01	1.03	1.12	1.11	55.3	60.3
Self-employment	2.07	1.98	1.85	2.00	2.01	2.09	2.09	27.4	15.9
Pensions and benefits	1.42	1.37	1.49	1.16	1.15	1.00	1.00	7.4	9.0
Property	2.74	2.75	2.43	1.84	2.14	1.94	3.51	4.8	1.7
(Top quintile adjustment)[a]	1.06	1.07	1.05	1.03	1.04	1.03	1.06		
Imputed rfents	0.46	0.44	0.46	0.44	0.45	0.43	0.48	5.1	13.1
Total	1.21	1.19	1.15	1.16	1.18	1.22	1.21	100.0	100.0
Source: La medicion de los ingresos en la encuesta CASEN, Cepal, various years									
[a] Adjustment factor when the whole discrepancy between NA ad HS is imputed to the top quintile proportionally to total market income									

13

| Table 2. Inequality effect of adjusting the NA/HS property income gap on the top quintile : rough calculation on Chile and Brazil | | | | | | | |

		Aggregate income by source (%)					Quintile shares [b] (%)	
		Household survey (HS) [c]	NA-HS gap as % of HS total income	National Accounts (NA)			Household survey	NA-Adjusted
Chile (2009)								
Labor income		75.7	22.2	84.4		0-20%	4.5	4.4
Property income		2.5	1.9	3.9		20-40%	8.2	8.0
Transfers		8.5	0.0	7.0		40-60%	11.9	11.7
Imputed rents		13.3	-6.3	4.6		60-80%	18.7	18.3
Total		100	17.8	100		80-100%	56.8	57.6
						Gini [d]	46.0	46.7
Chile (2011)								
Labor income		76.3	19.9	82.7		0-20%	4.8	4.6
Property income		1.7	3.4	4.8		20-40%	8.5	8.2
Transfers		9.0	0.0	7.4		40-60%	12.2	11.8
Imputed rents		13.1	-5.7	5.1		60-80%	19.1	18.4
Total		100	17.6	100		80-100%	55.5	57.0
						Gini [d]	44.8	46.0
Brazil (2005)								
Labor income		76.2	-4.1	62.6		0-20%	3.0	2.8
Property income		3.6	10.1	11.9		20-40%	6.5	6.1
Transfers		20.2	9.2	25.5		40-60%	11.0	10.3
						60-80%	18.6	17.4
Total		100.0	15.2	100.0		80-100%	60.9	63.4
						Gini [d]	51.2	53.0

[a] Adjustment consists of allocating the NA-HS property income gap to top quintile.

[b] For Brazil, the household survey quintile share are from Sedlac. For Chile the adjustment goes in the opposite direction. As Sedlac gives NA-adjusted quintile shares, the correction procedure estimates the HS quintile share which would have led to the Sedlac shares with the procedure described in [a].

The effect of NA/HS adjustment: an illustration

NA/HS consistency checks would be valuable

Table 3. Ratio of the mean income in household survey to the mean household final consumption expenditure per capita in National Accounts																
Country	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Bolivia	125.6		107.3	100.7	108.0	108.3			116.9	118.8	117.2	121.4	126.8		126.2	126.6
Brazil				83.8	84.1	85.5	82.0	82.0	83.9	86.1	83.2	83.9	81.7		77.9	81.3
Colombia					49.8	66.7	60.5	63.0	65.5			65.2	68.4	70.8	70.7	66.9
Costa-Rica					80.3	79.5	80.2	74.8	75.6	75.4	80.0	80.2	89.7			
Dominican Republic				92.4	88.1	82.2	69.4	57.6	59.6	56.7	59.3	48.2	54.6	49.4	46.5	49.8
Ecuador				46.6			65.7	86.6	69.9	74.8	75.0	66.3	66.2	70.0	69.2	70.9
El Salvador								57.1	55.6	53.2	53.6	49.1	55.1	52.3	50.5	51.4
Honduras					112.8	93.1	95.1	95.5	90.7	91.7	98.4	102.7	103.4	100.7	98.2	
Mexico		43.8		49.0		47.5		43.3	43.0	42.9		43.2		42.3		43.8
Paraguay	143.8		134.0		131.9	122.4	125.5	115.0	117.6	108.2	106.9	98.0	109.8	105.7	105.6	
Peru	73.7	81.0	81.4	70.3	67.4	75.8	71.5	72.2	67.3	71.5	74.8	72.7	77.2	77.7	76.1	76.9
Uruguay	75.6	80.3		82.7	82.2	75.6	71.1	69.8	70.3	68.6	68.5	70.9	81.5	74.0	73.0	69.6
Source: Sedlac and WDI, author's calculation																

16

NA/HS consistency checks would be valuable

Table 3. Ratio of the mean income in household survey to the mean household final consumption expenditure per capita in National Accounts																
Country	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Bolivia	125.6		107.3	100.7	108.0	108.3			116.9	118.8	117.2	121.4	126.8		126.2	126.6
Brazil				83.8	84.1	85.5	82.0	82.0	83.9	86.1	83.2	83.9	81.7		77.9	81.3
Colombia					49.8	66.7	60.5	63.0	65.5			65.2	68.4	70.8	70.7	66.9
Costa-Rica					80.3	79.5	80.2	74.8	75.6	75.4	80.0	80.2	89.7			
Dominican Republic				92.4	88.1	82.2	69.4	57.6	59.6	56.7	59.3	48.2	54.6	49.4	46.5	49.8
Ecuador				46.6			65.7	86.6	69.9	74.8	75.0	66.3	66.2	70.0	69.2	70.9
El Salvador								57.1	55.6	53.2	53.6	49.1	55.1	52.3	50.5	51.4
Honduras					112.8	93.1	95.1	95.5	90.7	91.7	98.4	102.7	103.4	100.7	98.2	
Mexico		43.8		49.0		47.5		43.3	43.0	42.9		43.2		42.3		43.8
Paraguay	143.8		134.0		131.9	122.4	125.5	115.0	117.6	108.2	106.9	98.0	109.8	105.7	105.6	
Peru	73.7	81.0	81.4	70.3	67.4	75.8	71.5	72.2	67.3	71.5	74.8	72.7	77.2	77.7	76.1	76.9
Uruguay	75.6	80.3		82.7	82.2	75.6	71.1	69.8	70.3	68.6	68.5	70.9	81.5	74.0	73.0	69.6
Source: Sedlac and WDI, author's calculation																

Other issues

- Non-response
- Eqivalence scales
- Imputed rents
- Spatial differences in the cost of living
- Multiple poverty lines